

Technische Universität Dortmund

Fakultät Erziehungswissenschaft, Psychologie und Bildungsforschung

**Wissen über Diversität im Kontext schulischer
Inklusion: Evaluation und Einsatz eines Tests zum
deklarativen Faktenwissen von (angehenden)
Lehrkräften**

Dissertation zur Erlangung des akademischen Grades
Doktorin der Philosophie (Dr. phil.)

vorgelegt von
M. Sc. Psychologin Lisa Tometten

Erstgutachterin: Prof. Dr. Ricarda Steinmayr

Zweitgutachterin: Prof. Dr. Anke Heyder

Disputation: 14. August 2023

Das dieser Arbeit zugrunde liegende Vorhaben (v. a. Studie 1 & 2) wurde im Rahmen der gemeinsamen „Qualitätsoffensive Lehrerbildung“ von Bund und Ländern mit Mitteln des Bundesministeriums für Bildung und Forschung unter dem Förderkennzeichen 01JA1630 gefördert. Das Teilprojekt zum Diversitätswissen von (angehenden) Lehrkräften im Institut für Psychologie der TU Dortmund wurde geleitet von Prof. Dr. Ricarda Steinmayr und begleitet von Prof. Dr. Anke Heyder.

Danksagung

Mein größter Dank gilt Ricarda Steinmayr und Anke Heyder, die mir die Möglichkeit gegeben haben, meine Doktorarbeit im Rahmen des Projekts „DoProfil – Dortmunder Profil für inklusionsorientierte Lehrer/-innenbildung“ zu schreiben. Ihr habt daran geglaubt, dass ich dieser Aufgabe gewachsen bin und habt mich auch über die Zeit an der TU Dortmund hinaus dabei begleitet. Danke! Für mich war diese Zeit sowohl beruflich als auch privat sehr bereichernd und sie hat mir viele neue Türen und Perspektiven geöffnet.

Aus dieser Zeit nehme ich außerdem viele Begegnungen, Gespräche und Treffen mit Kolleg:innen mit, die meinen (Arbeits-)Alltag und die Zeit in Dortmund sehr bereichern haben. Heike Knieriem, Linda Wirthwein, Anne Weidinger, Sebastian Bergold, Josi Michels, Patrick Paschke, Katharina Abad Borger, Laura Modler, Christian Blötner und Gabriel Brandenburg, es hat mir immer viel Spaß mit euch gemacht!

Danke an alle (angehenden) Lehrkräfte und Schüler:innen, die an den Erhebungen teilgenommen und mit ihrer Teilnahme diese Arbeit überhaupt erst möglich gemacht haben. Ein Dank geht auch an alle studentischen Hilfskräfte für ihre tatkräftige Unterstützung. Die Daten von Studie 3 wurden vor allem von den Studierenden Caroline Kirchner and Mark Fischer erhoben. Danke für euer Engagement in der Datenerhebung!

Ich bin meiner Familie sehr dankbar, die es mir ermöglicht hat, diese Arbeit abzuschließen. Ihr habt mir Zeitfenster für das Schreiben verschafft, mich motiviert und mir schriftlich und mündlich Rückmeldungen zu meinem Text gegeben. Danke, Mathis und Finja Tometten, Dorothee Schleiß-Grewe und Detlef Grewe, Ute und Dieter Tometten!

Schlussendlich danke ich allen, die mich und meine Familie auf unserem Weg unterstützt und dafür gesorgt haben, dass wir jetzt unseren Traum vom Hof verwirklichen und leben dürfen. Das ist ein großes Privileg, für das ich sehr dankbar bin!

Inhaltsverzeichnis

Zusammenfassung	9
Abstract	13
Einleitung	17
1 Schulische Inklusion	21
1.1 Definition schulischer Inklusion.....	21
1.1.1 Inklusionsmodelle, Heterogenität und Diversität	22
1.1.2 Praktische Umsetzung an Schulen	25
1.1.3 Fazit.....	27
1.2 Empirische Befunde	28
1.2.1 Effekte inklusiver vs. nicht-inklusive Beschulung auf Schüler:innen.....	28
1.2.2 Lehrkräfte als Akteur:innen schulischer Inklusion	30
2 Wissen als Teil professioneller Kompetenz von Lehrkräften	35
2.1 Pädagogisch-psychologisches Wissen und Wissen über Diversität	37
2.2 Empirische Zusammenhänge zu Lerngelegenheiten, Aspekten professioneller Kompetenz, der Performanz und dem Unterrichtserfolg von Lehrkräften.....	39
2.3 Bisherige Tests zur Erfassung von Wissen über Diversität.....	45
2.3.1 Tests zum pädagogisch-psychologischen Wissen mit Diversität als Facette	47
2.3.2 Tests zu einzelnen Diversitätsbereichen	51
3 Der vorliegende Wissenstest	55
3.1 Bedeutung der Diversitätsbereiche	55
3.1.1 Sonderpädagogischer Förderbedarf emotionale-soziale Entwicklung & Lernen .	55
3.1.2 Teilleistungsstörungen	59
3.1.3 Geschlecht.....	62
3.1.4 Kognitive Hochbegabung	64
3.2 Testaufbau und Items.....	67
3.3 Kriterien zur Evaluation dieses Tests	69
3.3.1 Klassische Testtheorie	70
3.3.2 Hauptgütekriterien	71
3.3.3 Nebengütekriterien.....	75
4 Ziele, Fragestellungen und Studienüberblick	79
Literaturverzeichnis I	83

5	Studien	101
5.1	Studie 1: Die Erfassung von Wissen über verschiedene Diversitätsbereiche im Kontext einer inklusionsorientierten Lehrkräftebildung	101
5.2	Studie 2: Teachers' knowledge about diversity: Correlations with attitudes, self-efficacy, and instructions	129
5.3	Studie 3: Links between teachers' knowledge about special educational needs and students' social participation and academic achievement in mainstream classes ..	169
6	Gesamtdiskussion	217
6.1	Zusammenfassung der Studienergebnisse	218
6.2	Interpretation der Ergebnisse zum Wissen und zu den Fehlannahmen über Diversität	220
6.2.1	Einordnung in die professionelle Kompetenz.....	221
6.2.2	Erfassung und Struktur	229
6.3	Stärken und Grenzen der Arbeit	231
6.4	Implikationen.....	236
6.4.1	Forschung.....	236
6.4.2	Praxis.....	243
6.5	Fazit	247
	Literaturverzeichnis II	251
7	Anhang	261
7.1	Liste der Einzelarbeiten	261

Zusammenfassung

Seit der Ratifizierung der UN-Behindertenrechtskonvention der Vereinten Nationen in Deutschland im Jahr 2009 (Beauftragter der Bundesregierung für die Belange von Menschen mit Behinderungen, 2018) ist das Bemühen um Inklusion im Bildungssystem nicht nur in Bezug auf Menschen mit Behinderungen gestiegen. Ziel ist vielmehr die Förderung und Anerkennung aller Schüler:innen, unabhängig von ihren individuellen Hintergründen und Merkmalen (Grosche, 2015). Lehrkräfte gelten als wichtige Akteur:innen in der Umsetzung schulischer Inklusion, und die Bedeutung ihrer professionellen Kompetenz, und damit auch ihres Professionswissens (vgl. COACTIV-Kompetenzmodell; Baumert & Kunter, 2013a), wird sowohl innerhalb als auch außerhalb inklusiver Settings untersucht. Wer diverse Klassen unterrichtet, sollte eine solide Wissensbasis für die Diversität der Schüler:innen mitbringen – das erscheint plausibel. So weist auch die Forschung zum pädagogisch-psychologischen Wissen mit (in)direktem Bezug zu Diversität auf die praktische Bedeutung von Lerngelegenheiten für (angehende) Lehrkräfte (z. B. Sciutto et al., 2000; Sousa et al., 2017; Yin et al., 2019) und Zusammenhänge zum Berufserfolg von Lehrkräften hin (z. B. König et al., 2017; Lauer mann & König, 2016; Voss et al., 2011, 2022). Bisherige Messinstrumente zum Wissen von (angehenden) Lehrkräften können wichtige Erkenntnisse für ihre Aus- und Weiterbildung liefern, haben aber zum Teil nur einen indirekten Bezug zu Diversität (Kunina-Habenicht et al., 2020) oder bilden nur einen isolierten Teil von Diversität ab (Ginsburg et al., 2021) und sind wenig ökonomisch (König et al., 2017) bzw. reliabel in Bezug auf die spezielle Diversitätsskala (Hohenstein et al., 2017), woraus sich zu schließende Forschungslücken ergeben. Um bestehenden Nachteilen zu begegnen, wurde in dieser Arbeit ein neuer Test vorgestellt, evaluiert und eingesetzt.

In *Studie 1* wurde dieser Test zum deklarativen Faktenwissen in einer Lang- (DiWi) und Kurzform (DiWi-K) mit den Subskalen sonderpädagogischer Förderbedarf (SPF) (1)

emotionale und soziale Entwicklung (ESE) und (2) Lernen, (3) Teilleistungsstörungen, (4) Geschlecht und (5) kognitive Hochbegabung vorgestellt und in drei unabhängigen Lehramtsstudierendenstichproben ($N_1 = 395$ Bachelor-Studierende, $N_2 = 462$ Bachelor- und Master-Studierende, $N_3 = 113$ Bachelor-Studierende) evaluiert. Der Reliabilitätsschätzer McDonalds Omega und die Retest-Reliabilität über einen Zeitraum von 3 Monaten ergaben zufriedenstellende Werte. Die 5-faktorielle Struktur des DiWi wurde durch die Daten in der ersten Studie gestützt, und das objektive Wissen hing mit weiteren wissensnahen Konstrukten zusammen (Abiturnote, selbsteingeschätztes Wissen, behandelte Diversitätsbereiche im Studium, Studienfach [sonderpädagogische Förderung vs. andere]). In der zweiten und dritten Studie bestätigten sich diese Ergebnisse für den DiWi-K, und es kamen die konvergenten Konstrukte Vorerfahrungen, Einstellungen, Selbstwirksamkeitserwartungen bezüglich schulischer Inklusion und pädagogisches Unterrichtswissen hinzu. Damit ist ein erstes nomologisches Netz (Hartig et al., 2020) rund um das Diversitätswissen gezeichnet, das Hinweise auf die Validität der Testwertinterpretationen gibt.

Da die Ergebnisse aus Studierendenstichproben nicht automatisch auf Lehrkräfte generalisierbar sind, wurde der DiWi in *Studie 2* in einer Stichprobe aus Lehrkräften verschiedener Schulformen ($N = 139$) weiter evaluiert. Außerdem wurden die Fehlannahmen (= was die Lehrkräfte fälschlicherweise zu wissen glaubten) ausgewertet. Auch hier hing das objektive Wissen mit den Einstellungen zu Inklusion, den Selbstwirksamkeitserwartungen im Unterrichten in einem inklusiven Schulsystem und dem Studienfach (sonderpädagogische Förderung vs. andere) zusammen. Hinzu kam eine positive Korrelation mit den selbstberichteten lernorientierten Unterrichtsinstruktionen. In Bezug auf den Besuch einer Fortbildung zu Inklusion zeigten sich keine eindeutigen Unterschiede im Wissen oder den Fehlannahmen. Die Lehrkräfte hatten ein höheres Wissen und mehr Fehlannahmen als die hinzugezogenen studentischen Vergleichsstichproben, die aus Studie 1 generiert wurden.

In *Studie 3* wurde das Wissen von Grundschullehrkräften mit Variablen auf Ebene der

Schüler:innen in Verbindung gebracht. Mit dem Ziel, inklusive Prozesse – also die Förderung und Anerkennung aller Schüler:innen (Grosche, 2015) – unterstützen zu können, wurden die subjektiven Angaben zu der sozialen Partizipation und die objektiven Daten zu den Schulleistungen von Grundschüler:innen ($N_1 = 421$, $N_2 = 715$) und das Wissen ihrer 23 bzw. 39 Klassenlehrer:innen in 25 bzw. 40 Klassen mit Schüler:innen mit und ohne SPF ESE bzw. Lernen erfasst. Die Multilevel-Analysen ergaben geringere Leistungen der Schüler:innen mit SPF im Vergleich zu ihren Peers ohne SPF. Schüler:innen mit SPF ESE fühlten sich schlechter sozial integriert. Das Wissen der Lehrkräfte über den SPF ESE hing positiv mit der sozialen Partizipation und den Leistungen aller Schüler:innen (unabhängig von ihrem SPF) in Klassen mit Kindern mit und ohne SPF ESE zusammen. Ein entsprechender Zusammenhang zum Wissen über den SPF Lernen in Klassen mit Kindern mit und ohne SPF Lernen zeigte sich nicht.

Zusammenfassend weisen die drei Studien auf die Eignung des Tests für den Einsatz in der Erforschung und Umsetzung der Aus- und Fortbildung von Lehrkräften hin. Die genutzten Reliabilitätsschätzer legen eine adäquate Messgenauigkeit nahe und die inhaltliche Unterteilung der Skalen erwies sich als passend zu den empirischen Daten. Ein breites nomologisches Netz rund um das Diversitätswissen gibt Hinweise auf die Validität und ordnet das Konstrukt plausibel in das COACTIV-Kompetenzmodell (Baumert & Kunter, 2013a; Kunter, Kleickmann et al., 2013) ein. Die Ergebnisse schließen sich damit an vorherige Erkenntnisse zu Tests mit unterschiedlichen Schwerpunkten an (z. B. König et al., 2017; Pflanzl et al., 2013). Die Gesamtheit der mit dem hier gemessenen Wissen zusammenhängenden Konstrukte deutet darauf hin, dass das Diversitätswissen ein ökonomisch erfassbarer Teil des Professionswissen ist, der im Zuge schulischer Inklusion eine Relevanz für Lehrkräfte und Schüler:innen aufweist. Abschließend werden Fragestellungen für zukünftige Forschung, die sich für die Evaluation und Nutzung des Tests ergeben, und Implikationen für den praktischen Einsatz des Tests diskutiert.

Abstract

Efforts to achieve inclusion in education have increased since the ratification of the UN Convention on the Rights of Persons with Disabilities in Germany in 2009 (Beauftragter der Bundesregierung für die Belange von Menschen mit Behinderungen [Federal Government Commissioner for Disability Issues], 2018). The aim of inclusion, however, is not only to place students with disabilities in mainstream classes, but rather to encourage and appreciate all students, regardless of their individual backgrounds and characteristics (Grosche, 2015). Teachers are considered as important actors in the implementation of inclusive education. The importance of their professional competence, and thus their professional knowledge (cf. COACTIV competence model; Baumert & Kunter, 2013a), is explored both within and outside inclusive settings. Those who teach diverse classes should have a solid knowledge base of student diversity - this seems plausible. Thus, research on pre- and in-service teachers' pedagogical-psychological knowledge (in)directly related to diversity highlights the practical importance of learning opportunities (e.g., Scitutto et al., 2000; Sousa et al., 2017; Yin et al., 2019) and links to teachers' professional success (e.g., König et al., 2017; Lauer mann & König, 2016; Voss et al., 2011, 2022). Current instruments on what pre- and in-service teachers know can provide important findings for their professional qualification and further training. Some tests are only indirectly related to diversity (Kunina-Habenicht et al., 2020), depict an isolated part of diversity (Ginsburg et al., 2021), and are not very economical (König et al., 2017) or reliable with regard to the specific diversity-related subscale (Hohenstein et al., 2017), resulting in research gaps that need to be filled. In order to overcome existing disadvantages, a new test was introduced, evaluated, and implemented in this thesis.

In *Study 1*, this test of declarative factual knowledge was therefore introduced in a long (DiWi) and short form (DiWi-K) with the subscales special educational needs (SEN) due

to (1) emotional and behavioural (EBD) and (2) learning difficulties (LD), (3) partial achievement disorders, (4) gender, and (5) cognitive giftedness and evaluated in three independent student teacher samples ($N_1 = 395$ Bachelor students, $N_2 = 462$ Bachelor and Master students, $N_3 = 113$ Bachelor students). McDonald's Omega as an estimator of reliability and retest reliability over a period of 3 months resulted in satisfactory values. The 5-factorial structure of the DiWi was supported by the data in the first study and objective knowledge was linked to other knowledge-related constructs (high school grade, self-assessed knowledge, diversity areas covered in university, teaching profession [special education vs. others]). In the second and third study, these results were confirmed for the DiWi-K and the convergent constructs of prior experience, attitudes, self-efficacy regarding inclusive teaching and pedagogical teaching knowledge were added. This draws a first nomological net (Hartig et al., 2020) around knowledge about diversity, which provides evidence for the validity of the test score interpretations.

Since the results from student samples cannot automatically be generalized to teachers, the DiWi was further evaluated in *Study 2* in a sample of teachers from different school types ($N = 139$). Also, teachers' misconceptions (= what teachers mistakenly believe they know) were evaluated. Again, objective knowledge was related to attitudes towards inclusion, self-efficacy in teaching in an inclusive school system and teaching profession (special education vs. others). In addition, there was a positive correlation with self-reported mastery-oriented instructional practices. No clear differences in knowledge or misconceptions were found for teachers' attendance of further training on inclusion. Teachers had higher knowledge and more misconceptions than the student comparison samples generated from *Study 1*.

In *Study 3*, primary school teachers' knowledge was linked to variables at student level. With the aim of supporting inclusive processes, the promotion and appreciation of all learners (Grosche, 2015), subjective data on primary school students' social participation and

objective data on their academic achievement ($N_1 = 421$, $N_2 = 715$) were collected. These were linked to knowledge of their 23 and 39 teachers in 25 and 40 classes with students with and without SEN EBD and LD. Multilevel analyses revealed lower academic achievement for students with SEN compared to their peers without SEN. Students with SEN EBD felt less socially integrated. Teachers' knowledge of SEN EBD was positively related to the social participation and academic achievement of all students (regardless of their SEN) in classes with students with and without SEN EBD. There was no corresponding correlation with knowledge about SEN LD in classes with students with and without SEN LD.

In summary, the three studies indicate that the test is appropriate for use in researching and implementing teacher education and training. The reliability estimators here used suggest adequate measurement accuracy and the classification of the subscales fit the empirical data. A broad nomological net around knowledge about diversity provides hints about validity and plausibly aligns the construct to the COACTIV competency model (Baumert & Kunter, 2013a; Kunter, Kleickmann et al., 2013). The results thus confirm previous findings on tests with different focuses (e.g., König et al., 2017; Pflanzl et al., 2013). The totality of constructs related to knowledge as measured here suggests that knowledge about diversity is an economically measurable part of teachers' professional knowledge that is relevant for both teachers and students in the context of inclusive education. Finally, questions for future research, which may arise for the evaluation and use of the test, and implications for the practical use of the test are discussed.

Einleitung

Im Zuge der Erstellung dieser Arbeit habe ich viele Gespräche mit Lehrkräften und Schulleiter:innen verschiedener Schulformen geführt. Dabei zeigte sich ein deutliches Muster: Inklusion? Ja! Umsetzbarkeit? Nein! Die Lehrkräfte, mit denen ich gesprochen habe, begrüßen die Diversität der Schüler:innen und das Bestreben um die Inklusion von allen Schüler:innen und im Speziellen von denjenigen mit besonderem pädagogischen Unterstützungsbedarf. Aber wie soll das Ganze umgesetzt werden, wenn zu wenige Ressourcen zur Verfügung stehen? Diese Arbeit wird viele von den berichteten Problemen wie zu große Klassen, zu wenig Personal und Vorbereitungszeit, eine zu laute Arbeitsumgebung und hinderliche Vorgaben nicht lösen können. Das System Schule ist komplex und lässt sich in seiner Gesamtheit nur schwer erfassen. Deshalb sehe ich es als Aufgabe der Wissenschaft, das Ganze in seine Einzelteile zu zerlegen, um Einflussbereiche isolieren, bearbeiten und verbessern zu können.

Ein erster Überblick deutet auf Vorteile schulischer Inklusion: Zusammenfassende Analysen zeigen neutrale bis positive und nur wenige negative Effekte inklusiver Beschulung für Schüler:innen mit und ohne SPF (Kart & Kart, 2021; Krämer et al., 2021; Ruijs & Peetsma, 2009; Szumski et al., 2017). Aktuelle Befragungen von Lehrkräften legen aber weder eine klare Favorisierung inklusiver im Vergleich zu nicht-inklusive Beschulung noch ein eindeutiges Gefühl, auf das Unterrichten inklusiver Klassen vorbereitet zu sein, ihrerseits nahe (Dignath et al., 2022). Auch sinken die positiven Einstellungen, das wahrgenommene Wissen und die Selbstwirksamkeitserwartungen bezogen auf schulische Inklusion von Lehramtsstudierenden im Übergang vom Studium zum Beruf (Mintz et al., 2020). Ein Realitätsschock, der das eben beschriebene Spannungsfeld zwischen der theoretischen und praktischen Umsetzung schulischer Inklusion widerspiegelt? Definitiv festhalten lässt sich, dass die Diversität der Schüler:innen im Kontext schulischer Inklusion eine besondere

Herausforderung für Lehrkräfte darstellt, auf die sie vorbereitet und mit der sie nicht allein gelassen werden sollten (Horne & Timmons, 2009; Mintz et al., 2020; Rowan et al., 2021).

Ziel dieser Arbeit ist die Erforschung eines Puzzleteils in diesem Gesamtgefüge – das Wissen über Diversität von (angehenden) Lehrkräften im Kontext schulischer Inklusion. Es wird ein neuer Wissenstest vorgestellt und evaluiert, mit dessen Hilfe das Verständnis um die Bedeutung dieses Wissens in Ergänzung zu bisherigen Tests (z. B. Sousa et al., 2017; Wildbrett et al., 2022) erweitert werden soll. Dazu werden Zusammenhänge zu verschiedenen Aspekten der Lehrkräftebildung, der professionellen Kompetenz von Lehrkräften und dem Unterrichtserfolg auf Basis eines gängigen Kompetenzmodells (Baumert & Kunter, 2013a; Kunter, Kleickmann et al., 2013; Voss et al., 2015) gesucht. Es werden soziale und akademische Variablen auf Seiten der Schüler:innen betrachtet, damit ein Beitrag zur Erforschung und zum Voranbringen schulischer Inklusion geleistet werden kann (vgl. Göransson & Nilholm, 2014; Grosche, 2015). Die Erkenntnisse können zum Beispiel genutzt werden, um das Bild zur professionellen Kompetenz von Lehrkräften zu erweitern und evidenzbasierte Wissensinhalte in der Aus- und Fortbildung von Lehrkräften zu prüfen und zu fördern.

Die Arbeit besteht aus drei empirischen Studien mit einer gemeinsamen theoretischen Einleitung, die mit den zugrundeliegenden Fragestellungen endet, und einer gemeinsamen Diskussion der Studienergebnisse. In Kapitel 1 des Theorieteils wird das Konzept schulische Inklusion (und damit einhergehend auch Heterogenität und Diversität) mit verschiedenen Konzeptualisierungen und der praktischen Umsetzung an Schulen vorgestellt. Daraus wird eine dieser Arbeit zugrundeliegende Definition schulischer Inklusion abgeleitet. Mit empirischen Befunden zu Effekten inklusiver vs. nicht-inklusive r Beschulung werden zunächst die Schüler:innen als primäre Empfänger:innen schulischer Inklusion in den Blick genommen. Im Anschluss werden die Lehrkräfte als Akteur:innen schulischer Inklusion beleuchtet. Eine Facette ihrer professionellen Kompetenz ist das Wissen, das in Kapitel 2

vorgestellt wird. Im Kontext schulischer Inklusion sind das pädagogisch-psychologische Wissen und das Wissen über Diversität von besonderer Bedeutung, und es werden Zusammenhänge zu formalen Lerngelegenheiten, verschiedenen Aspekten professioneller Kompetenz (v.a. berufliche Überzeugungen), der Performanz und dem Unterrichtserfolg aufgezeigt. Die Möglichkeiten zur Erfassung des Wissens über Diversität sind bislang begrenzt und bieten noch Entwicklungspotenzial: So liegen Tests zum pädagogisch-psychologischen Wissen mit Diversität als Subskala und Tests zu einzelnen Diversitätsbereichen vor. In Ergänzung dazu wird in Kapitel 3 ein Wissenstest mit den fünf Diversitätsbereichen SPF ESE, SPF Lernen, Teilleistungsstörungen, Geschlecht und kognitive Hochbegabung vorgestellt. Die Diversitätsbereiche spiegeln wichtige Aspekte schulischer Inklusion wider und sind als Auswahl aus allen in diesem Kontext relevanten Merkmalen zu verstehen. Es werden das Item- und Antwortformat und Beispielitems vorgestellt. Zur Testevaluation werden die Klassische Testtheorie und die Haupt- und Nebengütekriterien herangezogen. In Kapitel 4 folgen die sich daraus ableitenden Fragestellungen und ein Überblick über die drei empirischen Studien, die zu ihrer Prüfung durchgeführt wurden. Kapitel 5 besteht aus den drei Studien in Manuskriptform. Die zentralen Ergebnisse der einzelnen Studien werden in der Diskussion in Kapitel 6 zusammengefasst. Es folgen studienübergreifende Interpretationen und Stärken und Grenzen der Arbeit vor dem Hintergrund der bisherigen Forschung. Die Arbeit schließt mit Implikationen für Forschung und Praxis und einem studienübergreifenden Fazit.

1 Schulische Inklusion

Ein inklusives Schulsystem führt zu mehr Toleranz und einem besseren Miteinander, wirkt sich positiv auf die Persönlichkeitsentwicklung und das Selbstwertgefühl von Kindern aus und verbessert die Chancengleichheit in unserer Gesellschaft: So empfinden es ca. $\frac{3}{4}$ der 1500 Teilnehmenden an einer repräsentativen deutschlandweiten Umfrage (Hess et al., 2019). Diese positive Einstellung in tatsächlich gelingende schulische Inklusion umzusetzen, ist Aufgabe der im System Schule tätigen Praktiker:innen, ebenso wie der in der Forschung tätigen Wissenschaftler:innen (Göransson & Nilholm, 2014). Das folgende Kapitel dient als Grundlage und Überblick: Es werden verschiedene Definitionen von Inklusion aus wissenschaftlicher Perspektive und die praktische Umsetzung schulischer Inklusion an Schulen (in Deutschland) vorgestellt, damit im Fazit eine konkrete Definition für die vorliegende Arbeit dargelegt werden kann. Vor diesem Hintergrund werden abschließend empirische Befunde mit dem Fokus auf Schüler:innen als Empfänger:innen und Lehrkräften als Akteur:innen schulischer Inklusion dargestellt.

1.1 Definition schulischer Inklusion

Die eine Definition schulischer Inklusion zu finden, ist schwierig bis nahezu unmöglich. Inklusion wird vielfach als Ermöglichung physischer Anwesenheit und Beachtung sozialer und akademischer Bedürfnisse von Schüler:innen mit Behinderung bzw. sonderpädagogischem Förderbedarf (SPF) definiert (z. B. Kart & Kart, 2021; Ruijs & Peetsma, 2009). Behinderung und SPF sind dabei nicht gleichzusetzen: Eine Behinderung (z. B. Lern- und Entwicklungsstörungen, geistige Behinderung, Körperbehinderung) kann, muss aber nicht, zu einem Bedarf an sonderpädagogischer Förderung führen. In diesem Fall wird nach der Ausbildungsordnung sonderpädagogische Förderung (AO-SF) ein entsprechendes Feststellungsverfahren eingeleitet (Ministerium für Schule und Bildung des

Landes Nordrhein-Westfalen [MSB NRW], 2016). Dementsprechend gibt es Schüler:innen mit Behinderung, denen ein SPF attestiert wurde, und Schüler:innen mit Behinderung, die keine besondere Unterstützung benötigen. Die Unterscheidung zwischen Schüler:innen mit und ohne Behinderung bzw. SPF ist eine Inklusionsdefinition, die vielfach diskutiert wird und zu kurz greift (Göransson & Nilholm, 2014; Grosche, 2015). Im Folgenden werden deshalb verschiedene Konzeptionalisierungsansätze herausgegriffen und erläutert, die den Diskurs um schulische Inklusion im deutschsprachigen und internationalen Raum maßgeblich geprägt haben, und um die Begriffe Heterogenität und Diversität ergänzt. Auf dieser Basis wird die praktische Umsetzung schulischer Inklusion beleuchtet. Beide Perspektiven dienen als Grundlage für das Inklusionsverständnis in dieser Arbeit, das im Fazit vorgestellt wird.

1.1.1 Inklusionsmodelle, Heterogenität und Diversität

Ein im deutschen Sprachraum vorliegendes Indikatorenschema schulischer Inklusion (Krämer et al., 2016) mit 15 Dimensionen (z. B. Partizipation/Chancengleichheit, Respekt/Wertschätzung/Anerkennung, Diagnostik/Ergebnisse, Förderung/Unterstützung) ermöglicht eine sehr umfassende Definition. Ausgewählte Elemente sind zum Beispiel, dass alle Schüler:innen eine gemeinsame Schule besuchen, an der Ressourcen zur Verwirklichung von Inklusion flexibel und transparent eingesetzt werden. Alle Mitarbeitenden dieser Schule respektieren, wertschätzen und anerkennen die Individualität aller Schüler:innen. Dementsprechend werden Lernhindernisse beseitigt und es werden alle Schüler:innen gefördert, um individuelle Lernprozesse zu meistern. Die Lehrkraft und der Unterricht stellen methodische und didaktische Maßnahmen zur Förderung von Chancengleichheit und Gemeinsamkeit bereit. Lerninhalte, Lernumgebungen, Aktivitäten und Materialien werden individualisiert und differenziert für alle Voraussetzungen und Bedürfnisse der Schüler:innen geplant, angepasst und durchgeführt (Krämer et al., 2016). Klarer Vorteil dieses Schemas ist die Möglichkeit, Inklusion sehr vielschichtig und umfassend zu definieren. Eine möglichst prägnante Definition, wie sie für empirische Arbeiten meistens benötigt wird, auf Basis aller

15 Dimensionen erscheint allerdings nur schwer umsetzbar.

In einem Überblicksartikel über Reviews zur Erforschung schulischer Inklusion im internationalen Kontext wurden vier qualitativ unterschiedliche Definitionskategorien zusammengestellt, in denen aktuelle Studien prägnant verortet werden können (Göransson & Nilholm, 2014). Die Kategorien A bis D sind hierarchisch geordnet (wobei Kategorie D die strengsten Kriterien enthält und somit auch alle anderen Kategorien mit einschließt). Zur Kategorie A gehört die weit verbreitete Platzierungsdefinition. Inklusion wird also als Platzierung von Schüler:innen mit Behinderungen bzw. SPF in Regelklassen definiert. Kategorie B beinhaltet eine spezifizierte und individualisierte Definition. Demnach ist Inklusion als Erfüllung der sozialen und akademischen Bedürfnisse von Schüler:innen mit Behinderungen bzw. SPF zu verstehen. Zur Kategorie C zählt eine allgemeine individualisierte Definition mit Inklusion als Erfüllung der sozialen und akademischen Bedürfnisse aller Schüler:innen. Nach der Gemeinschaftsdefinition der Kategorie D ist Inklusion die Schaffung von Gemeinschaften mit spezifischen Merkmalen. Schüler:innen mit Behinderungen bzw. SPF werden also nicht gesondert hervorgehoben. Der Unterstützungsbedarf ist dementsprechend als ein Merkmal von vielen zu verstehen.

In einem im deutschsprachigen Raum häufig zitierten Diskussions- und Positionsartikel (Grosche, 2015) wiederum werden hierarchische Definitionsmodelle kritisiert. Nach einem Modell von Wocken (2009) zum Beispiel erhalten Menschen mit Behinderung bzw. SPF in der vierten Stufe (Integration) das Recht auf soziale Teilhabe. In dieser Stufe wird noch eine dichotome Kategorisierung vorgenommen. In der fünften Stufe (Inklusion) steht das Recht auf Selbstbestimmung und Gleichheit im Zentrum, und es wird keine Einteilung in Kategorien mehr vorgenommen. Grosche (2015) kritisiert unter anderem, dass die Stufen nicht distinkt sind, sich also überlappen (z. B. könnte eine Lehrkraft alle Schüler:innen gleichermaßen wertschätzen und trotzdem eine Diagnostik zum SPF vornehmen), und dass die letzte Stufe in unserer Gesellschaft bislang kaum umgesetzt wird.

Dementsprechend dürfte weder aus praktischer noch aus wissenschaftlicher Perspektive der Begriff Inklusion tatsächlich genutzt werden. Obwohl das Modell von Göransson und Nilholm (2014) einen wichtigen Rahmen zur Begriffseinordnung bietet, lässt sich die Kritik auch auf dieses Modell übertragen. Grosche (2015) definiert stattdessen die Förderung und Anerkennung als zwei grundlegende Ziele von Inklusion, die sich gegenseitig bedingen, aber auch behindern können. Zum Beispiel ermöglichen akademische Kompetenzen die Teilhabe in der Gesellschaft, und das Gefühl der Wertschätzung kann wiederum das Lernen erleichtern. Es kann aber auch sein, dass bei der passgenauen Förderung die Anerkennung leidet, und die Anerkennung von Individualität kann dazu führen, dass nicht gleichermaßen gefördert wird. Dementsprechend sollten in empirischen Studien unbedingt sowohl die Förderung als auch die Anerkennung als abhängige Variablen erhoben werden (Grosche, 2015).

Neben den genannten Konzeptionalisierungsansätzen schulischer Inklusion sind ergänzend die Begriffe Heterogenität und Diversität von Bedeutung. Ein umfassendes Inklusionsverständnis (vgl. Kategorie C und D nach Göransson & Nilholm, 2014; Grosche, 2015; Krämer et al., 2016; Stufe 5 nach Wocken, 2009) bezieht sich auf die Bedürfnisse einer heterogenen Schülerschaft. Dabei ist Heterogenität im Sinne einer gesellschaftlichen Vielfalt als Kontinuum verschiedener Merkmale zu verstehen und wird definiert als die Unterschiedlichkeit von Schüler:innen als Individuen und Teil einer Schulklasse bzw. Schule/Schulform (Zulliger & Tanner, 2013). Beispiele für Heterogenitätsmerkmale sind Leistung, Geschlecht, sozialer Hintergrund, Sprache, Religion, Intelligenz, kultureller Hintergrund, sexuelle Orientierung, Alter und Behinderungsmerkmale (Faulstich-Wieland, 2011; Grosche, 2015; Zulliger & Tanner, 2013). Heterogen ist dabei als eine relative, partielle, konstruierte und wertneutrale Eigenschaft zu verstehen (Lang et al., 2010): Ob eine Gruppe an Menschen als heterogen zu bezeichnen ist, hängt maßgeblich davon ab, mit welcher anderen Gruppe sie verglichen wird (= relativ). Außerdem können sich die Unterschiede zwischen den beschriebenen Personen im Laufe der Zeit vergrößern oder

verkleinern, die Beschreibung dieser Gruppe ist also nicht allumfassend (= partiell). Auch ist eine Gruppe nicht per se heterogen, sondern wird erst so bezeichnet, wenn ein konstruierter Maßstab (z. B. Leistung bei der Bezeichnung einer leistungsheterogenen Klasse) angelegt wird (= konstruiert). Vom Grund her ist Heterogenität weder etwas Gutes noch etwas Schlechtes (= wertneutral). Die Bezeichnung erfolgt allerdings immer vor dem Hintergrund bestimmter Interessen und in bestimmten Zusammenhängen und kann damit eine Wertung erhalten. Im Schulkontext birgt der Begriff Heterogenität zum Beispiel die Problematik, dass er häufig negativ aufgefasst und mit unüberwindbaren Hürden in Verbindung gebracht wird (Krüger-Potratz, 2011). Aus diesem Grund soll ergänzend der Begriff Diversität eingeführt werden, der zum Beispiel im Bereich der Personalentwicklung eher aus dem Blickwinkel der positiven Effekte einer an Verschiedenheit und Vielfalt ausgerichteten Unternehmensführung betrachtet wird. Die Begriffe Heterogenität und Diversität beschreiben in diesem Zusammenhang dieselben Ziele, nämlich „unterschiedlichen Ausgangslagen und Interessen gerecht zu werden, ohne auszugrenzen“ (Krüger-Potratz, 2011, S. 192). Im Folgenden wird der Begriff Diversität genutzt, der Begriff Heterogenität hingegen wird nur dann verwendet, wenn er zum Beispiel Teil eines Messinstruments ist und im Sinne eines Namens verwendet wird.

1.1.2 Praktische Umsetzung an Schulen

Die beiden Schriften, die die Umsetzung schulischer Inklusion in Deutschland und anderen Ländern maßgeblich beeinflusst haben, sind die Salamanca Erklärung und der Aktionsrahmen für Sonderpädagogik (Salamanca statement and framework for action on special needs education; United Nations Educational, Scientific and Cultural Organization [UNESCO], 1994) und die UN-Behindertenrechtskonvention als Übereinkommen über die Rechte von Menschen mit Behinderungen (Beauftragter der Bundesregierung für die Belange von Menschen mit Behinderungen, 2018). Letzteres wurde in Deutschland im Jahr 2009 ratifiziert und damit verbindlich. Dementsprechend ist der Schwerpunkt „Differenzierung,

Integration und Förderung – Diversität und Heterogenität als Bedingungen von Schule und Unterricht“ auch in den bildungswissenschaftlichen Standards der Lehrkräftebildung curricular verankert (Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland [KMK], 2014, S. 5). Nach dem Salamanca Statement Framework (UNESCO, 1994) sollen Schulen alle Kinder unabhängig von ihren körperlichen, intellektuellen, sozialen, emotionalen, sprachlichen und anderen Bedingungen aufnehmen und fördern. Dabei wird Sonderpädagogik als Prinzip beschrieben, von dem alle Kinder profitieren können, und inklusive Regelschulen werden als wirksamstes Mittel zur Erschaffung einer inklusiven Gesellschaft und zur Bekämpfung diskriminierender Einstellungen angesehen (vgl. Kategorie C und D nach Göransson & Nilholm, 2014). In Artikel 24 der UN-Behindertenrechtskonvention (Beauftragter der Bundesregierung für die Belange von Menschen mit Behinderungen, 2018) wird festgehalten, dass Menschen mit Behinderungen ein Recht auf Bildung ohne Diskriminierung und auf Basis von Chancengleichheit haben und nicht vom allgemeinen Bildungssystem ausgeschlossen werden dürfen (vgl. Kategorie A und B nach Göransson & Nilholm, 2014)

In Deutschland (hier konkret Nordrhein-Westfalen) sollen (angehende) Lehrkräfte die Begriffe und Merkmale von Diversität kennen, um die Bedeutung von Diversität und ihrer Anerkennung für Lernprozesse wissen und ihr professionelles Handeln im Umgang mit Diversität reflektieren (KMK, 2014). In Bezug auf das Bildungssystem werden auf Basis der Beschlüsse Kinder mit und ohne Behinderung bzw. SPF gemeinsam an allgemeinen Schulen unterrichtet. Dies wird als inklusive Bildung bezeichnet (MSB NRW, 2016; vgl. Platzierungsdefinition, Göransson & Nilholm, 2014). Abweichend können Eltern die Beschulung in einer Förderschule wählen. Das verstärkte Bemühen um Inklusion spiegelt sich auch in offiziellen Statistiken wider: Bundesweit wurden im Jahr 2020 43.7% der Schüler:innen mit SPF inklusiv an Regelschulen unterrichtet. Im Jahr 2011 waren es nur 25.0% (KMK, 2022). Am häufigsten vertreten waren 2020 die Förderschwerpunkte ESE,

Hören und Lernen, gefolgt von Sehen, Sprache, körperliche und motorische Entwicklung und zum Schluss geistige Entwicklung.

Diese Entwicklungen haben dazu geführt, dass Klassen und die zu unterrichtenden Schüler:innen in Bezug auf unterrichts- und leistungsrelevante Merkmale immer diverser wurden. Dadurch wurde allerdings nicht die Diversität der Schüler:innen in Bezug auf alle theoretisch denkbaren Merkmale erhöht (vgl. Diversität als relative, partielle, konstruierte und wertneutrale Eigenschaft, Lang et al., 2010). Im System Schule wird vor allem Homogenität in Bezug auf die Merkmale Alter und Leistung angestrebt, indem verschiedene Schulklassen und Schulformen gebildet werden (Faulstich-Wieland, 2011). Dementsprechend erscheint es plausibel, dass Inklusion und die damit einhergehende Diversität sowohl aus wissenschaftlicher als auch praktischer Perspektive als Herausforderung wahrgenommen werden und Inklusion nicht abgeschlossen ist, sondern als Prozess verstanden wird (Besand & Jügel, 2015; Gebhardt et al., 2012; Scharenberg, 2013; Wischer, 2007).

1.1.3 Fazit

Wenn alle Kriterien der vorgestellten Inklusionsmodelle erfüllt sein müssen, kann aktuell keine Studie und auch kein Klassenraum die Bezeichnung Inklusion für sich beanspruchen (Grosche, 2015). Im Rahmen von wissenschaftlichen Arbeiten sollte das Inklusionsverständnis deshalb vor allem eindeutig definiert werden. Dabei handelt es sich um eine zeitlich begrenzte, einzelne und reduzierte Definition. Keine Studie muss demnach dem Anspruch gerecht werden, alle Definitionen gleichermaßen zu bedienen. Wichtiger sind die klare Einordnung und die Reduktion auf Kriterien, die für das Forschungsdesign, die Stichprobe, die Variablen, etc. von Relevanz sind. Dieser Arbeit liegt deshalb ein Inklusionsverständnis zugrunde, das möglichst hohe Anforderungen an inklusive Prozesse und deren Erforschung stellt und gleichzeitig an die praktische Umsetzung an Schulen angepasst ist. Als Basis wird Inklusion in Anlehnung an Grosche (2015) deshalb als Förderung und Anerkennung aller Schüler:innen angesehen. Entgegen der höchsten Kategorie

von Wocken (2009), nach der Merkmale wie Behinderungen und SPF nicht mehr relevant sind, wird trotzdem eine explizite Benennung dieser Kategorien in Einklang mit der Umsetzung von Inklusion an Schulen (MSB NRW, 2016) vorgenommen. Da diese Merkmale aber nur als Auswahl aus allen Merkmalen zur Beschreibung von Diversität zu verstehen sind (vgl. Faulstich-Wieland, 2011; Grosche, 2015; Zulliger & Tanner, 2013), werden weitere Diversitätsdimensionen in den Blick genommen. Auf dieser Basis hat diese Arbeit das Ziel, Faktoren zu identifizieren, die Inklusion vorantreiben bzw. die Grundlage dafür zu schaffen (vgl. Göransson & Nilholm, 2014), und es werden sowohl soziale als auch akademische Variablen untersucht (Göransson & Nilholm, 2014; Grosche, 2015). Darauf aufbauend wird im Folgenden eine Auswahl an Studien vorgestellt, die vor allem zusammenfassende Artikel (Reviews, Meta-Analysen) sind und einen Überblick über empirische Befunde zu Schüler:innen und Lehrkräften im Kontext schulischer Inklusion ermöglichen.

1.2 Empirische Befunde

Ziel schulischer Inklusion ist die Förderung und Anerkennung aller Schüler:innen (Grosche, 2015), die folglich die primären Empfänger:innen sind, Inklusion natürlich aber auch selbst mitgestalten. Diesem Ziel entsprechend werden im Folgenden zuerst Studien vorgestellt, die sich mit den Effekten inklusiver vs. nicht-inklusive Beschulung auf die Schüler:innen befassen. In Anlehnung an die Forderungen von Göransson und Nilholm (2014) werden sowohl akademische als auch soziale Variablen in den Blick genommen. Im Anschluss folgen Studien, in denen Variablen auf Seiten der Lehrkräfte als Umsetzer:innen schulischer Inklusion betrachtet werden.

1.2.1 Effekte inklusiver vs. nicht-inklusive Beschulung auf Schüler:innen

Die wohl am häufigsten genutzte Definition unterscheidet zwischen Regelschulklassen mit Schüler:innen mit SPF (= inklusive Beschulung) und Regelschulklassen ohne Schüler:innen mit SPF bzw. Förderschulklassen mit Schüler:innen mit SPF (= nicht-inklusive

Beschulung) und folgt somit der oben beschriebenen Platzierungsdefinition (Göransson & Nilholm, 2014). Dies bedeutet allerdings nicht, dass damit automatisch nur die Kriterien dieser einen Kategorie erfüllt sind (vgl. Grosche, 2015). Entsprechen Studien dem Ziel der Erfüllung der sozialen und akademischen Bedürfnisse von Schüler:innen mit Behinderungen bzw. SPF und der Erfüllung der sozialen und akademischen Bedürfnisse aller Schüler:innen, können ebenfalls die nächsthöheren Kategorien bedient werden.

In einem systematischen Überblicksartikel zu inklusiver vs. nicht-inklusive Beschulung und sozio-emotionalen und akademischen Variablen bei Schüler:innen mit und ohne (leichtem bis moderatem) SPF konnten neutrale bis positive Effekte inklusiver Beschulung zusammengefasst werden, aus Einzelstudien gingen auch negative Effekte hervor (Ruijs & Peetsma, 2009). Methodische Unterschiede und Mängel in den Primärstudien wie die unterschiedliche Dauer der Inklusion (in gesamten Unterricht vs. nur in einigen Stunden) und fehlende Kontrollgruppen lassen allerdings keine eindeutigen Schlüsse zu. Auch in einem vergleichbaren Überblicksartikel zu Effekten gemeinsamer Beschulung von Schüler:innen mit und ohne Behinderung (SPF mit eingeschlossen) auf sozio-emotionale und akademische Variablen bei Schüler:innen ohne Behinderung wurden neutrale bis positive Effekte aufgezeigt (Kart & Kart, 2021). In einer Meta-Analyse mit dem Fokus auf Schulleistungen von Schüler:innen ohne SPF in inklusiven vs. nicht-inklusive Regelklassen zeigten sich kleine positive Effekte inklusiver Beschulung ($d = 0.12$, Szumski et al., 2017). Eine Meta-Analyse mit Stichproben aus Schüler:innen mit und ohne den konkreten SPF Lernen ergab Effekte zugunsten der inklusiven Beschulung für die akademischen Leistungen ($d = 0.35$), nicht aber die psychosozialen Variablen ($d = 0.00$) von Schüler:innen mit SPF Lernen (Krämer et al., 2021). Schüler:innen ohne SPF Lernen unterschieden sich nicht in akademischen ($d = -0.14$) und psychosozialen Variablen ($d = 0.06$) in Abhängigkeit der Beschulung. Zusammenfassend sprechen die Ergebnisse für neutrale bis positive Effekte inklusiver Beschulung, in Einzelstudien konnten aber auch negative Effekte aufgezeigt

werden. Als Moderatoren der Zusammenhänge erwiesen sich das Studiendesign (kleinere Effekte in Längsschnittstudien, Krämer et al., 2021), das Publikationsjahr (größere Effekte in älteren Studien, Krämer et al., 2021), das Ausmaß des Förderbedarfs (positivere Effekte bei leichtem SPF, Szumski et al., 2017), das Land der Erhebung (positive Effekte in den USA und Kanada und neutrale in Europa, Szumski et al., 2017) und die Schulform (negativere Effekte in weiterführenden Schulen, Kart & Kart, 2021; Krämer et al., 2021).

Neben diesen studien- und stichprobenspezifischen Moderatoren spielen Lehrkräfte eine Schlüsselrolle in der direkten Umsetzung schulischer Inklusion (de Boer et al., 2011; Hehir et al., 2016; Ruijs & Peetsma, 2009). In einem Review zu pädagogischen Ansätzen zur erfolgreichen Inklusion von Kindern mit SPF heben die Autor:innen die Bedeutung von Lehrkräften und ihrer wahrgenommenen Verantwortlichkeit für alle Lernenden hervor (Rix et al., 2009). Eine Meta-Analyse zum Selbstkonzept von Schüler:innen mit SPF Lernen ergab in Einklang mit den oben genannten Studien keine Unterschiede zwischen den untersuchten inklusiven und nicht-inklusive Settings (Elbaum, 2002). Zusatzanalysen deuten aber auf Moderatoreffekte von besonderer Unterstützung in Regelschulklassen hin. Fanden zum Beispiel Co-Teaching oder eine außerunterrichtliche Beratung der Regelschullehrkraft durch eine Lehrkraft für sonderpädagogische Förderung statt, hatten die betreffenden Schüler:innen ein höheres Selbstkonzept als in Regelschulklassen ohne besondere Unterstützung. In einer weiteren Meta-Analyse waren verhaltensbasierte Interventionen in inklusiven Settings für Schüler:innen mit Autismusspektrumstörung gleichermaßen effektiv, wenn sie von Wissenschaftler:innen und Lehrkräften durchgeführt wurden (Camargo et al., 2016). Lehrkräfte sind demnach nicht der einzige, aber ein wichtiger Faktor in der Umsetzung schulischer Inklusion.

1.2.2 Lehrkräfte als Akteur:innen schulischer Inklusion

Die wohl am häufigsten betrachteten Variablen im Kontext schulischer Inklusion auf Lehrkräfteebene sind die Einstellungen zu inklusiver Beschulung und die

Selbstwirksamkeitserwartungen im Umgang mit Inklusion. Beide Konstrukte gelten als wichtige Gelingensbedingungen schulischer Inklusion (de Boer et al., 2011; Fernandes et al., 2021; Lautenbach & Heyder, 2019; Wray et al., 2022) und beziehen sich auf die individuelle Perspektive einer Person auf ein Thema (vgl. Überzeugungen, Pajares, 1992). Sie können mit Items wie „Je mehr Zeit Kinder mit besonderen pädagogischen Bedürfnissen in einer Regelklasse verbringen, desto eher werden sie von anderen Mitschülern ihrer Klasse gut behandelt“ (Einstellungen, Kunz et al., 2010) und „Ich bin mir sicher, dass ich auch bei größten Leistungsunterschieden für jedes Kind ein angemessenes Lernangebot bereithalten kann“ (Selbstwirksamkeitserwartungen, Gebhardt et al., 2015) erfasst werden. Es liegen auch Studien vor, in denen die kognitive Komponente von Einstellungen Wissensaspekte enthält (de Boer et al., 2011). Dennoch werden Einstellungen und Selbstwirksamkeitserwartungen im Folgenden vom Wissen (= objektive, in sich stimmige und einem Gruppen-/Forschungskonsens entsprechenden Fakten) getrennt (vgl. Baumert & Kunter, 2013a; Voss et al., 2011).

Dass Einstellungen und Selbstwirksamkeitserwartungen eng miteinander verknüpft sind, konnte in einer Meta-Analyse festgehalten werden, in der ein positiver Zusammenhang von $r = 0.35$ vorlag (Yada et al., 2022). In einem Überblicksartikel lagen bei den Lehrkräften negative bis neutrale Einstellungen gegenüber inklusiver Beschulung vor und keine betrachtete Studie berichtete positive Einstellungen (de Boer et al., 2011). Die kognitiven und emotionalen Einstellungen gegenüber inklusiven vs. nicht-inklusive Klassen von (angehenden) Lehrkräften lagen in einer neueren Meta-Analyse im Schnitt etwas über dem Mittel der Skala und zeigten keine eindeutige Favorisierung einer der beiden Beschulungsformen – unabhängig davon ob Lehramtsstudierende oder Lehrkräfte befragt wurden (Dignath et al., 2022). Die Selbstwirksamkeitserwartungen im Unterrichten von inklusiven Klassen lagen etwas höher, wobei sich Lehramtsstudierende selbstwirksamer fühlten als Lehrkräfte. Diejenigen, die für sonderpädagogische Förderung ausgebildet waren,

hatten positivere kognitive Einstellungen als diejenigen mit anderen Studienhintergründen ($d = 0.41$). Ein systematisches Review über die Einstellungen von (hier speziell spanischen) Lehrkräften waren tendenziell positiv, zum Teil aber nicht eindeutig (Lacruz-Pérez et al., 2021). Als Einflussgrößen für die Entwicklung von positiven Einstellungen gelten zum Beispiel die Erfahrung, die (angestrebte) Schulform und das betrachtete Diversitätsmerkmal (Ruberg & Porsch, 2017). Bei der Entwicklung von hohen Selbstwirksamkeitserwartungen spielen zum Beispiel die Lehramtsausbildung, das berufliche Lernen, der Erfahrungsaustausch mit Menschen mit Behinderungen und das Wissen eine wichtige Rolle (Wray et al., 2022).

Einstellungen, Selbstwirksamkeitserwartungen und Wissen im Kontext schulischer Inklusion erwiesen sich insgesamt als veränderbare Konstrukte: Die in einer Meta-Analyse zusammengefassten Interventionen konnten die kognitiven Einstellungen von (angehenden) Lehrkräften positiv verändern ($d = 0.63$, Dignath et al., 2022), was besonders für Ansätze mit praktischen Lernerfahrungen galt. Der Kontakt zu einer Person mit SPF im Rahmen einer Intervention hatte hingegen keinen signifikanten Effekt auf die Einstellungsänderung. Auch die in einem Review zusammengefassten Interventionen mit Lehramtsstudierenden konnten die Einstellungen in den meisten Studien positiv verändern (Lautenbach & Heyder, 2019). Dies galt sowohl für informationsbasierte kognitive Interventionen als auch für solche mit einer Kombination aus Informationen und praktischen Erfahrungen. Interventionen für lehrplangestützte Messungen zur Identifikation von Lernbehinderungen und für eine datenbasierte individualisierte Unterrichtsgestaltung konnten das Wissen ($ES \leq 2.28$) deutlicher steigern als die Fertigkeiten ($ES \leq 1.96$) und die Selbstwirksamkeitserwartungen ($ES \leq 0.78$, Gesel et al., 2021). Diese Meta-Analyse wurde allerdings nicht explizit mit Bezug zu schulischer Inklusion durchgeführt. Ein Überblicksartikel zu Lehrkräfte-Fortbildungen zu Inklusion weist insgesamt auf positive Veränderungen in den Bereichen Einstellungen/Wahrnehmung, Wissen und Strategien/Kompetenzentwicklung hin (Tristani &

Bassett-Gunter, 2020).

Während die Konstrukte Einstellungen und Selbstwirksamkeitserwartungen im Kontext schulischer Inklusion bereits vielfach erforscht wurden, weshalb auch zahlreiche zusammenfassende Artikel zu ihnen vorliegen (z. B. de Boer et al., 2011; Fernandes et al., 2021; Lautenbach & Heyder, 2019; Wray et al., 2022), ergab die Literaturrecherche keine vergleichbaren Artikel zum Wissen von Lehrkräften. Dies gilt sowohl für das pädagogisch-psychologische Wissen als auch das Wissen über Diversität. In Kapitel 2 wird dieses Wissen von (angehenden) Lehrkräften definiert und in ihre professionelle Kompetenz eingeordnet. Sowohl theoretisch als auch empirisch lassen sich Zusammenhänge zu Variablen auf Lehrkräfte- und Schüler:innenebene herstellen, die darauf aufbauend vorgestellt werden. Im Anschluss werden bisherige Tests zur Erfassung von Wissen über Diversität präsentiert, die einen Rahmen für den in Kapitel 3 vorgestellten Wissenstest bilden.

2 Wissen als Teil professioneller Kompetenz von Lehrkräften

Um das Wissen von (angehenden) Lehrkräften eindeutig zu definieren, ist eine Abgrenzung von Überzeugungen (Pajares, 1992) und beruflichen Verhaltensweisen (Voss et al., 2015) ratsam. Wissen wird als eher kognitives Konstrukt angesehen, wohingegen Überzeugungen eher affektiv und evaluativ sind (Nespor, 1987). Im Gegensatz zum Konstrukt der Überzeugungen können für die Quantifizierung von Wissen Aussagen als eindeutig falsch oder richtig bewertet werden (Voss et al., 2015). Dabei müssen die als Wissen klassifizierten Aussagen in sich konsistent sein und einem Gruppen- bzw. wissenschaftlichen Konsens entsprechen (Baumert & Kunter, 2013a; Voss et al., 2011). Bei Überzeugungen hingegen ist diese Bedingung nicht gegeben. Dieses Verständnis von Wissen dient als Grundlage für den hier vorgestellten Wissenstest, auch wenn Forschungsarbeiten vorliegen, in denen Überzeugungen als Teil des Wissens angesehen werden, womit die gegenseitige Abhängigkeit der beiden Konstrukte deutlich gemacht wird (vgl. Pajares, 1992). Das Wissen ist nicht gleichzusetzen mit den tatsächlich gezeigten Verhaltensweisen, obwohl es auch hier Überschneidungen geben mag. So könnte es sein, dass eine Lehrkraft weiß, dass Förderziele positiv und realistisch formuliert sein sollten, dies in der Praxis aber nicht umsetzt. Wissen ist demnach eine notwendige, nicht aber zwingend ausreichende Voraussetzung für erfolgreiches Unterrichtshandeln (Voss et al., 2015).

Ein Rahmenmodell zur Einordnung des Wissens von (angehenden) Lehrkräften entstand im Zuge des Forschungsprogramms COACTIV (Professionelle Kompetenz von Lehrkräften, kognitiv aktivierender Unterricht und die mathematische Kompetenz von Schüler:innen; Kunter, Baumert et al., 2013), das eine wichtige Grundlage für die in dieser Arbeit untersuchten Zusammenhänge bildet (für einen Überblick s. Abb. 1). Hinter dem Modell steht die Annahme, dass die verschiedenen Bereiche der professionellen Kompetenz und ihre Bedeutung für die erfolgreiche Bewältigung von Aufgaben im Lehrberuf empirisch

mess- und analysierbar sind (Baumert & Kunter, 2006). Nach dem Modell setzt sich die professionelle Kompetenz aus motivationalen Orientierungen, beruflichen Überzeugungen, der Fähigkeit zur Selbstregulation und dem Professionswissen zusammen. Letzteres ist wiederum unterteilt in die Bereiche Fachwissen, fachdidaktisches Wissen, pädagogisches Wissen, Organisationswissen und Beratungswissen.

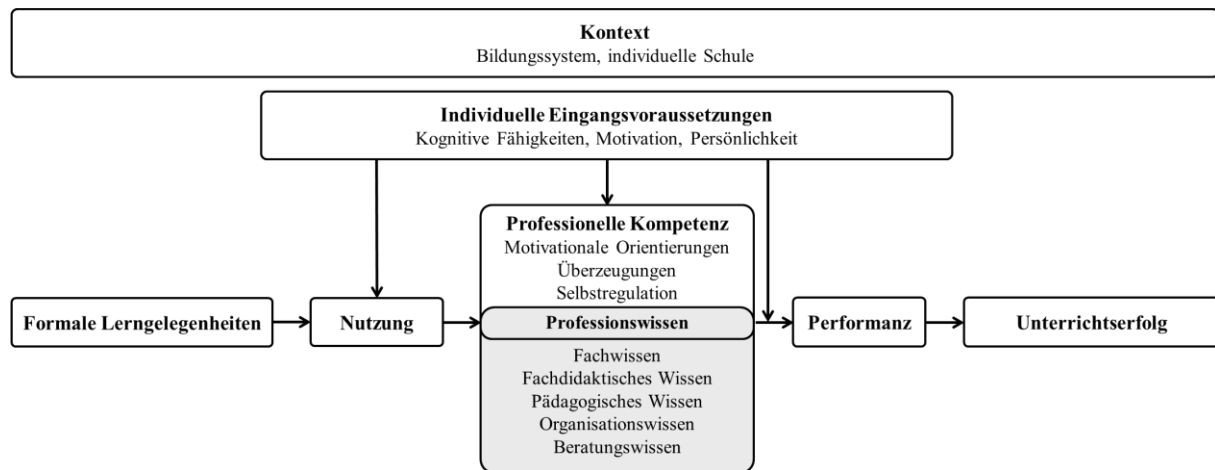


Abbildung 1. Eigene Darstellung des COACTIV-Kompetenzmodells (in Anlehnung an Kunter, Kleickmann et al., 2013, S. 67; Voss et al., 2015, S. 190)

Damit baut das Modell auf der in der Literatur wohl am häufigsten anzutreffenden Unterscheidung von Wissensarten innerhalb des Professionswissens von (angehenden) Lehrkräften nach Shulman (1987) auf, der (1) das Fachwissen (engl. content knowledge, CK), welches vertieftes fachliches Hintergrundwissen beinhaltet, (2) das fachdidaktische Wissen (engl. pedagogical content knowledge, PCK), unter dem Kenntnisse zur Übermittlung von fachspezifischen Inhalten verstanden werden, und (3) das hier fokussierte pädagogische Wissen (engl. general pedagogical knowledge, GPK) voneinander abgrenzte (z. B. Voss et al., 2015). Außerdem definierte Shulman die Kategorien Wissen über Lernende und ihre Charakteristiken, Wissen über Bildungszusammenhänge (z. B. Funktionsweise der Gruppe/des Klassenzimmers, Verwaltung/Finanzierung von Schulbezirken) und Wissen über

Ziele, Zwecke und Werte von Bildung (Shulman, 1987), die sich allerdings weniger durchgesetzt haben (Pflanzl et al., 2013). Die für diese Arbeit bedeutendste Wissensform ist das pädagogische Wissen, auch pädagogisch-psychologisches Wissen genannt, für das ab hier die Abkürzung PPK (engl. pedagogical-psychological knowledge; vgl. Heyder, Bergold & Steinmayr, 2018; Voss et al. 2011) genutzt wird.

Neben dem, was Lehrkräfte wissen, könnte auch das, was sie fälschlicherweise zu wissen glauben (= Fehlannahmen, vgl. Heyder, Bergold & Steinmayr, 2018; Schmiedeler, 2013), in der Entwicklung ihrer professionellen Kompetenz wichtig sein. Da zu den Fehlannahmen von (angehenden) Lehrkräften kein zugrunde liegendes Modell und bedeutend weniger Forschung vorliegen als zum Wissen, werden ihnen im Folgenden keine eigenen Kapitel gewidmet. Vielmehr werden Befunde zum Wissen nur an den jeweiligen Stellen um Befunde zu den Fehlannahmen ergänzt, wenn diese vorliegen.

2.1 Pädagogisch-psychologisches Wissen und Wissen über Diversität

Das PPK „umfasst Kenntnisse über das Lernen und Lehren, die sich auf die Gestaltung von Unterrichtssituationen beziehen und die fachunabhängig, das heißt auf verschiedene Fächer und Bildungsbereiche anzuwenden sind“ und ist „erlern- und vermittelbar“ (Voss et al., 2015, S. 194). Wie auch andere Facetten des Professionswissens wird das PPK durch typische Lerngelegenheiten in der Lehramtsausbildung und dem Beruf geformt und ist von diesen abhängig (vgl. König & Blömeke, 2009). Es ist wichtiger Bestandteil des theoretischen fachunspezifischen bildungswissenschaftlichen Wissens von (angehenden) Lehrkräften, bezieht sich dabei aber vor allem auf unterrichtsrelevantes Wissen, das pädagogisch-psychologisch orientiert ist, und weniger auf zum Beispiel bildungstheoretisches und bildungshistorisches Wissen, Wissen über Systemfunktionen bzw. Schulorganisation und Wissen über typische Beratungs- und Konsultationsaufgaben von Lehrkräften (Kunina-Habenicht et al., 2012). Trotz der Forderung nach einer breiteren Definition des PPK im Lehramtsstudium (z. B. Hohenstein et al., 2017), liegt dieser Arbeit

das gerade genannte Verständnis zugrunde, und das Organisations- und Beratungswissen werden als eigenständige Wissensbereiche definiert (vgl. Voss et al., 2011). Dies liegt vor allem daran, dass für den Bereich der Diversität nicht alle weiteren Aspekte notwendigerweise eingeschlossen und mitgedacht werden müssen. Der Vorteil der sehr spezifischen Erfassung bringt an dieser Stelle automatisch den Nachteil mit sich, dass nicht alle theoretisch relevanten Teilbereiche des PPK gleichermaßen Beachtung finden können.

Zur Struktur des PPK seien zwei Ansätze genannt, die in der gängigen Literatur häufig zitiert werden: (1) Eine Expert:innenbefragung über die wesentlichen Inhalte des bildungswissenschaftlichen Begleitstudiums der Lehramtsausbildung mit Fokus auf Erziehungswissenschaften, Psychologie und Soziologie ergab 104 praktisch relevante Themen, die neun Inhaltsbereichen zugeordnet werden können (Kunina-Habenicht et al., 2012). Diese sind breiter gefasst und übersteigen somit das hier zugrunde liegende Verständnis des PPK. Es handelt sich um die Bereiche Unterricht (34 Themen), Lehrerberuf (11 Themen), Bildungssystem und Schulorganisation (9 Themen), Bildungstheorie und Bildungsgeschichte (6 Themen), Entwicklungsprozesse (9 Themen), Lernprozesse (11 Themen), Sozialisationsprozesse (7 Themen), Heterogenität und soziale Konflikte (7 Themen) und Diagnostik und Evaluation (10 Themen). (2) Im Rahmen des Forschungsprogramms COACTIV entstand das eben beschriebene Kompetenzmodell, auf dessen Grundlage ein Test zum PPK konstruiert wurde (Voss et al., 2011). Dem PPK werden die acht Teilbereiche Lernprozesse, Heterogenität, Altersstufen und Lernbiographien, Klassenführung, Interaktion/Kommunikation, Lehr-Lern-Methoden und -Konzepte, Diagnostik und Gestaltung der Lernumgebung zugeordnet (Voss et al., 2015).

Trotz der unterschiedlichen Konzeptualisierung haben beide Ansätze gemeinsam, dass sie Diversität als unterrichtsrelevantes Thema einstufen (Kunina-Habenicht et al., 2012; Voss et al., 2015). Das Wissen über Diversität wird damit als elementarer Bestandteil der professionellen Kompetenz von (angehenden) Lehrkräften betitelt. Auch die bereits von

Shulman (1987) benannte Facette Wissen über Lernende und ihre Charakteristiken lässt sich diesem Teilbereich zuordnen. Dennoch liegen bislang keine Modelle und empirische Studien zur Struktur des Wissens über Diversität vor. Erste Hinweise zu Inhalten können aus Studien, in denen Tests zum PPK mit Bezug zu Diversität zum Einsatz kommen, generiert werden (für Details s. Kapitel 2.3.1 Tests zum PPK mit Diversität als Facette). Dementsprechend könnten folgende Inhalte Teil des Diversitätswissens sein: Lernprozesse, Dispositionen/Unterschiede, Diagnose, Klassenführung, Strukturierung und Binnendifferenzierung/Individualisierung (König et al., 2017), Differenzierungsmaßnahmen und Methodenvielfalt im Unterricht (König & Blömeke, 2009, 2010), individuelle Lernprozesse, Umgang mit Diversität und Diagnostik bei Lernstörungen (Wildbrett et al., 2022), Geschlecht und sozialer Hintergrund (Kunina-Habenicht et al., 2020), Vorwissen, Intelligenz, Interesse, sozioökonomischer Status, Lern- und Entwicklungsstörungen, Geschlechterdifferenzen, Aufmerksamkeitsdefizit-/Hyperaktivitätsstörung (ADHS), Begabung (Hohenstein et al., 2017) und Testangst, ADHS, Dyslexie, Essstörungen, geistige Fähigkeiten, Begabungen und verschiedene ethnische Hintergründe (Voss et al., 2011).

2.2 Empirische Zusammenhänge zu Lerngelegenheiten, Aspekten professioneller Kompetenz, der Performanz und dem Unterrichtserfolg von Lehrkräften

Das Wissen von (angehenden) Lehrkräften (hier v.a. GPK/PPK, Wissen über einzelne Diversitätsbereiche) und seine Bedeutung für Lehrkräfte und Schüler:innen wurde bereits in verschiedenen Stichproben im Kontext Schule (Studierende verschiedener Lehramtsstudiengänge, Referendar:innen, Lehrkräfte verschiedener Schulformen), mit und ohne direkten Inklusionsbezug untersucht. Vereinzelt liegen auch Studien zu den Fehlannahmen der Lehrkräfte vor. Den Annahmen des COACTIV-Modells (Kunter, Kleickmann et al., 2013; Voss et al., 2015) folgend, werden verschiedene Aspekte professioneller Kompetenz durch Lerngelegenheiten geformt. Die Performanz und der Unterrichtserfolg wiederum werden maßgeblich durch die professionelle Kompetenz

beeinflusst. Im Folgenden werden Studien berichtet, in denen Zusammenhänge zwischen dem Wissen/den Fehlannahmen und Lerngelegenheiten, Aspekten professioneller Kompetenz (v.a. Überzeugungen), der Performanz und dem Unterrichtserfolg der Lehrkräfte (vgl. Kunter, Kleickmann et al., 2013) untersucht wurden. Mit dem Unterrichtserfolg werden zusätzlich die Schüler:innen als Empfänger:innen schulischer Inklusion fokussiert. Viele dieser Studien sind korrelativer Natur, weshalb keine kausalen Schlüsse gezogen werden können. Weiterführende Überlegungen zu den Wirkrichtungen werden in der Diskussion vorgenommen (vgl. Kapitel 6.2.1 zur Einordnung in die professionelle Kompetenz im Rahmen der Interpretation der Ergebnisse).

Lerngelegenheiten. Lerngelegenheiten für (angehende) Lehrkräfte ergeben sich vor allem in der akademischen Grundausbildung und durch Praxiserfahrungen (Voss et al., 2015). Die Abiturnote spiegelt zwar nicht eine konkrete Lerngelegenheit in Bezug auf das Wissen über Diversität wider, schließt aber die Fähigkeit, Lerngelegenheiten in Leistung umzusetzen, zum Beispiel durch kognitive Kompetenzen und Lernmotivation (vgl. individuelle Eingangsvoraussetzungen; Kunter, Kleickmann et al., 2013), mit ein und ist der umfassendste Prädiktor von Studienerfolg (Wolf, 2017). So hing eine bessere Note mit einem höheren PPK allgemein (Hohenstein et al., 2017; Wildbrett et al., 2022) und einem höheren PPK speziell für inklusiven Unterricht (König et al., 2017) von Lehramtsstudierenden zusammen. Den Lerngelegenheiten im Studium entsprechend hatten (angehende) Lehrkräfte für sonderpädagogische Förderung in verschiedenen Stichproben und mit verschiedenen Tests gemessen ein höheres Wissen in dem jeweiligen Bereich als Lehrkräfte anderer Lehrämter (Heyder, Vaskova et al., 2018; Sanz-Cervera et al., 2017; Sousa et al., 2017). Ein fortgeschrittenes Studium ging ebenfalls mit einem höheren Wissen einher (Sanz-Cervera et al., 2017; Wildbrett et al., 2022). Auch wussten Studierende weniger über generelle pädagogische Inhalte, wenn diese (noch) nicht in ihrem Curriculum vorkamen, und Lehrkräfte hatten insgesamt größere Schwierigkeiten mit spezifischen theoretischen Inhalten (Malva et

al., 2020). Die Vorerfahrungen mit und der Kontakt zu Personen im Allgemeinen und Schüler:innen im Speziellen in Bezug auf verschiedene Diversitätsfacetten hingen positiv mit dem Wissen der (angehenden) Lehrkräfte zusammen (Sciutto et al., 2000; Yin et al., 2019). Zusätzliche Trainings scheinen darüber hinaus mit einem höheren Wissen einherzugehen (Sanz-Cervera et al., 2017; Schmiedeler, 2013; Yin et al., 2019).

Demgegenüber hing das Wissen über Hochbegabung nicht mit der beruflichen und privaten Kontakthäufigkeit mit hochbegabten Personen zusammen (Heyder, Bergold & Steinmayr, 2018), was auch für die Fehlannahmen galt. In einer Studie zum Wissen und den Fehlannahmen über die Autismusspektrumstörung von Lehramtsstudierenden ging ein fortgeschritteneres Studium mit mehr Fehlannahmen einher (Sanz-Cervera et al., 2017). Studierende für sonderpädagogische Förderung hatten allerdings weniger Fehlannahmen als Studierende anderer Lehramtsstudiengänge.

Aspekte professioneller Kompetenz. Da Einstellungen zu inklusiver Beschulung und Selbstwirksamkeitserwartungen im Umgang mit Inklusion als besonders wichtig für das Gelingen schulischer Inklusion gelten (de Boer et al., 2011; Fernandes et al., 2021; Lautenbach & Heyder, 2019; Wray et al., 2022), werden diese nun für die konzeptionelle Einordnung des Wissens über Diversität herangezogen. Sowohl für die Einstellungen (Ghanizadeh et al., 2006; Lu et al., 2020) als auch die Selbstwirksamkeitserwartungen (Sciutto et al., 2000) ergaben sich in bisherigen Studien positive Zusammenhänge zum Wissen über verschiedene Diversitätsbereiche von (angehenden) Lehrkräften. Schlussendlich korrelierte auch das pädagogische Wissen spezifisch für inklusiven Unterricht in einer Studierendenstichprobe positiv mit Überzeugungen zur adaptiven Förderung von Lernenden mit SPF im Regelunterricht, positiv mit Überzeugungsfacetten zur adaptiven Förderung von allen Lernenden im Unterricht (z. B. aktive Reaktionsform beim Umgang mit Diversität) und negativ mit anderen Überzeugungsfacetten (z. B. passive Reaktionsform beim Umgang mit Diversität, König et al., 2017). Dieses Ergebnis legt darüber hinaus bereits einen positiven

Zusammenhang zwischen dem Wissen und der Performanz nahe (vgl. Abschnitt Performanz).

Allerdings liegen auch im Bereich der Einstellungen und Selbstwirksamkeitserwartungen Studien vor, in denen keine signifikanten Zusammenhänge gefunden wurden: So korrelierten die Einstellungen zur Förderung hochbegabter Schüler:innen nicht mit dem Wissen über Hochbegabung (Heyder, Bergold & Steinmayr, 2018). Auch das PPK hing zwar mit den unterrichtsbezogenen, nicht aber den allgemeinen Selbstwirksamkeitserwartungen von Lehrkräften zusammen (Laueremann & König, 2016). Auch das generelle PPK von Lehramtsstudierenden hing nicht mit ihren Selbstwirksamkeitserwartungen bezogen auf Unterrichtsstrategien, Klassenmanagement und die Einbeziehung der Schüler:innen zusammen (Depaepe & König, 2018). Zusammenfassend deuten empirische Befunde auf positive Zusammenhänge zwischen dem Wissen und den Einstellungen und Selbstwirksamkeitserwartungen hin, es liegen aber auch vereinzelt Studien vor, in denen keine signifikanten Zusammenhänge gefunden wurden.

In Bezug auf die Fehlannahmen liegt nur ein Befund vor, nach dem Lehrkräfte mit mehr Fehlannahmen über Hochbegabung eine niedrigere Einstellung zur Förderung hochbegabter Schüler:innen aufwiesen (Heyder, Bergold & Steinmayr, 2018). Zu den Selbstwirksamkeitserwartungen und Fehlannahmen von Lehrkräften gibt es bislang keine empirischen Studien. Eine neuere Studie legt darüber hinaus einen Zusammenhang zwischen Fehlannahmen und Stereotypen nahe: Lehramtsstudierende, die glaubten, dass Mathematik ein eher maskulines Fach ist, hatten mehr Fehlannahmen über Geschlechterunterschiede in Bezug auf dieses Fach (Dersch et al., 2022).

Performanz. Ein hohes Wissen sollte sich in einer hohen Performanz widerspiegeln (Voss et al., 2015). Mit dieser Annahme in Einklang stehen Befunde zur Unterrichtsqualität, die Vorteile von einem hohen Lehrkräfte-Wissen für Schüler:innen nahelegen: Das PPK (inklusive einer Skala über die Heterogenität von Schüler:innen) von Lehramtsstudierenden korrelierte positiv mit der kognitiven Aktivierung und dem Unterrichtstempo aus Sicht der

unterrichteten Schüler:innen (Voss et al., 2011). Dieses Ergebnis konnte mit einer Stichprobe aus Referendar:innen für die schülerberichteten Facetten Unterstützung der Schüler:innen, weniger Störungen im Klassenzimmer und Monitoring repliziert werden (Voss et al., 2022). Bei Lehrkräften mit einem hohen Wissen variierte die Unterrichtsqualität darüber hinaus weniger in Abhängigkeit der Klasse. Auch das GPK von Berufsschullehrkräften hing positiv mit allen Facetten der schülerberichteten Unterrichtsqualität zusammen (König & Pflanzl, 2016). In einer weiteren Studie mit Lehramtsstudierenden hingegen konnte das GPK zwar die selbstberichteten Unterrichtsinstruktionen in Bezug auf die Unterstützung beim Lernen, nicht aber in Bezug auf die kognitive Aktivierung und das Klassenmanagement vorhersagen (Depaepe & König, 2018). Das GPK von Gymnasiallehrkräften hing positiv mit den kompetenzorientierten Unterrichtspraktiken zusammen, die mittels Videoaufnahmen analysiert wurden (Gess-Newsome et al., 2017). Auch das pädagogische Unterrichtswissen (allerdings nur die Dimensionen Strukturierung von Unterricht und Klassenführung, nicht aber Heterogenität) von Berufsschullehrkräften korrelierte positiv mit der Methoden-, Fach- und Beziehungskompetenz der Lehrkräfte und der Anwendung von Lernstrategien der Schüler:innen (Einprägungsstrategien, Anstrengung und Ausdauer) aus Sicht der Schüler:innen (Pflanzl et al., 2013).

Befunde zu den Fehlannahmen und der Performanz von Lehrkräften liegen bislang nicht vor.

Unterrichtserfolg. Der Unterrichtserfolg kann sich sowohl auf Lehrkräfteebene (z. B. Berufserfolg, berufliches Wohlbefinden) als auch auf Ebene der Schüler:innen (z. B. Lernerfolg, motivationale und emotionale Entwicklung) manifestieren (vgl. Kunter, Kleickmann et al., 2013; Voss et al., 2015). Auf Ebene der Lehrkräfte konnte das GPK (inklusive des Wissens über den Umgang mit der Diversität der Schüler:innen) von Grundschul- und Gesamtschullehrkräften ihr Empfinden von Burnout negativ vorhersagen. Das galt sowohl für den direkten Zusammenhang zwischen dem Wissen und der Skala

Depersonalisation als auch für den indirekten Zusammenhang mit allen drei Burnoutskalen (Depersonalisation, emotionale Erschöpfung, Leistungsminderung) über die Selbstwirksamkeitserwartungen (Lauermann & König, 2016). Auch in einer Längsschnittstudie mit Lehramtsstudierenden während und nach dem Ende ihres Referendariats konnte das bildungswissenschaftliche Wissen die Veränderung in der emotionalen Erschöpfung der Teilnehmenden vorhersagen. Das Wissen konnte hier eine Verschlechterung abpuffern: Je größer das Wissen war, desto weniger verschlechterte sich die emotionale Erschöpfung (Dicke et al., 2015). Auch das Wissen über Klassenführung konnte die emotionale Erschöpfung (negativ) und die Arbeitszufriedenheit (positiv) von Lehramtskandidat:innen in ihrem Vorbereitungsdienst längsschnittlich vorhersagen (Klusmann et al., 2012). Vor dem Hintergrund, dass die emotionale Erschöpfung von Lehrkräften negativ mit den Leistungen ihrer Schüler:innen zusammenhängt (Klusmann et al., 2016), und in Einklang mit den genannten Zusammenhängen zwischen dem Wissen der Lehrkräfte und der Unterrichtsqualität (z. B. König & Pflanzl, 2016), erscheinen auch Zusammenhänge zwischen dem Wissen und dem Lernerfolg der Schüler:innen plausibel.

Zwar gibt es bereits Forschung, die positive Zusammenhänge zwischen dem fachdidaktischen Wissen zur Vermittlung eines speziellen Fachs und dem Lernerfolg der Schüler:innen herstellt (Baumert & Kunter, 2013b, Keller et al., 2017; Kunter, Klusmann et al., 2013), doch in Bezug auf das PPK/GPK und Wissen über Diversität gibt es bislang kaum Erkenntnisse (vgl. Voss et al., 2015). Einige Studien legen einen Zusammenhang zwischen der adaptiven Lehrkompetenz der Lehrkräfte, eine Facette des handlungsnahen pädagogischen Wissens (Voss et al., 2015), und den Leistungen der Schüler:innen nahe (Beck et al., 2008; Brühwiler, 2014). Dabei erwies sich der Zusammenhang als unabhängig von den kognitiven Lernvoraussetzungen, dem Geschlecht, der sozialen Herkunft und dem Sprachhintergrund der Schüler:innen (Brühwiler, 2014). Besonders wirksam war eine hohe adaptive Lehrkompetenz in leistungsheterogenen Klassen, was die Bedeutung für schulische Inklusion hervorhebt. In

einer Studie mit Gymnasiallehrkräften zeigte sich demgegenüber kein Zusammenhang zwischen deren GPK und den Leistungen der Schüler:innen (hier Wissen im Fach Biologie, Gess-Newsome et al., 2017). Zum Diversitätswissen liegen nach aktueller Recherche keine Erkenntnisse vor. Dies gilt auch für die Fehlannahmen von (angehenden) Lehrkräften.

Fazit. Die genannten Studien zum PPK/GPK und zum diversitätsbezogenen Wissen legen positive Zusammenhänge mit formalen Lerngelegenheiten wie dem Studienfach (Sousa et al., 2017) und zusätzlichem Training (Sanz-Cervera et al., 2017), den Einstellungen (Lu et al., 2020), den Selbstwirksamkeitserwartungen (Sciutto et al., 2000) und der Unterrichtsqualität (Gess-Newsome et al., 2017) und negative Zusammenhänge mit Burnout (Dicke et al., 2015) nahe. Zum Zusammenhang zwischen dem Wissen und dem Lernerfolg bzw. der motivationalen und emotionalen Entwicklung der Schüler:innen liegen bislang kaum Studien vor, weshalb kein eindeutiges Fazit gezogen werden kann. Dies gilt in Bezug auf die genannten Wissensbereiche vor allem für das Wissen über Diversität, das bislang kaum erforscht ist, was auf einen Mangel an umfassenden Messinstrumenten zurückzuführen sein könnte. Im Folgenden werden bisherige Tests vorgestellt, die das Wissen über Diversität in unterschiedlichem Maße (als Subskala in einem übergeordneten Test, als Test zu einer einzelnen Diversitätsdimension) abbilden. Darauf aufbauend wird in Kapitel 3 ein neues Messinstrument vorgestellt.

2.3 Bisherige Tests zur Erfassung von Wissen über Diversität

Um das Wissen über Diversität erfassbar zu machen, wurden bereits einige Messinstrumente entwickelt, die sich hinsichtlich ihrer Konzeptualisierung stark unterscheiden. So liegen im englisch- und deutschsprachigen Raum Tests zum PPK mit Diversität als Facette (z. B. Voss et al., 2011) und Tests zu einzelnen Diversitätsdimensionen (z. B. Heyder, Bergold & Steinmayr, 2018) vor. Es kommen Vignetten (= Beschreibung von Beispielsituationen; z. B. König et al., 2017) und inhaltlich eindeutig richtige bzw. falsche, allgemein gehaltene Aussagen (z. B. Haimour & Obaidat, 2013) zum Einsatz, wobei es Tests

gibt, die nur eine Art Itemformat nutzen (z. B. Sousa et al., 2017) und andere, in denen verschiedene Formate eingesetzt werden (z. B. Voss et al., 2011). Ebenso liegen verschiedene Antwort-Formate vor, wobei es auch hier Tests gibt, die nur eine Art Antwortformat nutzen (z. B. Sciutto et al., 2000) und andere, in denen verschiedene Formate eingesetzt werden (z. B. König & Blömeke, 2009). Es kommen offene (z. B. König & Blömeke, 2009), Single-Choice- (z. B. König et al., 2017), Multiple-Choice-Antworten (z. B. Kunina-Habenicht et al., 2020), Likertskalen (z. B. Wildbrett et al., 2022) und Richtig-Falsch-Antworten (z. B. Ghanizadeh et al., 2006) vor, wobei zum Teil eine zusätzliche Weiß-nicht-Option (z. B. Soriano-Ferrer & Echegaray-Bengoa, 2014) gegeben wird. Neben den im Folgenden beschriebenen Tests wird das Wissen in einigen Studien auch über die Angabe zu besuchten inklusionsbezogenen Lehrveranstaltungen (Schwab et al., 2017) oder im Selbstbericht erfasst, entweder mit einem Fragebogen mit mehreren Subskalen (z. B. Kuyini & Desai, 2007) oder mit einem Item (z. B. Heyder, Bergold & Steinmayr, 2018). Teilweise wird keine detaillierte Angabe zur Erfassung des subjektiven Wissens gemacht (z. B. Schmiedeler, 2013). Da in diesen Studien das subjektive und objektiv gemessene Wissen gar nicht (z. B. Heyder, Bergold & Steinmayr, 2018) oder nur moderat positiv (z. B. Schmiedeler, 2013) miteinander korrelieren und Lehramtsstudierende ihr Wissen nicht akkurat einschätzen können (Ernst et al., 2023), ist die Nutzung objektiver Tests wichtig für die Erfassung des Wissens von (angehenden) Lehrkräften. Die folgenden Unterkapitel stellen den Versuch einer Systematisierung der sehr unterschiedlichen Messinstrumente dar, die bereits existieren. Es handelt sich dabei nicht um eine erschöpfende Auflistung, sondern um eine Auswahl an für diese Arbeit inhaltlich und methodisch relevanten Tests. Tests, die zur Evaluation einer spezifischen Intervention entwickelt wurden und keiner explizit beschriebenen Prüfung der Gütekriterien unterzogen wurden (z. B. Autismus, Hart & More, 2013) oder nur aus einzelnen wenigen Items bestehen (z. B. Angststörungen, Headley, 2013), werden zum Beispiel nicht beschrieben. Aus der Recherche ging nur ein Test zur expliziten Erfassung von

Fehlannahmen hervor (Dersch et al., 2022), der im Rahmen der Tests zu einzelnen Diversitätsdimensionen vorgestellt wird.

2.3.1 Tests zum pädagogisch-psychologischen Wissen mit Diversität als Facette

Das PPK ist bereits über verschiedene Messinstrumente erfassbar. Die sieben folgenden Tests sind alle dem PPK zuzuordnen und enthalten mit unterschiedlich starkem Bezug zu Inklusion Skalen bzw. Items zur Diversität der Schüler:innen. Sie unterscheiden sich in den genutzten Item- und Antwortformaten, sodass alle eingangs genannten in den Tests zu finden sind.

Ein sehr umfangreicher Test mit 44 Items liegt zum pädagogischen Wissen für inklusiven Unterricht (General Pedagogical Knowledge for Inclusive Teaching – GPK-IT; König et al., 2017) mit den beiden Anforderungsdimensionen Diagnose (Wissen über Lernprozesse, Dispositionen/Unterschiede und methodisches Wissen über Diagnose) und Intervention (Wissen über Klassenführung, Strukturierung und Binnendifferenzierung/Individualisierung) vor. Es kommen sowohl Vignetten als auch allgemeine Fragen zum Einsatz, die im Single-Choice-Format mit jeweils 3 Distraktoren beantwortet werden. Bei der Evaluation in einer Studierendenstichprobe lag die Expected A Posteriori Estimation zur Schätzung der internen Konsistenz des Gesamttests bei einem noch akzeptablen Wert von .60, wobei sie in den verschiedenen Substichproben zum Teil geringer ausfiel (EAP = .51-.67). Im Rahmen der Vergleichsstudie TEDS-M (Teacher Education and Development Study – Learning to Teach Mathematics) entstand ein Test zum pädagogischen Unterrichtswissen in einer Lang- (König & Blömeke, 2009) und einer Kurzform (König & Blömeke, 2010) mit 5 Subskalen, die sowohl deklaratives als auch prozedurales Wissen erfassen. Die Skala zum Umgang mit Heterogenität erfasst das Wissen über Differenzierungsmaßnahmen und Methodenvielfalt im Unterricht (z. B. Vor- und Nachteile offener Unterrichtskonzepte) mit 15 bzw. 4 Items. Es kommen alle eingangs genannten Item- und Antwortformate vor. Die Schätzung der internen Konsistenz der Langversion in einer

Studierendenstichprobe war mit $EAP = .85$ sehr gut (König & Blömeke, 2009). Ein vignettenbasierter Test zum pädagogisch-psychologischen Wissen (PPK-V; Wildbrett et al., 2022) mit 3 Subskalen ermöglicht eine handlungsnah und gleichzeitig ökonomische Erfassung des Wissens. Die beiden Skalen Lernen und Lernende (z. B. individuelle Lernprozesse, Umgang mit Heterogenität) und Diagnostik und Evaluation (z. B. Diagnostik bei Lernstörungen) enthalten Items mit Bezug zu der Diversität der Schüler:innen. McDonalds Omega zur Schätzung der Reliabilität war in einer Studierendenstichprobe für den Generalfaktor PPK akzeptabel ($\omega = .73$), für die beiden diversitätsbezogenen Subskalen hingegen eher niedrig (Lernen und Lernende: $\omega = .54$, Diagnostik und Evaluation: $\omega = .61$). Pro Vignettenstamm liegen 5 bis 6 Handlungsoptionen vor, die auf einer 6-stufigen Likertskala beurteilt werden.

In einem Test zum pädagogisch-psychologischen Professionswissen von Lehrkräften (ProwiN-Test; Lenske et al., 2015) werden die beiden Skalen zum deklarativen und konditional-prozeduralen Wissen anhand der Wissensart differenziert. Die fünf Inhaltsbereiche, von denen einer die Individualisierung darstellt, sind pro Skala mit mindestens einem Item vertreten. Die relativ geringe Itemanzahl pro Inhaltsbereich erlaubt keine Bildung einer Skala zur Individualisierung, und die Studienergebnisse geben keine differenzierte Auskunft über diesen Teilbereich, dessen isolierte Erfassung nicht Teil der Studienziele war. Jede der 28 Aufgaben besteht aus einem Itemstamm, der als allgemeine Frage formuliert ist, und mehreren Antwortoptionen, zu denen die Zustimmung der Teilnehmenden auf einer vierstufigen Skala ($1 = \text{sehr} - 4 = \text{gar nicht}$) erfasst wird, die im Zuge der Auswertung wieder dichotomisiert wird. Dieses Antwortformat unterscheidet sich von vielen anderen Tests, da hier nicht davon ausgegangen wird, dass die Aussagen ohne Einschränkung als eindeutig richtig oder falsch bewertet werden können.

Im Zuge einer Expert:innenbefragung zu bildungswissenschaftlichen Inhalten der universitären Lehrer:innenbildung wurden 104 Themen aus 9 Inhaltsbereichen extrahiert (u.a.

Lernprozesse, Sozialisationsprozesse, Heterogenität/soziale Konflikte; Kunina-Habenicht et al., 2012). Daraus entstand ein umfangreicher revidierter Test mit den Dimensionen Unterrichtsgestaltung, Lernen und Entwicklung, Diagnostik und Evaluation, Bildungstheorie, Schule als Bildungsinstitution und Lehrerberuf als Profession in einer Lang- und Kurzform mit 119 bzw. 65 Items (BilWiss-2.0-Test; Kunina-Habenicht et al., 2020). Zwar gibt es auch hier einzelne Items mit Bezug zur Diversität der Schüler:innen (z. B. Heterogenität bzgl. des Geschlechts/sozialen Hintergrunds), allerdings gibt es keine explizite Skala dazu. Es kommen Multiple-Choice-Aufgaben mit jeweils vier Antwortalternativen zum Einsatz.

Des Weiteren gibt es zwei Tests, die im Rahmen des PPK verschiedene Diversitätsfacetten zu einer Skala zusammenfassen. Ein Test zum pädagogisch-psychologischen Wissen in der Lehramtsausbildung (Hohenstein et al., 2017) enthält 10 Subskalen, von denen eine das Wissen über die Bedeutung von und den Umgang mit individuellen Besonderheiten der Lernenden erfasst. Dazu gehören verschiedene Diversitätsfacetten (Vorwissen, Intelligenz, Interesse, sozioökonomischer Status, Lern- und Entwicklungsstörungen, Geschlechterdifferenzen, ADHS, Begabung). Es kommen Vignetten und eindeutig richtige bzw. falsche Aussagen zum Einsatz, und für die Beantwortung werden Multiple-Choice- und Richtig-Falsch-Antwortformate genutzt. Bei der Erprobung des Tests in einer Stichprobe aus Lehramtsstudierenden lag die interne Konsistenz der Skala zu individuellen Besonderheiten bei $\alpha = .30$, weshalb sie aus den Analysen ausgeschlossen wurde und die Items in Folgeuntersuchungen überarbeitet werden. Ein weiterer Test zum pädagogisch-psychologischen Wissen (Voss et al., 2011) mit 4 Subskalen enthält eine Skala zur Heterogenität der Schüler:innen mit 9 Items über die Identifizierung von und den Umgang mit individuellen Schwierigkeiten und Begabungen (Testangst, ADHS, Dyslexie, Essstörungen, geistige Fähigkeiten und Begabungen, verschiedene ethnische Hintergründe). Die Items bestehen aus Vignetten und eindeutig richtigen bzw. falschen allgemeinen Aussagen, deren Beantwortung offen oder im Multiple-Choice-Format vorgenommen wird. In

einer Erprobung des Tests mithilfe einer Referendar:innen-Stichprobe war die interne Konsistenz der Skala zur Heterogenität mit $\alpha = .52$ sehr niedrig, und die Autor:innen empfehlen für die Zukunft die Nutzung mehrerer Subskalen, um die heterogenen Inhaltsfelder skalenspezifisch abzubilden.

Zusammenfassend lässt sich festhalten, dass die bereits bestehenden Tests wichtige Ansätze zur Erfassung des Wissens über Diversität liefern und gleichzeitig mit einigen Nachteilen einhergehen (die zum Teil auch daher rühren, dass die Erfassung von Diversität nicht erklärtes Ziel der entsprechenden Studien war). Der GPK-IT (König et al., 2017) ist eng an die Anforderungen schulischer Inklusion geknüpft und ermöglicht eine differenzierte Erfassung verschiedener Anforderungsdimensionen. Hinsichtlich der gewählten Item- und Antwortformate ist er eher aufwändig. Der Fokus liegt auf der Diagnose und Intervention, und verschiedene Diversitätsdimensionen werden nicht explizit in den Blick genommen. Eine ökonomischere Erfassung von Diversität bietet die Subskala zu Heterogenität des Tests zum pädagogischen Unterrichtswissen (König & Blömeke, 2009, 2010). Er werden verschiedene Diversitätsdimensionen zusammengefasst, innerhalb derer aber nicht weiter differenziert wird. Beide Verfahren erwiesen sich als reliabel. Der ökonomische vignettenbasierte Test (Wildbrett et al., 2022) enthält Items mit direktem Bezug zu Diversität, die allerdings zwei verschiedenen Skalen zugeordnet sind. Der ProwiN- (Lenske et al., 2015) und BilWiss-2.0-Test (Kunina-Habenicht et al., 2020) beinhalten zwar Items zur Diversität der Schüler:innen, allerdings kann daraus keine Subskala gebildet werden. Die beiden Diversitätsskalen aus den Tests von Hohenstein et al. (2017) und Voss et al. (2011) ermöglichen aktuell keine reliable Erfassung aufgrund einer niedrigen internen Konsistenz. Aus den vorgestellten Tests lässt sich festhalten, dass das pädagogisch-psychologische Wissen trotz Nutzung einzelner Subskalen ein eher heterogenes Konstrukt ist (Wildbrett et al., 2022). Um dieser Heterogenität gerecht zu werden und Diversität reliabel abbilden zu können, können Tests zu einzelnen Diversitätsdimensionen einen wichtigen Beitrag leisten (Voss et al., 2011). Solche

Tests werden im Folgenden vorgestellt.

2.3.2 Tests zu einzelnen Diversitätsbereichen

Zu einigen Diversitätsdimensionen der gerade beschriebenen allgemeinen Tests liegen bereits spezifische Tests vor, die zum Teil sehr ähnlich strukturiert sind und sich als reliable Ergänzung zu den oben beschriebenen Subskalen erweisen. Die meisten dieser Tests haben die Nutzung allgemeiner Aussagen und das Richtig-Falsch-Antwortformat mit Weiß-nicht-Option gemeinsam, wobei sie sich zum Teil in der Auswertung der Antworten unterscheiden. In den meisten Tests werden das Wissen und die Fehlannahmen getrennt voneinander ausgewertet (Wissen: $1 = \text{korrekte Antwort}$, $0 = \text{inkorrekte/weiß-nicht Antwort}$; Fehlannahmen: $1 = \text{inkorrekte Antwort}$, $0 = \text{korrekte/weiß-nicht Antwort}$; z. B. Heyder, Bergold & Steinmayr, 2018; Scituro et al., 2000), wobei die Auswertung der Fehlannahmen nicht in jeder Studie auch tatsächlich vorgenommen wird (z. B. Yin et al., 2019). Alternativ findet eine Auswertung mittels dreistufiger Skala statt ($1 = \text{inkorrekte Antwort}$, $2 = \text{weiß-nicht Antwort}$, $3 = \text{korrekte Antwort}$; z. B. Sousa et al., 2017). Andere als die beschriebenen Formate werden im Folgenden explizit benannt. Ein aktueller Test wurde außerdem zur Erfassung von Fehlannahmen konzipiert (Dersch et al., 2022).

Es liegen verschiedene Tests zum Wissen über Dyskalkulie und Dyslexie vor, die jeweils in drei Subskalen unterteilt sind: Das Wissen über Merkmale (10 Items, $\alpha = .66$), Symptome (19 Items, $\alpha = .88$) und Interventionsstrategien (11 Items, $\alpha = .90$) bei Dyskalkulie wurde in einer Stichprobe aus Grundschullehrkräften evaluiert (40 Items gesamt, Sousa et al., 2017). Das Wissen über allgemeine Informationen (17 Items, $\alpha = .69$), Symptome/Diagnose (32 Items, $\alpha = .64$) und Behandlung (9 Items, $\alpha = .67$) bei Dyslexie wurde in einer Stichprobe aus Grundschullehrkräften evaluiert (36 Items gesamt, $\alpha = .76$, Soriano-Ferrer & Echegaray-Bengoia, 2014) und in verschiedenen Versionen und Sprachen eingesetzt (z. B. Dodur & Kumaş, 2020; Yin et al., 2019).

Ähnlich aufgebaut ist auch ein Test zum Wissen über ADHS mit den Skalen zu

allgemeinen Informationen (15 Items, $\alpha = .71$), Symptomen/Diagnose (9 Items, $\alpha = .71$) und Behandlung (12 Items, $\alpha = .71$), der nach der Evaluation in einer Stichprobe aus Grundschullehrkräften (36 Items gesamt, $\alpha = .86$; Knowledge of Attention Deficit Disorders Scale, KADDS; Sciotto et al., 2000) vielfach eingesetzt wurde (z. B. Perold et al., 2010; Schmiedeler, 2013; Sciotto et al., 2016).

Ein Test zum Wissen über kognitive Hochbegabung mit 38 Items ($\alpha = .79$) ohne separate Subskalen, der ebenfalls die Charakteristiken dieser Schüler:innen fokussiert, wurde in einer Stichprobe aus Gymnasial- und Gesamtschullehrkräften getestet (Heyder, Bergold & Steinmayr, 2018) und noch nicht in weiteren Stichproben eingesetzt. Items aus diesem Test finden sich in dem in Kapitel 3 beschriebenen Wissenstest wieder.

Zu dem Bereich Geschlecht liegt nach aktueller Recherche nur ein Test vor, der explizit Mathe-Geschlechter-Fehlannahmen, nicht aber das Wissen von (angehenden) Lehrkräften erfasst (Math-Gender Misconception Questionnaire, MGMQ; Dersch et al., 2022). Die 15 Items repräsentieren gängige Fehlannahmen aus den Bereichen Einfühlen–Systematisieren ($\omega = 0.88$), Kompensation der Mädchen ($\omega = 0.76$) und Nicht-Kompensierbarkeit der Mädchen ($\omega = 0.72$). Ausgeglichen wird der Test durch 15 Füllitems, die korrekt sind und nicht ausgewertet werden. Das Richtig-Falsch-Antwortformat wird in diesem Test durch eine Angabe der Antwortsicherheit ergänzt, sodass durch die Multiplikation der Antwort ($1 = richtig$, $-1 = falsch$) mit der Antwortsicherheit ($0 = sehr unsicher$, $4 = sehr sicher$) und der Bildung eines Mittelwerts intervallskalierte Werte für die einzelnen Subskalen berechnet werden können.

Außerdem gibt es einen Test zum Wissen über die Autismusspektrumstörung (30 Items, $\alpha = .93$) mit einem Fokus auf die Charakteristiken dieser Schüler:innen, der in einer Stichprobe aus Sonder- und Regelschullehrkräften erprobt (Autism Knowledge Questionnaire, AKQ; Haimour & Obaidat, 2013) und bereits auf Spanisch übersetzt und bei Lehramtsstudierenden evaluiert wurde (Sanz-Cervera et al., 2017). Zur internen Konsistenz

von acht Items zur Messung des Wissens über die Autismusspektrumstörung werden in einer Studie mit Grundschullehrkräften keine Angaben gemacht (Ghanizadeh et al., 2006). Vergleichbare Items kommen in einer anderen Studie mit Gymnasiallehrkräften mit offenen, Single-Choice- und Ja-Nein-Antwortformaten vor (Brook et al., 2000), wobei letztere vergleichbar mit dem Richtig-Falsch-Antwortformat sind. In der gleichen Studie werden 12 Items zur Messung des Wissens über Lernschwierigkeiten eingesetzt. Auch hier werden keine Angaben zur internen Konsistenz gemacht. Dies war auch der Fall in einer Studie mit Lehrkräften an Grundschulen und weiterführenden Schulen zu einem Wissenstest über Angststörungen und evidenzbasierte Strategien zum Abbau von Ängsten (Teacher Knowledge of Anxiety Assessment, TKA; Ginsburg et al., 2021). Die 15 Items wurden im offenen und Multiple-Choice-Format beantwortet und basieren zum Teil auf Fallvignetten.

Die hier beschriebenen Tests behandeln Dyskalkulie (Sousa et al., 2017), ADHS (Sciutto et al., 2000), Geschlecht (Dersch et al., 2022), Hochbegabung (Heyder, Bergold & Steinmayr, 2018), Autismusspektrumstörung (Haimour & Obaidat, 2013) und Angststörungen (Ginsburg et al., 2021) und bilden damit einen Teilbereich der im Schulwesen relevanten Diversitätsbereiche ab. Nicht in allen Studien werden Angaben zur internen Konsistenz gemacht. Für die Tests, bei denen die interne Konsistenz untersucht wurde, werden meist mindestens akzeptable Werte berichtet. Es wird jeweils ein Diversitätsbereich abgebildet, und in keinem der Tests werden mehrere Diversitätsbereiche in den Blick genommen.

3 Der vorliegende Wissenstest

In Ergänzung zu den oben genannten Tests, wurde an der TU Dortmund im Rahmen des Projekts DoProfiL (Dortmunder Profil für inklusionsorientierte Lehrer:innenbildung) ein Test zum Wissen über Diversität mit den Subskalen Förderschwerpunkt ESE, Förderschwerpunkt Lernen, Teilleistungsstörungen, Geschlecht und kognitive Hochbegabung entwickelt (für Details zur ersten Veröffentlichung der Entwicklung und Pilotierung s. Heyder, Vaskova et al., 2018).

3.1 Bedeutung der Diversitätsbereiche

Die Auswahl der fünf Diversitätsbereiche erfolgte auf der Basis bildungswissenschaftlicher Standards für die Lehramtsausbildung (KMK, 2014), einer Delphi-Studie zu bildungswissenschaftlichen Inhalten in der Lehrer:innenbildung (Kunina-Habenicht et al., 2012), empirischen Forschungsarbeiten zu Einstellungen von Lehrkräften (z. B. de Boer et al., 2011) und der Prävalenz von Kindern mit SPF (KMK, 2016). Die Auswahl der Diversitätsbereiche ist nicht erschöpfend und hat somit nicht den Anspruch, Diversität in seiner Gesamtheit abzubilden. Schüler:innen können sich noch in vielen weiteren Merkmalen unterscheiden (vgl. Faulstich-Wieland, 2011; Grosche, 2015; Zulliger & Tanner, 2013), und eine Erweiterung der Testinhalte ist langfristig denkbar. Im Folgenden werden die inhaltlichen Bereiche definiert, und es werden empirische Befunde zu betroffenen Schüler:innen und ihren Lehrkräften vorgestellt.

3.1.1 Sonderpädagogischer Förderbedarf emotionale-soziale Entwicklung & Lernen

Obwohl berechtigte Kritik an den teils schwammigen Kategorien zum SPF besteht (Grosche, 2015), bilden diese die aktuelle Situation an deutschen allgemeinbildenden Schulen, Förderschulen und Schulen für Kranke ab (MSB NRW, 2016) und sind wichtig für die Zuteilung von personellen Ressourcen (MSB NRW, 2022). In Deutschland wurden 2020

knapp 600.000 Schüler:innen sonderpädagogisch gefördert (wobei auch Schüler:innen ohne abgeschlossenes Feststellungsverfahren mitgezählt wurden). Damit bildete diese Schüler:innengruppe 7.7% aller Schüler:innen im Alter der Vollzeitschulpflicht ab. Der SPF Lernen bildete dabei mit 39.2% den größten Bereich ab. Der SPF ESE war mit 17.8% am zweithäufigsten vertreten, und in diesem Bereich ist seit 2011 ein besonders starker Anstieg zu verzeichnen (KMK, 2022). Nach den aktuellsten Zahlen aus dem Schuljahr 2016/2017 hatten europaweit 4.4% aller Schüler:innen einen festgestellten SPF (Ramberg et al., 2018). Die Zahlen schwanken von 1.1% bis 20.5% für die einzelnen Länder deutlich, was zum Teil daran liegt, dass die Definitionen und Feststellungsverfahren unterschiedlich sind und es keine europa- oder weltweit allgemeingültigen Definitionen gibt. Dies führt vor allem für den SPF ESE in der internationalen Literatur zu uneinheitlichen Bezeichnungen, die entweder soziale, emotionale und behaviorale Schwierigkeiten oder nur zwei der genannten Bereiche mit einschließen (vgl. Banks et al., 2017; Nelson et al., 2004). In Deutschland können Lern- und Entwicklungsstörungen, die mit erheblichen Beeinträchtigungen einhergehen, zu einem Bedarf an sonderpädagogischer Unterstützung führen, wobei dieser Bedarf einen oder mehrere Förderschwerpunkte betreffen kann (MSB NRW, 2016). Ein Bedarf an sonderpädagogischer Unterstützung im Förderschwerpunkt ESE besteht nach der AO-SF in Nordrhein-Westfalen, wenn „sich eine Schülerin oder ein Schüler der Erziehung so nachhaltig verschließt oder widersetzt, dass sie oder er im Unterricht nicht oder nicht hinreichend gefördert werden kann und die eigene Entwicklung oder die der Mitschülerinnen und Mitschüler erheblich gestört oder gefährdet ist“ (MSB NRW, 2016, § 4 [4]). Ein Bedarf an sonderpädagogischer Unterstützung im Förderschwerpunkt Lernen besteht, wenn „Lern- und Leistungsausfälle schwerwiegender, umfänglicher und langdauernder Art“ (MSB NRW, 2016, § 4 [2]) vorliegen.

Bei der Erforschung schulischer Inklusion mit Bezug zu den Förderschwerpunkten werden häufig alle Förderschwerpunkte zusammen betrachtet (z. B. Huber & Wilbert, 2012;

Krull et al., 2014), wobei der SPF Lernen und zum Teil auch der SPF ESE in Anlehnung an die Prävalenz an deutschen Schulen meistens den größten Anteil ausmachen (z. B. Henke et al., 2017; Schwab et al., 2015). Vereinzelt liegen auch Studien vor, die nur einen isolierten SPF in den Blick nehmen (ESE, Crede et al., 2019). Wie auch in offiziellen Statistiken (Ramberg et al., 2018; KMK, 2022) werden in der Forschung zum Teil unterschiedliche Voraussetzungen für die Annahme eines SPF definiert. Im deutschsprachigen Raum wird der SPF meistens durch die Lehrkraft angegeben. Zugrunde gelegt wird dabei für eine dichotome Erfassung (kein SPF vs. SPF) zum Beispiel die subjektive Einschätzung der Lehrkraft, wenn an den Erhebungsschulen keine sonderpädagogischen Gutachten vorgenommen werden (Krull et al., 2014), oder ein abgeschlossenes Feststellungsverfahren nach AO-SF (Henke et al., 2017; Schwab et al., 2015; Spörer et al., 2015). In Anlehnung an die Realität an Schulen in NRW kommt alternativ eine dreistufigen Skala zum Einsatz, in der festgehalten wird, ob die betreffenden Schüler:innen bereits eine Diagnostik durchlaufen hatten, ob ein Verfahren eingeleitet, aber noch nicht abgeschlossen wurde, oder ob die Schüler:innen präventiv als Schüler:innen mit SPF im inklusiven Unterricht beschult wurden (Crede et al., 2019).

Obwohl Überblicksartikel neutrale bis positive Effekte inklusiver Beschulung auf die Schulleistungen von Schüler:innen mit SPF hindeuten (Ruijs & Peetsma, 2009), scheinen diese weiter hinter ihren Peers ohne SPF zurückzuliegen. Schüler:innen mit SPF ESE zeigten zum Beispiel schlechtere Lese-, Rechtschreib- und Matheleistungen im Vergleich zu Schüler:innen ohne SPF (Nelson et al., 2004). Auch gemessen in Noten schnitten diese Schüler:innen schlechter ab (Sainio et al. 2019). In Einklang mit ihrer Diagnose lagen auch Schüler:innen mit SPF Lernen (ebenso wie solche mit SPF ESE) in der PISA Vergleichsstudie in den Lese- und Matheleistungen hinter ihren Peers zurück (Gebhardt et al., 2015).

Schüler:innen mit SPF und geringen Schulleistungen an Grundschulen haben ein erhöhtes Risiko, sozial ausgegrenzt zu werden. Je ausgeprägter der Unterstützungsbedarf und

je geringer die Schulleistung, desto geringer nahmen die Schüler:innen ihre soziale Integration in der Klasse und das Angenommensein durch die Lehrkraft wahr (Huber & Wilbert, 2012). Auch unabhängig von den Schulleistungen zeigten sich Unterschiede zuungunsten der Schüler:innen mit SPF in der sozialen Integration (Crede et al., 2019; Krawinkel et al., 2017; Krull et al., 2014), in der Beliebtheit (Krawinkel et al., 2017), im schulischen Selbstkonzept (Crede et al., 2019; DeVries et al., 2018; Krull et al., 2014), im Klassenklima (Crede et al., 2019) und in den Freundschaften mit Peers (Banks et al., 2017). Dementgegen stehen Befunde, nach denen keine Unterschiede in der Schuleinstellung (Crede et al., 2019; Schwab et al., 2015), in schulischen/sozialen Selbstkonzepten (Spörer et al., 2015), im wahrgenommenen Klassenklima (Henke et al., 2017; Krull et al., 2014), in der sozialen Integration (DeVries et al., 2018; Henke et al., 2017) und im Gefühl des Angenommenseins durch die Lehrkraft (Crede et al., 2019; Henke et al., 2017; Krull et al., 2014) bestehen. Auch in Bezug auf die Effekte inklusiver vs. nicht-inklusive Beschulung auf die soziale Entwicklung von Schüler:innen mit SPF liegen gemischte Befunde vor (Ruijs & Peetsma, 2009). Zusammenfassend scheint es bei Vorliegen von Unterschieden vor allem solche zum Nachteil von Schüler:innen mit SPF in sozialen und akademischen Variablen zu geben. Das kann zum Teil, aber nicht gänzlich, durch die Diagnosen an sich erklärt werden könnte (vgl. MSB NRW, 2016).

Die Rolle der Lehrkraft. Die Varianz im Ausgrenzungsrisiko von Schüler:innen mit SPF auf Klassenebene (Huber & Wilbert, 2012) deutet darauf hin, dass es Variablen auf dieser Ebene gibt, die das Risiko beeinflussen können. In Klassen mit Schüler:innen mit und ohne SPF waren die Unterschiede in der sozialen Partizipation zwischen diesen Schüler:innen kleiner, je positiver das wahrgenommene Klassenklima war und je stärker sich die Lehrkraft laut eigener Angaben bei der Leistungsbeurteilung an individuellen Bezugsnormen orientierte (Krawinkel et al., 2017). Auch fühlten sich Schüler:innen mit SPF besser sozial integriert, je positiver die Einstellungen der Lehrkräfte zur Inklusion in regulären Klassen im Vergleich zu

Förderschulen waren (Heyder et al., 2020) – ein praktisch sehr relevanter Befund, da Lehrkräfte vor allem in Bezug auf die Inklusion von Schüler:innen mit Lernbeeinträchtigungen und Verhaltensproblemen besonders negativ eingestellt sind (für einen Überblick s. de Boer et al., 2011). In beiden Studien gaben die Schüler:innen in Klassen mit diesen Merkmalen eine höhere soziale Partizipation bzw. Integration an, wohingegen die Angaben der Schüler:innen ohne SPF nicht variierten. Erkenntnisse zum Zusammenhang zwischen Lehrkräftemerkmalen und den Leistungen von Schüler:innen mit und ohne SPF liegen nach aktuellen Recherchen nicht vor. Dies gilt auch im Speziellen für das Wissen der Lehrkräfte. Dieses Wissen könnte aber für die inklusive Beschulung von Schüler:innen mit SPF an Regelschulen von besonderer Bedeutung sein (Ruijs & Peetsma, 2009).

3.1.2 Teilleistungsstörungen

Besondere Schwierigkeiten im Lesen und/oder Rechtschreiben und im Rechnen (KMK, 1991; KMK, 2007; Landesregierung Nordrhein-Westfalen, 2020) gelten als Teilleistungsstörungen. Beide werden im internationalen Klassifikationssystem ICD-10 aufgeführt und gehören zu den umschriebenen Entwicklungsstörungen schulischer Fertigkeiten (World Health Organization [WHO], 2016). Es lassen sich für erstere alternative Bezeichnungen wie Lese- und Rechtschreibstörung (LRS) (KMK, 1991; WHO, 2016), Legasthenie (Bundesverband Legasthenie und Dyskalkulie e. V., 2018) und Dyslexie (Novita, 2016) und für letztere Bezeichnungen wie Rechenstörung (KMK, 2007) und Dyskalkulie (Bastos et al., 2016; Landesregierung Nordrhein-Westfalen, 2020) finden. Für eine einheitliche Bezeichnung werden im Folgenden in Anlehnung an die internationale Literatur die Begriffe Dyslexie und Dyskalkulie genutzt. Bei Schüler:innen mit dieser Diagnose sind die Schwierigkeiten nicht durch einen Mangel an Gelegenheit zu lernen, eine Intelligenzminderung oder irgendeine erworbene Hirnschädigung oder -krankheit zu erklären (WHO, 2016). Eine umfassende Analyse der Lernsituation, allgemeine Fördermaßnahmen (im Rahmen des Regelunterrichts) und zusätzliche Fördermaßnahmen (schulische

Förderkurse) werden durch die Lehrkräfte durchgeführt. Zusätzliche Fördermaßnahmen kommen dabei für die Unterstützung von Schüler:innen der Grundschule und auch der weiterführenden Schule infrage, wobei unterschiedliche Bedingungen gelten. Sie kommen für Schüler:innen der Klassen 1 und 2 in Betracht, wenn die notwendigen Voraussetzungen für das Lesen- und Schreibenlernen noch fehlen. Bei Schüler:innen der Klassen 3 bis 6 müssen Leistungsausfälle mindestens drei Monate bestehen, und Schüler:innen der Klassen 7 bis 9 bzw. 10 können zusätzlich gefördert werden, wenn in Einzelfällen die besonderen Schwierigkeiten noch nicht behoben werden konnten (KMK, 1991). Ein Nachteilsausgleich (z. B. Ausweitung der Arbeitszeit bei Klassenarbeiten) steht in Deutschland nur Schüler:innen mit einer diagnostizierten Dyslexie, nicht aber mit einer diagnostizierten Dyskalkulie zu (KMK, 2007). Im Gegensatz zum SPF, unter den Teilleistungsstörungen in der Regel nicht fallen (KMK, 1991), gibt es in Deutschland keine landesweiten Daten über die Prävalenz von Schüler:innen mit Dyslexie und Dyskalkulie. Eine entsprechende Statistik ist auch nicht vorgesehen (Landesregierung Nordrhein-Westfalen, 2020). Internationale Studien legen Prävalenzraten von 3 bis 9% nahe (Barbiero et al., 2012; Bastos et al., 2016; Leonardi et al., 2021; Wong et al., 2017), wobei die stark schwankenden Zahlen auf unterschiedliche Kriterien, Tests, Altersstufen und Länder zurückgeführt werden könnten (Barbiero et al., 2012). Aus Studien im deutschsprachigen Raum gehen Prävalenzen einer Lernschwäche bzw. –störung im Lesen, Rechtschreiben oder Rechnen von 13 bis 25% bei Grundschulkindern hervor (vgl. Thomas et al., 2015).

Im Gegensatz zur Erforschung sozialer und akademischer Variablen bei Schüler:innen mit SPF liegen bislang weniger Erkenntnisse über diese bei Schüler:innen mit Teilleistungsstörungen vor, wobei zu vermuten ist, dass auch hier eine Benachteiligung von Schüler:innen mit Teilleistungsstörungen besteht (Ziesig Deng, 2016). So legen es zumindest längsschnittliche Analysen nahe, nach denen schlechtere Leistungen in Lesen und Mathematik von Grundschüler:innen eine geringere Akzeptanz durch Peers im Folgeschuljahr

vorhersagen konnten (Laursen et al., 2021). Auch liegen bereits Hinweise vor, dass Schüler:innen mit Dyslexie domänenspezifisch, aber nicht generell mehr Angst und ein geringeres Selbstwertgefühl haben als Schüler:innen ohne Dyslexie (Novita, 2016). Auch Schüler:innen mit Dyskalkulie haben ein erhöhtes Risiko, mathematikbezogene Ängste zu entwickeln, was sich wiederum negativ auf die Leistungen auswirken kann (Semeraro et al., 2020). Darüber hinaus berichten Eltern von Schüler:innen mit anhaltender Dyskalkulie mehr Aufmerksamkeitsprobleme bei ihren Kindern als Eltern von Schüler:innen mit nicht-anhaltender Dyskalkulie (Shalev et al., 1998).

Sowohl das Lesen und Schreiben als auch das Rechnen sind wichtige Grundlagen für nahezu alle Lebenslagen. Bei Erwachsenen mit Problemen in diesen Bereichen konnten deshalb bereits langfristige Nachteile wie ein niedrigeres generelles Selbstwertgefühl (Riddick et al., 1999), eine geringere Lebenszufriedenheit und soziale Unterstützung (Kalka & Lockiewicz, 2018), weniger ökonomische Ressourcen (Parsons & Bynner, 2005) und eine höhere Neigung, mit Darlehen in Verzug zu geraten (Gerardi et al., 2013), nachgewiesen werden. Um diesen Nachteilen vorzubeugen, sollen frühe Interventionen erfolgen, zu denen bereits zahlreiche zusammenfassende Artikel vorliegen (Al Otaiba et al., 2018; Austin et al., 2017; Chodura et al., 2015; Monei & Pedro, 2017).

Die Rolle der Lehrkraft. Ein Großteil der Verantwortung bei der Diagnose und Förderung der betreffenden Schüler:innen liegt bei den Lehrkräften selbst, in Einzelfällen unterstützt durch Schulpsycholog:innen und besonders erfahrene Lehrkräfte (vgl. KMK, 1991 für Bestimmungen in Nordrhein-Westfalen; für internationale Perspektiven s. z. B. Sousa et al., 2017). Vor allem Wissenschaftler:innen und Lehrkräfte mit entsprechender Ausbildung sehen das Wissen über die individuellen Bedürfnisse aller und der betroffenen Schüler:innen als wichtig für das erfolgreiche Unterrichten des entsprechenden Faches an (Sheppard & Wieman, 2020; Sousa et al., 2017). Studien decken allerdings immer wieder Mängel in verschiedenen Wissensbereichen von (angehenden) Lehrkräften auf (Ness & Southall, 2010;

Peries et al., 2021; Washburn et al., 2017).

3.1.3 Geschlecht

Obwohl es sich beim Geschlecht nur um eines von vielen Merkmalen handelt, mit denen Menschen beschrieben werden können, wird es als wichtigste und mächtigste soziale Kategorie betitelt (Eckes & Trautner, 2000). So finden sich in den unterschiedlichsten Kulturen Annahmen über männliche und weibliche Rollenbilder, Rechte und Verantwortlichkeiten. Dies gilt auch für das Bildungssystem: Der Wechsel vom geschlechtergetrennten zum gemeinsamen Unterricht von Jungen und Mädchen sollte die Gleichberechtigung der Geschlechter fördern, sodass dieses Modell in Deutschland mittlerweile zum Standard geworden ist, obwohl sich gezeigt hat, dass diese Änderung allein dafür nicht ausreicht (Hannover & Wolter, 2021). Es zeigen sich Unterschiede zwischen Jungen und Mädchen in den Schulleistungen, Bildungsabschlüssen und Berufschancen. Diese Unterschiede sind nicht allein auf die tatsächlichen Fähigkeiten zurückzuführen, sondern zum Beispiel auf Erwartungen und Bewertungen durch die Lehrkraft, das Selbstkonzept über die eigenen Fähigkeiten, die eigene Motivation und Geschlechtsstereotype, die für beide Geschlechter nachteilig und einschränkend sind (für einen Überblick s. Hannover & Wolter, 2021).

Empirische Studien (Keller et al., 2021; Steinmayr et al., 2019) und die internationale Schulleistungserhebung PISA (Mostafa & Schwabe, 2019) deuten auf schlechtere Leseleistungen bei den Jungen und schlechtere Mathematikleistungen bei den Mädchen mit unterschiedlich großen Effekten hin. Unabhängig davon betreiben Mädchen einen höheren Lernaufwand (Workman & Heyder, 2020) und haben tendenziell bessere Schulnoten (Steinmayr & Kessels, 2017; Voyer & Voyer, 2014). Damit erreichen sie höhere Schulabschlüsse (Bundesministerium für Bildung und Forschung [BMBF], 2022), haben allerdings weniger Erfolg im Beruf (Steinmayr & Kessels, 2017). Männer entscheiden sich eher für MINT-Berufe (in den Bereichen Mathematik, Informatik, Naturwissenschaften und

Technik), in denen Frauen unterrepräsentiert sind (Institut der deutschen Wirtschaft, 2021).

Mögliche Hemmnisse für schulische Leistungen, die Entwicklung eines guten Fähigkeitsselbstkonzepts und den Aufbau von Motivation können vielschichtig sein. An dieser Stelle seien deshalb ein paar Beispielbefunde herausgegriffen: Der geringere Lernaufwand, den Jungen im Vergleich zu Mädchen betreiben, kann zum Teil durch die höheren sozialen Kosten erklärt werden, die Jungen bei ihren Peers in Form von mehr Schikanen und weniger Beliebtheit erleben (Workman & Heyder, 2020). Dies könnte mit männlichen Rollenbildern einhergehen: In einer Längsschnittstudie mit 4200 Jugendlichen konnte der höhere gefühlte Druck der Geschlechterkonformität eine Zunahme des schulischen Fehlverhaltens bei männlichen, nicht aber bei weiblichen Jugendlichen vorhersagen (Heyder et al., 2021). Für die Bedeutung geschlechtsspezifischer Normen sprechen auch Befunde, nach denen die intrinsische Motivation zum Lesen besser durch das Ausmaß der Identifikation mit dem weiblichen bzw. männlichen Geschlecht (engl. gender) als durch das tatsächlich bei Geburt zugewiesene Geschlecht (engl. sex) erklärt werden kann (McGeown et al., 2012). Die Identifikation mit dem weiblichen Geschlecht hing in dieser Studie stärker mit der Motivation zusammen als die Identifikation mit dem männlichen Geschlecht. Die großen Geschlechterunterschiede im mathematischen Fähigkeitsselbstkonzept zuungunsten der Mädchen konnten in einer Grundschulstichprobe nicht durch die tatsächlichen Leistungen, wohl aber durch die Schulnoten und Eltern- und Lehrereinschätzungen teilweise erklärt werden (Steinmayr et al., 2019). Die höheren Erwartungen, die Lehrkräfte an die Mathematikleistungen von Jungen haben, hängen mit den Leistungen der Schüler:innen zusammen und benachteiligen dementsprechend die Mädchen (Muntoni et al., 2019). Außerdem scheinen Mädchen stärker durch die Kompetenzeinschätzungen der Lehrkräfte beeinflusst zu werden – hier in ihrem Leseselbstkonzept und dem empfundenen Wert des Lesens (Boerma et al., 2016). Insgesamt scheinen sich Geschlechtsunterschiede im Laufe der Schulzeit (hier von der dritten bis zur sechsten Klasse) eher zu vergrößern, was am stärksten

für Schüler:innen mit einem niedrigen sozio-ökonomischer Status gilt (Becker & McElvany, 2018).

Die Rolle der Lehrkraft. Nach den Standards für die Lehrer:innenbildung müssen Lehrkräfte die Bedeutung geschlechtsspezifischer Einflüsse auf Bildungs- und Erziehungsprozesse kennen (KMK, 2014), und auch die Delphi-Studie zu bildungswissenschaftlichen Inhalten in der Lehrer:innenbildung ergab das Geschlecht als wichtigen Teilbereich von Diversität (Kunina-Habenicht et al., 2012). Dennoch scheinen Stereotype und die zugrunde liegenden Fehlannahmen in unterschiedlichem Ausmaß unter (angehenden) Lehrkräften verbreitet zu sein. In einer Studie mit Lehramtsstudierenden lag bei knapp 50% der Befragten mindestens eine der drei erfragten Fehlannahmen vor, insgesamt wurden aber mehr als die Hälfte aller inkorrekten Aussagen auch als solche erkannt (Dersch et al., 2022). In einer Interviewbefragung berichteten Lehrkräfte unterschiedlich über die störenden und akademischen Verhaltensweisen von Jungen im Vergleich zu Mädchen. Analog zu gängigen Geschlechtsstereotypen wurden Jungen unabhängig vom Alter und der Schulklasse eher als ungestüm, wettbewerbsorientiert, aktiv und gut in Mathematik und Mädchen eher als geschwätzig, fleißig, schüchtern und gut im Lesen beschrieben. Viele der befragten Lehrkräfte gaben außerdem an, Unterrichtsentscheidungen basierend auf dem Geschlecht anstelle von dem individuellen Lernverhalten zu treffen (Morrissette et al., 2017). Dass die Erwartung bzw. Einschätzung von Lehrkräften eine wichtige Einflussgröße ist, konnte bereits nachgewiesen werden (z. B. Steinmayr et al., 2019), was bislang nicht für das Diversitätswissen im Kontext schulischer Inklusion gilt.

3.1.4 Kognitive Hochbegabung

Der Begriff Hochbegabung wird uneinheitlich genutzt, bezeichnet im Allgemeinen aber eine stark überdurchschnittliche Intelligenz (Rost, 2000), die über verschiedene Intelligenztests gemessen werden kann (z. B. Terman, 1925). Hochbegabung wird dementsprechend in Studien zum Teil leicht unterschiedlich operationalisiert, zum Beispiel

über den Intelligenzquotienten (IQ) mit einem Cut-Off von 125 bzw. 130 oder mit dem IQ als kontinuierliche Variable ohne direkte Unterscheidung zwischen Hochbegabten und Nicht-Hochbegabten (Francis et al., 2016). Um Kohorteneffekte bei Intelligenztestleistungen auszuschließen, sollten aktuelle Normstichproben genutzt werden (Rost, 2000), denn bei dem IQ handelt es sich um einen normierten Wert. Einen IQ von ≥ 130 oder mehr haben dementsprechend ca. 2.1% der Bevölkerung und einen Intelligenzquotienten von ≥ 145 haben nur ca. 0.1% der Bevölkerung, wobei keine offiziellen Statistiken zur Prävalenz von hochbegabten Schüler:innen in den verschiedenen Schulformen vorliegen. Unter diesen Schüler:innen befinden sich auch sogenannte Underachiever mit einem hohen IQ und erwartungswidrigen Minderleistungen, deren Anteil gemessen an allen Hochbegabten und abhängig von der jeweiligen Stichprobe auf 9% bis 50% geschätzt wird. Neben dem IQ können zur Diagnostik auch weitere Leistungsdaten (z. B. Schulleistung, Kreativität), leistungsrelevante Konstrukte (z. B. Motivation, Interessen) und individuelle Entwicklungsbedingungen (z. B. verfügbare Ressourcen, Konflikte) herangezogen werden (Preckel & Vock, 2021).

Sowohl in der Wissenschaft als auch in der Praxis liegen verschiedene Annahmen über Hochbegabte vor: Die Harmoniehypothese besagt, dass Hochbegabte eine durchschnittliche oder überdurchschnittliche psychosoziale Anpassung haben, und die Disharmoniehypothese geht von einer unterdurchschnittlichen psychosozialen Anpassung aus. Umfragen zeigen, dass Personen eher an die Disharmoniehypothese glauben, was für die Allgemeinbevölkerung (Baudson, 2016), aber auch für Lehrkräfte im Speziellen gilt (Baudson & Preckel, 2016). Empirische Studien stützen aber zu einem großen Teil die Harmoniehypothese. Demnach erbringen Hochbegabte zum Beispiel höhere Leistungen, sind intrinsisch motivierter (Bergold et al., 2020) und haben eine vergleichbare Schulzufriedenheit wie ihre normalbegabten Peers (Guignard et al., 2021). Zusammenfassende Analysen legen ebenfalls eine bessere sozio-emotionale Anpassung, weniger Verhaltensprobleme (Francis et

al., 2016), eine bessere mentale Gesundheit, ein besseres psychisches Wohlbefinden (Jones, 2013) und eine höhere emotionale Intelligenz (Alabbasi et al., 2021; Ogurlu, 2021) nahe. Unterschiede zuungunsten der hochbegabten Schüler:innen, hier in Bezug auf die Verbindung zu Peers, könnten hingegen darauf zurückzuführen sein, dass das Label für Lehrkräfte mit negativen Stereotypen assoziiert ist, die es unbedingt zu vermeiden bzw. zu korrigieren gilt (Guignard et al., 2021).

Die Rolle der Lehrkraft. Lehrkräfte geben an, dass sie in der Aus- und Weiterbildung wenig über Hochbegabung lernen, dabei ist das Wissen über Hochbegabung wichtig für die Identifikation und Förderung hochbegabter Schüler:innen (Heller et al., 2005). Außerdem gehen Fehlannahmen in diesem Bereich mit negativeren Einstellungen zur Förderung Hochbegabter einher (Heyder, Bergold & Steinmayr, 2018). Befeuert werden diese Fehlannahmen unter Umständen noch durch Massenmedien, die in der Lage sind, die Einstellungen gegenüber Hochbegabten negativ wie positiv zu beeinflussen: Negativ stereotypisierende Berichterstattung führt dabei eher zu negativeren Einstellungen und evidenzbasierte Berichterstattung eher zu positiveren Einstellungen (Bergold et al., 2021). Dies ist vor allem relevant, da Lehrkräfte angeben, ihr Wissen zu einem großen Teil aus den Medien zu haben und nur zu einem geringeren Teil aus universitären Veranstaltungen (Heyder, Bergold & Steinmayr, 2018). Ein Ausbau evidenzbasierten Wissens und ein Abbau negativer Stereotype erscheint deshalb wichtig für den Umgang mit hochbegabten Schüler:innen. Dies gilt auch für hochbegabte Underachiever, die von Lehrkräften häufig nicht sicher erkannt werden, da diese sich in der Identifikation vor allem an Schulnoten orientieren (Rost & Hanses, 1997) und hohe akademische Leistungen fälschlicherweise als notwendige Bedingung für Hochbegabung einschätzen (Heyder, Bergold & Steinmayr, 2018). Dementsprechend werden auch keine Fördermaßnahmen eingeleitet, die zum Beispiel bei den vorhandenen Wissenslücken der Schüler:innen, der Optimierung des eigenen Unterrichts, der Schließung von Motivationsdefiziten, einer differenzierten Aufgabenstellung, aber auch bei

professioneller pädagogischer und pädagogisch-psychologischer Hilfe ansetzen könnten (Sparfeldt & Rost, 2012).

3.2 Testaufbau und Items

Um das Wissen und die Fehlannahmen über die oben beschriebenen Inhaltsbereiche erfassbar zu machen, wurden für den hier vorgestellten Test eindeutig richtige bzw. falsche Aussagen erstellt. In vielen vorhergehenden Tests zu einzelnen Diversitätsbereichen wird ein Richtig-Falsch-Antwortformat genutzt (z. B. Haimour & Obaidat, 2013; Scitutto et al., 2000), wobei die Weiß-nicht-Option eingefügt wird, um Ratetendenzen zu minimieren (z. B. Heyder, Vaskova et al., 2018; Sanz-Cervera et al., 2017). Eine Aussage kann als korrekt bzw. inkorrekt gewertet werden, wenn ein breiter Forschungskonsens darüber besteht (Voss et al., 2011). Dieses Format eignet sich nicht nur zur Messung des Wissens, sondern auch zum Aufdecken von Fehlannahmen, also, was Personen fälschlicherweise glauben zu wissen (Scitutto et al., 2000). Dabei stellt das Format eine ökonomische und vergleichbar reliable und valide Alternative zum Multiple-Choice-Format dar (Tasdemir, 2010). Deswegen wurden für den vorliegenden Test eindeutig richtige bzw. falsche Aussagen, in einem ausgewogenen Verhältnis, zum deklarativen Faktenwissen von Wissenschaftler:innen der Pädagogischen Psychologie und Sonderpädagogik und Lehrkräften an inklusiven Schulen (Lehramt für Gymnasien oder Sonderpädagogik) formuliert (vgl. Heyder, Vaskova et al., 2018). Die Pilotierung erfolgte an Stichproben aus mehreren hundert Studierenden und etwas weniger Lehrkräften. Drei Expertinnen aus Lehrkräftebildung und Schulpraxis bestimmten unabhängig voneinander die inhaltlich wichtigsten Items, von denen diejenigen mit mindestens zwei Nennungen und einer Trennschärfe $> .20$ ausgewählt und bei Bedarf geringfügig sprachlich optimiert wurden. Daraus entstanden zwei Testversionen, die in der vorliegenden Arbeit evaluiert werden: Der DiWi (= Diversitätswissen, Langversion) umfasst 142 Items aus den Bereichen Förderschwerpunkt ESE (35 Items) und Lernen (28 Items), Teilleistungsstörungen (27 Items), Geschlecht (26 Items) und kognitive Hochbegabung (26

Items). Der DiWi-K (= Diversitätswissen, Kurzversion) umfasst 36 Items aus den Bereichen Förderschwerpunkt ESE (7 Items) und Lernen (7 Items), Teilleistungsstörungen (8 Items), Geschlecht (8 Items) und kognitive Hochbegabung (6 Items). Alle Items der Kurzversion sind ebenfalls in der Langversion enthalten. Dazu gehören die folgenden Beispielitems (inklusive korrekter Antwort):

Förderschwerpunkt ESE

- Angst ist behandlungsbedürftig, wenn sie die Entwicklung und den Alltag des Kindes oder der/des Jugendlichen beeinträchtigt. (*richtig*)
- Wertschätzende Äußerungen bei erwünschtem Verhalten führen zu einer Überforderung der Schüler:innen mit sonderpädagogischem Unterstützungsbedarf im Bereich ESE. (*falsch*)

Förderschwerpunkt Lernen

- Mit sinkender sozialer Schichtzugehörigkeit steigt die Wahrscheinlichkeit einer Lernbeeinträchtigung. (*richtig*)
- Schüler:innen mit sonderpädagogischem Unterstützungsbedarf im Bereich Lernen sind gemäß schulrechtlicher Definition unterdurchschnittlich intelligent. (*falsch*)

Teilleistungsstörungen

- Bei der Diagnostik einer Teilleistungsstörung muss ausgeschlossen werden, dass eine unzureichende Beschulung/Förderung stattgefunden hat. (*richtig*)
- Lese- und Rechtschreibschwierigkeiten treten stets gemeinsam auf. (*falsch*)

Geschlecht

- Gendergerechte Sprache bei Berufsbeschreibungen führt dazu, dass Mädchen und Jungen sich stärker auch für geschlechtsuntypische Berufe interessieren. (*richtig*)
- An Förderschulen sind Mädchen und Jungen ähnlich häufig vertreten. (*falsch*)

Kognitive Hochbegabung

- Kognitive Hochbegabung im engeren Sinne wird durch eine weit überdurchschnittliche Intelligenz definiert. (*richtig*)
- Wenn jemand viele Einsen und Zweien bekommt, ist er oder sie in der Regel kognitiv hochbegabt. (*falsch*)

3.3 Kriterien zur Evaluation dieses Tests

Als erste Annäherung an die Evaluation des Tests zum Diversitätswissen von (angehenden) Lehrkräften mit den oben beschriebenen Subskalen können zwei Veröffentlichungen herangezogen werden, in denen Items der hier vorgestellten Testversionen vorkommen (Heyder, Bergold & Steinmayr, 2018; Heyder, Vaskova et al, 2018). In einer ersten Pilotierung wurden die Items Studierenden und Lehrkräften vorgelegt, und auf dieser Basis wurden die hier vorgestellten Testversionen erstellt. Dazu wurden Items ausgewählt und zum Teil vorher sprachlich optimiert, die eine hohe interne Konsistenz und eine breite Spanne an Itemschwierigkeiten ermöglichen. Erste Ergebnisse zeigen ein höheres Wissen über die beiden Förderschwerpunkte ESE und Lernen bei Studierenden für sonderpädagogische Förderung im Vergleich zu Studierenden anderer Lehramtsstudiengänge (Heyder, Vaskova et al, 2018). Eine erste Evaluation der Subskala zur kognitiven Hochbegabung in einer Stichprobe aus Lehrkräften an Gesamtschulen und Gymnasien erfolgte mit einer 38-Item-Version, in der mehr als die hier vorgestellten Items vorkamen. Die Analysen ergaben eine zufriedenstellende interne Konsistenz von $\alpha = .79$. Das Wissen der Lehrkräfte korrelierte nicht mit ihrem selbsteingeschätzten Wissen, ihren Einstellungen zur Förderung Hochbegabter und dem (beruflichen und privaten) Kontakt zu Hochbegabten. Allerdings hatten Lehrkräfte mit mehr Fehlannahmen negativere Einstellungen zur Förderung Hochbegabter (Heyder, Bergold & Steinmayr, 2018).

Die beiden Studien geben erste Hinweise auf die Eignung des Tests für den Einsatz

bei (angehenden) Lehrkräften. In den nächsten Schritten soll der Test weiter evaluiert werden. Als Basis wird die Klassische Testtheorie herangezogen. Wissenschaftliche Tests sollen Anforderungen zur Qualitätssicherung – den Gütekriterien – genügen, zu denen vor allem die Hauptgütekriterien Objektivität, Reliabilität und Validität, aber auch die Nebengütekriterien Ökonomie, Nützlichkeit, Zumutbarkeit, Fairness und Unverfälschbarkeit gehören (Moosbrugger & Kelava, 2020).

3.3.1 Klassische Testtheorie

Die wohl am häufigsten verbreiteten Theorien zur Überprüfung von Fragebögen und Tests sind die Klassische Testtheorie (KTT) und die Item-Response-Theorie (IRT). Sie weisen große Überschneidungen auf und stellen nicht zwangsläufig zwei konkurrierende Alternativen zueinander dar, sondern können als Ergänzung betrachtet werden (Moosbrugger et al., 2020). In der hier vorgestellten Arbeit wird die KTT als Grundlage für die Testevaluation herangezogen. Aber auch die IRT kann einen wichtigen Beitrag zur Evaluation des Tests leisten, weshalb sie abschließend diskutiert wird (vgl. Kapitel 6.3 zu Stärken und Grenzen der Arbeit und 6.4.1 zu Implikationen für die Forschung).

In der KTT wird, analog zur IRT, im Fall eines Leistungstests angenommen, dass die (messbare) Beantwortung der Testitems von latenten (nicht messbaren) Fähigkeitsmerkmalen abhängt. Die Itemantworten sind deshalb als Indikatoren für die individuellen Merkmalsausprägungen (hier Wissen und Fehlannahmen) zu sehen. Auf dieser Basis kann eine Beurteilung der Reliabilität des Tests und der Validität der Testwertinterpretationen ökonomisch und praktikabel vorgenommen werden. In der Grundgleichung der KTT setzt sich jede beobachtete Itemvariable (= Testwert pro Item) aus einem wahren Wert (True Score) und einem Messfehler zusammen. Wird ein aufgestelltes Modell überprüft, besteht die grundsätzliche Annahme, dass alle Itemvariablen dieselbe latente Variable messen (im Fall von mehreren Subskalen messen alle Itemvariablen einer Subskala dieselbe latente Variable) und dass die Messfehler unabhängig sind und nicht miteinander korrelieren

(Eindimensionalität, Moosbrugger et al., 2020). Diese Struktur lässt sich mittels konfirmatorischer Faktorenanalysen prüfen. Für den vorliegenden Test bedeutet dies, dass geprüft werden muss, ob sich die verschiedenen Inhalte auch tatsächlich voneinander abgrenzen lassen. Wie auch bei anderen Wissenstests (Horn et al., 2004; Kunina-Habenicht et al., 2020) wird angenommen, dass die Testinhalte verschiedenen Faktoren zugeordnet werden können, denen ein gemeinsamer Generalfaktor zugrunde liegt. Dazu wird berechnet, wie weit das aufgestellte Modell mit fünf Faktoren (SPF ESE, SPF Lernen, Teilleistungsstörungen, Geschlecht, kognitive Hochbegabung) und einem Generalfaktor (Diversitätswissen) durch die empirischen Daten abgebildet werden kann.

3.3.2 Hauptgütekriterien

Objektivität. Ein Test ist dann objektiv, wenn er unabhängig von Ort, Zeit, Testleiter:in und Auswerter:in durchgeführt werden kann. Er soll sowohl in der Durchführung, als auch in der Auswertung und Interpretation objektiv sein (Moosbrugger & Kelava, 2020). Der hier vorgestellte Wissenstest ist so konstruiert, dass er diesen Anforderungen in hohem Maße genügen kann. Er liegt in schriftlicher Form vor, sodass eine direkte Interaktion mit einem/einer Testleiter:in nicht zwingend nötig ist. Da es sich nicht um einen Speedtest handelt, bei dem die Testdauer für die Vergleichbarkeit der Ergebnisse von Bedeutung wäre, wird keine Testzeit vorgegeben. Die Instruktionen sind für alle Teilnehmenden gleich und liegen schriftlich vor: *Bitte kreuzen Sie für jede Aussage an, ob sie Ihrer Meinung nach richtig oder falsch ist. Wenn Sie sich bei einer Aussage unsicher sind, antworten Sie bitte mit "weiß nicht". Raten Sie bitte nicht und lassen Sie bitte keine Aussage aus.* Das Richtig-Falsch-Weiß-nicht-Antwortformat ermöglicht eine objektive Auswertung, die nicht von den auswertenden Personen abhängt. Der Test ist so skaliert, dass ein Mittelwert oder Summenwert auf Intervallskalenniveau aus den korrekten bzw. inkorrekten Antworten pro Subskala bzw. Gesamttest, unter der Voraussetzung der Unkorreliertheit der Messfehler, gebildet werden kann, um das Wissen bzw. die Fehlannahmen abzubilden. Größere Werte

stehen dementsprechend für ein höheres Wissen bzw. mehr Fehlannahmen. Dieses Vorgehen ist laut KTT gerechtfertigt, wenn die Testwerte einen guten empirischen Prädiktor für andere Konstrukte darstellen (vgl. Abschnitt Validität). Eine objektive Interpretation der Testergebnisse im Sinne von Einteilungskriterien (z. B. in niedriges und hohes Wissen) ist aktuell noch nicht möglich, da keine Normstichproben vorliegen. Die Ergebnisse können deshalb nur vorsichtig im Licht vorheriger Studien interpretiert werden (z. B. Heyder, Bergold & Steinmayr, 2018; Scituito et al., 2000; Sousa et al., 2017; Yin et al., 2019). Eine Interpretation der Testwerte im Sinne von festgelegten Kriterien ist für die Nutzung ohne Einzelfalldiagnostik aber nicht zwingend nötig und sollte deshalb zum aktuellen Zeitpunkt vermieden werden.

Reliabilität. Unter der Voraussetzung, dass die Objektivität eines Tests gegeben ist, beschreibt die Reliabilität die Exaktheit oder auch Fehlerfreiheit einer Messung: Je fehlerfreier ein Test ist, desto messgenauer ist er (Moosbrugger & Kelava, 2020). Aus den vier gängigen Reliabilitätsvarianten (Retest-Reliabilität, Paralleltest-Reliabilität, Split-Half-Reliabilität, Cronbachs Alpha), die jeweils das Verhältnis der True-Score-Varianz an der Gesamtvarianz schätzen, werden im Rahmen dieser Arbeit Cronbachs Alpha und die Retest-Reliabilität genutzt.

Cronbachs Alpha ist das gängigste Reliabilitätsmaß, das in vielen Studien angegeben wird. Es handelt sich dabei um eine Form der Split-Half-Reliabilität, bei der jedes Item als eigenständiger Testteil betrachtet wird. Je höher die Korrelation der Testteile untereinander, desto höher ist Cronbachs Alpha. Als Voraussetzungen für die Berechnung müssen die Annahmen der Eindimensionalität und essentiellen τ -Äquivalenz gegeben sein, die mittels konfirmatorischer Faktorenanalyse überprüft werden können. Dazu wird im ersten Schritt ein Modell definiert, in dem alle Items auf einem Faktor laden und die Fehlervarianzen unabhängig voneinander sind (Eindimensionalität). Im zweiten Schritt werden die Faktorladungen gleichgesetzt (essentielle τ -Äquivalenz). Ob das restriktive Modell einen

guten Modellfit aufweist, kann anschließend mittels χ^2 -Differenztest überprüft werden. Anders als häufig berichtet gibt Cronbachs Alpha somit nicht automatisch das Ausmaß der internen Konsistenz bzw. Eindimensionalität an, da ein mehrdimensionales Konstrukt (z. B. aufgrund von Methodeneffekten) ebenfalls einen hohen Wert erreichen kann (Gäde et al., 2020). Als Alternative zu den strengen Annahmen von Cronbachs Alpha können modellbasierte Methoden der Reliabilitätsschätzung genutzt werden. Für mehrdimensionale Tests (wie den hier vorliegenden) kann McDonalds Omega (ω) eine Schätzung liefern für (1) den Anteil der wahren Varianz an der Gesamtvarianz der Testwerte, (2) den Anteil der wahren Varianz an der Gesamtvarianz der Testwertvariablen, der nur auf den Generalfaktor zurückgeführt wird, und (3) den Anteil der wahren Varianz an der Gesamtvarianz der Testwertvariablen, der nur auf die subskalenspezifischen Faktoren zurückgeführt wird (Schermelleh-Engel & Gäde, 2020).

Die Retest-Reliabilität ist eine Form der Test-Test-Reliabilität. Dazu wird der Test mit denselben Teilnehmenden zu zwei Messzeitpunkten durchgeführt und die Korrelation zwischen den Testwerten ermittelt. Die Bestimmung der Retest-Reliabilität ist nicht ganz unproblematisch. Zum einen können die sehr restriktiven Voraussetzungen (Eindimensionalität, essentiellen τ -Parallelität der Messungen) mittels konfirmatorischer Faktorenanalysen zur Überprüfung der strikten Messinvarianz nicht in jedem Fall geprüft werden. Zum anderen haben unsystematische Veränderungen und Erinnerungseffekte einen Einfluss auf das Ergebnis (Gäde et al., 2020). Im Fall des Diversitätswissens von angehenden Lehrkräften bedeutet dies, dass eine Vorlesungssitzung zwischen den beiden Messzeitpunkten die Testwerte zum zweiten Messzeitpunkt systematisch beeinflussen kann, was sich nicht auf die Korrelation auswirkt, da sich die Werte aller Teilnehmenden verändern. Bei unsystematischen Veränderungen hingegen kann die Retest-Reliabilität unterschätzt werden (z. B. wenn nur einige Studierende die betreffende Vorlesungssitzung besuchen, aber nicht alle). Dementsprechend darf das gewählte Intervall nicht zu groß sein, um nicht zu viele

Gelegenheiten für unsystematische Veränderungen zuzulassen. Erinnerungseffekte können im Gegensatz dazu dafür sorgen, dass die Retest-Reliabilität überschätzt wird. Ist das Intervall zu klein gewählt, erinnern sich die Studierenden an ihre Antworten zum ersten Messzeitpunkt und kreuzen allein deshalb die gleichen Antworten an.

Validität. Unter der Voraussetzung, dass die Objektivität und Reliabilität eines Tests gegeben sind, beschreibt die Validität „das Ausmaß, in dem empirische Befunde und theoretische Argumente die Interpretationen von Testwerten für die beabsichtigten Verwendungen von Tests unterstützen“ (Hartig et al., 2020, S. 530). Damit ist dieses Gütekriterium nicht mehr als generelle Eigenschaft eines Tests zu verstehen, sondern vielmehr als das Ausmaß der Belastbarkeit von Testwertinterpretationen. In diesem Sinne können auf Basis eines Modells, in dem theoretische Zusammenhänge definiert werden, empirische Befunde für diese Zusammenhänge und somit Hinweise auf die Validität der Testwertinterpretationen gesammelt werden. Dieses Vorgehen ist hypothetisch-deduktiv und arbeitet von der Theorie zur Empirie. Fundamental ist dabei die Güte des zugrunde liegenden Modells. Aus diesem Grund wurde für die Sammlung von Validitätshinweisen für das mit dem DiWi(-K) gemessene Diversitätswissen das weiter oben beschriebene COACTIV-Kompetenzmodell (Baumert & Kunter, 2013a; Kunter, Kleickmann et al., 2013) herangezogen. In dieser Arbeit geht es explizit nicht darum, das Modell zu validieren, sondern das Modell wird als theoretische Grundlage herangezogen, anhand derer ein nomologisches Netz rund um das Diversitätswissen gezeichnet wird. Das nomologische Netz umfasst alle theoretischen und empirischen Zusammenhänge, und seine Korrektheit kann schrittweise überprüft werden, indem getestet wird, ob die gemessenen Testwerte zuverlässige Indikatoren für weitere latente Konstrukte darstellen. Theoretisch angenommene Zusammenhänge sollten sich folglich unter Hinzunahme des zu prüfenden Konstrukts bzw. seiner gemessenen Testwerte empirisch bestätigen lassen. Aus diesem Grund wurden für die vorliegende Arbeit Konstrukte erfasst, die sich als formale Lerngelegenheiten, Aspekte professioneller

Kompetenz, Performanz und Unterrichtserfolg in das COACTIV-Modell (Baumert & Kunter, 2013a; Kunter, Kleickmann et al., 2013; Voss et al., 2015) einordnen lassen (Abb. 2). In diesem Fall sollen die Testwertinterpretationen für die Gruppendiagnostik in empirischen Studien (z. B. Grundlagenforschung, Evaluation von Aus- und Weiterbildungsmaßnahmen von Lehramtsstudierenden und Lehrkräften) validiert werden. Es findet hingegen keine Validierung von Testwertinterpretationen in der Einzelfalldiagnostik (z. B. zur Ableitung von kriteriumsgeleiteten Einzelfallentscheidungen) statt, weshalb von dieser Art der Interpretation abzusehen ist (vgl. Objektivität).

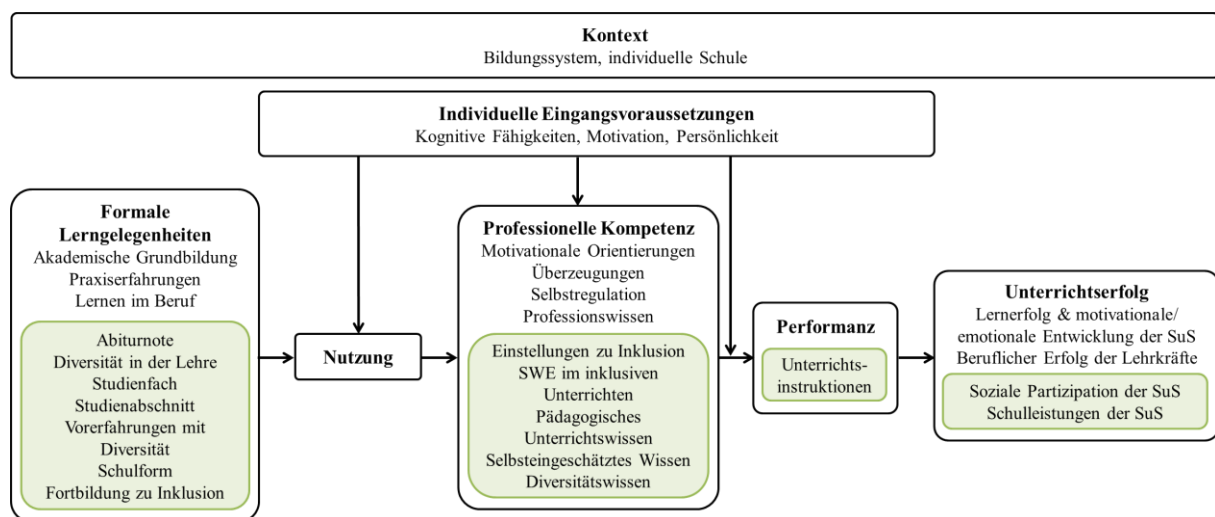


Abbildung 2. Eigene Darstellung des COACTIV-Kompetenzmodells (in Anlehnung an Kunter, Kleickmann et al., 2013, S. 67; Voss et al., 2015, S. 190) mit Einordnung aller studienrelevanten Variablen

3.3.3 Nebengütekriterien

Ökonomie. Der Test ist ökonomisch, beansprucht also möglichst wenig finanzielle und zeitliche Ressourcen (Moosbrugger & Kelava, 2020). Er liegt in einer Lang- und Kurzform vor, und die jeweilige Version kann passend für die angestrebten Ziele eingesetzt werden. Die Items liegen schriftlich vor, sodass neben einer Paper-and-Pencil- auch eine Online-

Befragung möglich ist. Für die Auswertung der Antworten gibt es eine Auswertungssyntax (s. Steinmayr et al., 2022).

Nützlichkeit. Ob der Test nützlich ist, also ob das gemessene Merkmal praktische Relevanz aufweist (Moosbrugger & Kelava, 2020), kann beurteilt werden, wenn Zusammenhänge zu weiteren unterrichtsrelevanten Variablen aufgedeckt werden (vgl. Validität). Aus dem COACTIV-Modell (Baumert & Kunter, 2013a; Kunter, Kleickmann et al., 2013; Voss et al., 2015) und aus empirischen Studien zum pädagogisch-psychologischen Wissen von (angehenden) Lehrkräften (König et al., 2017; Lauermann & König, 2016; Voss et al., 2011) lässt sich eine hohe theoretische Relevanz des Diversitätswissen für Lehrkräfte, Schüler:innen und das Unterrichten diverser Lerngruppen ableiten. Diese wird im Zuge der Evaluation empirisch geprüft.

Zumutbarkeit. Der Test ist zumutbar, da er keine zeitlich, psychisch und körperlich hohe Belastung von den Teilnehmenden fordert (Moosbrugger & Kelava, 2020). Die Bearbeitung im Rahmen der empirischen Forschung ist stets freiwillig und kann abgelehnt werden. Sollte er in einem Setting eingesetzt werden, in dem eine Teilnahme Pflicht ist, sollte eine sorgfältige Abwägung der Kosten und Nutzen stattfinden.

Fairness. Der Test sollte fair sein, da die Testwerte nicht auf Basis einer systematischen Benachteiligung bestimmter Personen aufgrund von ethnischen, soziokulturellen oder geschlechtsspezifischen Merkmalen zustande kommen (Moosbrugger & Kelava, 2020). Er richtet sich an (angehende) Lehrkräfte und soll nicht in lehramtsfernen Stichproben eingesetzt werden. Er kann darüber hinaus in verschiedene Sprachen übersetzt werden und kann sowohl als Paper-and-Pencil- als auch Online-Version genutzt werden. Eine Audio-Version bzw. das Vorlesen durch einen/eine Testleiter:in ist in Einzelfällen ebenfalls denkbar. Soll eine Interpretation der Testwerte im Vergleich zu anderen Personen stattfinden, müssen Normen gebildet werden, die verschiedene Merkmalsausprägungen berücksichtigen, sodass die Testwerte anhand vergleichbarer Stichproben ohne

Benachteiligungen eingeordnet werden können.

Unverfälschbarkeit. Da es sich um einen Leistungstest handelt, ist das Testergebnis nach oben unverfälschbar, die Testwerte können also nicht durch gezielte Vortäuschung eines unzutreffenden Testverhaltens verbessert werden (Moosbrugger & Kelava, 2020). Mit einer gezielten Verfälschung nach unten ist auf Basis von Phänomenen wie der sozialen Erwünschtheit nicht zu rechnen. Kreuzt eine Person aufgrund von fehlender Motivation immer die gleiche Antwort an oder ist ein Muster im Antwortverhalten erkennbar, sollte der entsprechende Fragebogen aus den Analysen ausgeschlossen werden.

4 Ziele, Fragestellungen und Studienüberblick

In dieser Arbeit wird ein Test zum Wissen über verschiedene Diversitätsbereiche von (angehenden) Lehrkräften evaluiert. Das damit erfasste Diversitätswissen wird darauf aufbauend in Zusammenhang mit verschiedenen unterrichtsrelevanten Variablen im Kontext schulischer Inklusion erforscht. Bisher bestehende Tests bieten ein breites Spektrum an Messinstrumenten mit verschiedenen Vor- und Nachteilen, das durch den vorliegenden Test erweitert werden soll. In Anlehnung an das COACTIV-Modell, ein Rahmenmodell zur professionellen Kompetenz von Lehrkräften (Kunter, Baumert et al., 2013), existieren aktuell verschiedene Tests zum pädagogisch-psychologischen Wissen, die einen Bezug zu Inklusion in Form von einzelnen Items oder Subskalen herstellen (z. B. König & Blömeke, 2009; Wildbrett et al., 2022). Außerdem liegen Tests vor, die einzelne Diversitätsbereiche abbilden (z. B. Ginsburg et al., 2021; Sousa et al., 2017). Der hier vorgestellte Test bildet mit fünf Subskalen verschiedene Diversitätsbereiche ab, ermöglicht damit eine sowohl spezifischere als auch umfassendere Messung und wird der Heterogenität des Konstrukts Diversität (vgl. Voss et al., 2011) damit eher gerecht als vorherige Messinstrumente. Darüber hinaus sind in bestehenden Tests unterschiedliche Item- und Antwortformate vertreten, die sich in der Ökonomie stark unterscheiden (z. B. König et al., 2017; Voss et al., 2011). In Anlehnung an bereits evaluierte und vielfach genutzte Tests (z. B. Scitutto et al., 2000; Sousa et al., 2017) kommen in dem hier vorgestellten Test eindeutig richtige bzw. falsche Aussagen zum Einsatz, die mit einem ökonomischen Richtig-Falsch-Weiß-nicht-Antwortformat eingeschätzt werden, das eine vergleichbar reliable und valide Alternative zum Multiple-Choice-Format darstellt (Tasdemir, 2010). Dieses Format soll die Ratewahrscheinlichkeit verringern und ermöglicht neben der Auswertung des Wissens auch eine Auswertung der Fehlannahmen von (angehenden) Lehrkräften (vgl. Scitutto et al., 2000).

Zur Evaluation des Wissenstests werden Lehramtsstudierende, Lehrkräfte und

Schüler:innen befragt bzw. getestet. Damit werden zusätzlich zu den Erkenntnissen aus Studierendenstichproben (Depaepe & König, 2018; König et al., 2017; Sanz-Cervera et al., 2017) auch solche aus praxisnahen Lehrkräftestichproben gewonnen. Die Fragestellungen zur Evaluation des vorgestellten Wissenstests lassen sich wie folgt zusammenfassen:

- 1) Welches Modell (1-Faktor-Modell vs. hierarchisches 5-Faktoren-Modell mit Generalfaktor) kann die Struktur des Wissens über die fünf Diversitätsbereiche am besten abbilden?
- 2) Wie hängen formale Lerngelegenheiten, Aspekte professioneller Kompetenz (berufliche Überzeugungen, Facetten des Professionswissens) und die Unterrichtsperformanz mit dem Diversitätswissen zusammen?
- 3) Wie hängt das Diversitätswissen mit dem Unterrichtserfolg auf Ebene der Schüler:innen zusammen?

Studie 1 (*Die Erfassung von Wissen über verschiedene Diversitätsbereiche im Kontext einer inklusionsorientierten Lehrkräftebildung*) besteht aus drei Teilstudien mit dem Fokus der ersten Evaluation des Tests (Fragestellung 1 & 2). In Studie 1.1 wird die Faktorstruktur des DiWi in einer Stichprobe aus 395 Lehramtsstudierenden im Bachelor überprüft, und es werden Hinweise auf die Validität der Testwertinterpretationen gesammelt. Der DiWi wird außerdem mit dem DiWi-K korreliert, der in Studie 1.2 und 1.3 weiter evaluiert wird. In diesen Teilstudien werden anhand von zwei Stichproben aus 462 Bachelor- und Master-Studierenden und 113 Bachelor-Studierenden weitere Validitätshinweise gesammelt, und die Retest-Reliabilität wird über einen Zeitraum von drei Monaten berechnet. Wie auch bei bisherigen Tests (vgl. Kunina-Habenicht et al., 2020) wird äquivalent zu den fünf Inhaltsbereichen eine fünf-faktorielle Struktur mit übergeordnetem Faktor angenommen. Auf dieser Basis und in Anlehnung an das COACTIV-Modell (Kunter, Kleickmann et al., 2013) werden zur Sammlung von Hinweisen auf ein nomologisches Netz rund um das Diversitätswissen (Hartig et al., 2020) die Variablen Abiturnote, selbsteingeschätztes Wissen,

Anzahl der in Lehrveranstaltungen behandelten Diversitätsbereiche, Studienfach (sonderpädagogische Förderung vs. andere Lehramtsstudiengänge), Studienabschnitt (Bachelor vs. Master), Vorerfahrungen mit Diversität, Einstellungen zu Inklusion, Selbstwirksamkeitserwartungen in der Gestaltung inklusiven Unterrichts und pädagogisches Unterrichtswissen herangezogen. Die Ergebnisse aus Studie 1 werden zusammengefasst und inklusive weiterer Informationen zur Testauswertung und den Itemkennwerten beider Testversionen öffentlich zugänglich gemacht (Steinmayr et al., 2022).

In Studie 2 (*Teachers' knowledge about diversity: Correlations with attitudes, self-efficacy, and instructions*) wird der DiWi in einer Stichprobe aus 139 Lehrkräften im Beruf evaluiert, weil die Ergebnisse nicht automatisch auf diese Stichprobe generalisiert werden können (Fragestellung 2). Hier werden die Variablen Ausbildungsstand (Bachelor-Studium, Master-Studium, Praxistätigkeit), beruflicher Kontakt zu Schüler:innen verschiedener Diversitätshintergründe, Einstellungen zu Inklusion, Selbstwirksamkeitserwartungen für das Unterrichten an einer inklusiven Schule, Unterrichtsinstruktionen, Studienfach (sonderpädagogische Förderung vs. andere Lehramtsstudiengänge), Schulform (Grundschule, Förderschule, andere) und Fortbildung zum Thema Inklusion zur Sammlung von Validitätshinweisen herangezogen. Außerdem findet eine Auswertung der Fehlannahmen statt.

In Studie 3 (*Links between teachers' knowledge about special educational needs and students' social participation and academic achievement in mainstream classes*) erfolgt eine detailliertere Betrachtung der beiden Subskalen zu den Förderschwerpunkten ESE und Lernen (Fragestellung 3). Auf Ebene der Lehrkräfte werden neben dem Wissen über die Förderschwerpunkte auch das Alter, die Jahre an der aktuellen Schule, die Einstellungen zu Inklusion und die Selbstwirksamkeitserwartungen im Unterrichten von Schüler:innen mit SPF und auf Ebene der Schüler:innen die soziale Partizipation und die schulischen Leistungen (im Lesen und in Mathematik) erfasst. Die zwei gebildeten Substichproben bestehen aus den 23

(bzw. 39) Klassenlehrer:innen und ihren 421 (bzw. 715) Schüler:innen in 25 (bzw. 40) Klassen mit Schüler:innen mit und ohne SPF ESE (bzw. SPF Lernen). In Multilevel-Analysen wird überprüft, wie das Wissen der Lehrkräfte mit der sozialen Partizipation und den Leistungen der Schüler:innen mit und ohne SPF ESE bzw. Lernen zusammenhängen. Damit leistet diese Arbeit einen Beitrag dazu, bei der Erforschung inklusiver Prozesse Faktoren zu identifizieren, die Inklusion vorantreiben und sowohl soziale als auch akademische Variablen in den Blick zu nehmen (Göransson & Nilholm, 2014).

Literaturverzeichnis I

- Al Otaiba, S., Rouse, A. G., & Baker, K. (2018). Elementary grade intervention approaches to treat specific learning disabilities, including dyslexia. *Language, Speech, and Hearing Services in Schools, 49*(4), 829–842. https://doi.org/10.1044/2018_lshss-dyslc-18-0022
- Alabbasi, A. M. A., Ayoub, A. E., & Ziegler, A. (2021). Are gifted students more emotionally intelligent than their non-gifted peers? A meta-analysis. *High Ability Studies, 32*(2), 189–217. <https://doi.org/10.1080/13598139.2020.1770704>
- Austin, C. R., Vaughn, S., & McClelland, A. M. (2017). Intensive reading interventions for inadequate responders in grades K–3: A synthesis. *Learning Disability Quarterly, 40*(4), 191–210. <https://doi.org/10.1177/0731948717714446>
- Banks, J., McCoy, S., & Frawley, D. (2017). One of the gang? Peer relations among students with special educational needs in Irish mainstream primary schools. *European Journal of Special Needs Education, 33*(3), 396–411. <https://doi.org/10.1080/08856257.2017.1327397>
- Barbiero, C., Lonciari, I., Montico, M., Monasta, L., Penge, R., Vio, C., Tressoldi, P. E., Ferluga, V., Bigoni, A., Tullio, A., Carrozzi, M., & Ronfani, L. (2012). The submerged dyslexia iceberg: How many school children are not diagnosed? Results from an Italian study. *PLoS ONE, 7*(10), e48082. <https://doi.org/10.1371/journal.pone.0048082>
- Bastos, J. A., Cecato, A. M. T., Martins, M. R. I., Grecca, K. R. R., & Pierini, R. (2016). The prevalence of developmental dyscalculia in Brazilian public school system. *Arquivos de Neuro-psiquiatria, 74*, 201–206. <https://doi.org/10.1590/0004-282x20150212>
- Baudson, T. G. (2016). The mad genius stereotype: Still alive and well. *Frontiers in Psychology, 7*, 368. <https://doi.org/10.3389/fpsyg.2016.00368>
- Baudson, T. G., & Preckel, F. (2016). Teachers' conceptions of gifted and average-ability students on achievement-relevant dimensions. *Gifted Child Quarterly, 60*(3), 212–225. <https://doi.org/10.1177/0016986216647115>
- Baumert, J., & Kunter, M. (2006). Stichwort: Professionelle Kompetenz von Lehrkräften. *Zeitschrift für Erziehungswissenschaft, 9*(4), 469–520. <https://doi.org/10.1007/s11618-006-0165-2>
- Baumert, J., & Kunter, M. (2013a). The COACTIV model of teachers' professional competence. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss & M. Neubrand (Eds.), *Cognitive activation in the mathematics classroom and professional competence of teachers: Results from the COACTIV project* (pp. 29–53). Springer. <https://doi.org/10.1007/978-1-4614-5149-5>
- Baumert, J., & Kunter, M. (2013b). The effect of content knowledge and pedagogical content knowledge on instructional quality and student achievement. In M. Kunter, J. Baumert,

- W. Blum, U. Klusmann, S. Krauss & M. Neubrand (Eds.), *Cognitive activation in the mathematics classroom and professional competence of teachers: Results from the COACTIV project* (pp. 175–205). Springer. https://doi.org/10.1007/978-1-4614-5149-5_9
- Beauftragter der Bundesregierung für die Belange von Menschen mit Behinderungen. (2018). *UN-Behindertenrechtskonvention*. Bundesministerium für Arbeit und Soziales. https://www.behindertenbeauftragter.de/SharedDocs/Downloads/DE/AS/PublikationenErklaerungen/Broschuere_UNKonvention_KK.pdf?__blob=publicationFile&v=8
- Beck, E., Baer, M., Guldemann, T., Bischoff, S., Brühwiler, C., Müller, P., Niedermann, R., Rogalla, M., & Vogt, F. (Eds.) (2008). *Adaptive Lehrkompetenz. Analyse und Struktur, Veränderung und Wirkung handlungssteuernden Lehrerwissens*. Waxmann.
- Becker, M., & McElvany, N. (2018). The interplay of gender and social background. A longitudinal study of interaction effects in reading attitudes and behaviour. *British Journal of Educational Psychology*, 88(4), 529–549. <https://doi.org/10.25656/01:16917>
- Bergold, S., Hastall, M. R., & Steinmayr, R. (2021). Do mass media shape stereotypes about intellectually gifted individuals? Two experiments on stigmatization effects from biased newspaper reports. *Gifted Child Quarterly*, 65(1), 75–94. <https://doi.org/10.1177/0016986220969393>
- Bergold, S., Wirthwein, L., & Steinmayr, R. (2020). Similarities and differences between intellectually gifted and average-ability students in school performance, motivation, and subjective well-being. *Gifted Child Quarterly*, 64, 285–303. <https://doi.org/10.1177/0016986220932533>
- Besand, A., & Jugel, D. (2015). Inklusion und politische Bildung – gemeinsam denken. In C. Dönges, W. Hilpert, & B. Zurstrassen (Eds.), *Didaktik der inklusiven politischen Bildung* (pp. 45–59). Bundeszentrale für politische Bildung. https://doi.org/10.1007/978-3-531-18671-9_4
- Boerma, I. E., Mol, S. E., & Jolles, J. (2016). Teacher perceptions affect boys' and girls' reading motivation differently. *Reading Psychology*, 37(4), 547–569. <https://doi.org/10.1080/02702711.2015.1072608>
- Brook, U., Watemberg, N., & Geva, D. (2000). Attitude and knowledge of attention deficit hyperactivity disorder and learning disability among high school teachers. *Patient Education and Counseling*, 40(3), 247–252. [https://doi.org/10.1016/s0738-3991\(99\)00080-4](https://doi.org/10.1016/s0738-3991(99)00080-4)
- Brühwiler, C. (Eds.) (2014). *Adaptive Lehrkompetenz und schulisches Lernen. Effekte handlungssteuernder Kognitionen von Lehrpersonen auf Unterrichtsprozesse und Lernergebnisse der Schülerinnen und Schüler*. Waxmann. <https://doi.org/10.1007/978-3-658-23187-3>
- Bundesministerium für Bildung und Forschung. (2022). *Bildung auf einen Blick 2022 – OECD-Indikatoren*. wbv Media. <https://read.oecd.org/10.1787/dd19b10a-de?format=pdf>

- Bundesverband Legasthenie und Dyskalkulie e. V. (2018). *Legasthenie: Ratgeber zum Thema Legasthenie – Erkennen und Verstehen*. https://www.bvl-legasthenie.de/images/static/pdfs/bvl/1_BVL_Ratgeber-Legasthenie_2018.pdf
- Camargo, S. P. H., Rispoli, M., Ganz, J., Hong, E. R., Davis, H., & Mason, R. (2016). Behaviorally based interventions for teaching social interaction skills to children with ASD in inclusive settings: A meta-analysis. *Journal of Behavioral Education, 25*, 223–248. <https://doi.org/10.1007/s10864-015-9240-1>
- Chodura, S., Kuhn, J. T., & Holling, H. (2015). Interventions for children with mathematical difficulties: A meta-analysis. *Zeitschrift für Psychologie, 223*(2), 129–144. <https://doi.org/10.1027/2151-2604/a000211>
- Crede, J., Wirthwein, L., Steinmayr, R., & Bergold, S. (2019). Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf im Bereich emotionale und soziale Entwicklung und ihre Peers im inklusiven Unterricht. *Zeitschrift für Pädagogische Psychologie, 1*–15. <https://doi.org/10.1024/1010-0652/a000244>
- de Boer, A., Pijl, S. J., & Minnaert, A. (2011). Regular primary schoolteachers' attitudes towards inclusive education: A review of the literature. *International Journal of Inclusive Education, 15*(3), 331–353. <http://doi.org/10.1080/13603110903030089>
- Depaepe, F., & König, J. (2018). General pedagogical knowledge, self-efficacy and instructional practice: Disentangling their relationship in pre-service teacher education. *Teaching and Teacher Education, 69*, 177–190. <https://doi.org/10.1016/j.tate.2017.10.003>
- Dersch, A. S., Heyder, A., & Eitel, A. (2022). Exploring the nature of teachers' math-gender stereotypes: The math-gender misconception questionnaire. *Frontiers in Psychology, 13*, 1728. <https://doi.org/10.3389/fpsyg.2022.820254>
- DeVries, J. M., Voß, S., & Gebhardt, M. (2018). Do learners with special education needs really feel included? Evidence from the Perception of Inclusion Questionnaire and Strengths and Difficulties Questionnaire. *Research in Developmental Disabilities, 83*, 28–36. <https://doi.org/10.1016/j.ridd.2018.07.007>
- Dicke, T., Parker, P. D., Holzberger, D., Kunina-Habenicht, O., Kunter, M., & Leutner, D. (2015). Beginning teachers' efficacy and emotional exhaustion: Latent changes, reciprocity, and the influence of professional knowledge. *Contemporary Educational Psychology, 41*, 62–72. <https://doi.org/10.1016/j.cedpsych.2014.11.003>
- Dignath, C., Rimm-Kaufman, S., van Ewijk, R., & Kunter, M. (2022). Teachers' beliefs about inclusive education and insights on what contributes to those beliefs: A meta-analytical study. *Educational Psychology Review, 34*, 2609–2660. <https://doi.org/10.1007/s10648-022-09695-0>
- Dodur, H. M. S., & Kumaş, Ö. A. (2020). Knowledge and beliefs of classroom teachers about dyslexia: the case of teachers in Turkey. *European Journal of Special Needs Education, 35*(1), 1–15. <https://doi.org/10.1080/0890/1903030089>

- 36(4), 593–609. <https://doi.org/10.1080/08856257.2020.1779980>
- Eckes, T., & Trautner, H. M. (2000). Developmental social psychology of gender: An integrative framework. In T. Eckes & H. M. Trautner (Eds.), *The developmental social psychology of gender* (pp. 3–32). Lawrence Erlbaum Associates Publishers. <https://doi.org/10.4324/9781410605245>
- Elbaum, B. (2002). The self-concept of students with learning disabilities: A meta-analysis of comparisons across different placements. *Learning Disabilities Research & Practice*, 17(4), 216–226. <https://doi.org/10.1111/1540-5826.00047>
- Ernst, H. M., Wittwer, J., & Voss, T. (2023). Do they know what they know? Accuracy in teacher candidates' self-assessments and its influencing factors. *British Educational Research Journal*, 00, 1–25. <https://doi.org/10.1002/berj.3860>
- Faulstich-Wieland, H. (2011). *Umgang mit Heterogenität und Differenz*. Schneider. <https://doi.org/10.14361/9783839429617-003>
- Fernandes, P. R. D. S., Jardim, J., & Lopes, M. C. D. S. (2021). The soft skills of special education teachers: Evidence from the literature. *Education Sciences*, 11(3), 125. <https://doi.org/10.3390/educsci11030125>
- Francis, R., Hawes, D. J., & Abbott, M. (2016). Intellectual giftedness and psychopathology in children and adolescents: A systematic literature review. *Exceptional Children*, 82(3), 279–302. <https://doi.org/10.1177/0014402915598779>
- Gäde, J. C., Schermelleh-Engel, K., & Werner, C. S. (2020). Klassische Methoden der Reliabilitätsschätzung. In H. Moosbrugger & A. Kelava (Eds.), *Testtheorie und Fragebogenkonstruktion* (pp. 305–334). Springer. https://doi.org/10.1007/978-3-662-61532-4_14
- Gebhardt, M., Rauch, D., Mang, J., Sälzer, C., & Stanat, P. (2012). Mathematische Kompetenz von Schülerinnen und Schülern mit Zuwanderungshintergrund. In M. Prenzel, C. Sälzer, E. Klieme, & O. Köller (Eds.), *PISA 2012: Fortschritte und Herausforderungen in Deutschland* (pp. 275–308). Waxmann
- Gebhardt, M., Schwab, S., Hessels, M. G., & Nusser, L. (2015). Einstellungen und Selbstwirksamkeit von Lehrerinnen und Lehrern zur schulischen Inklusion in Deutschland-eine Analyse mit Daten des nationalen Bildungspanels Deutschlands (NEPS). *Empirische Pädagogik*, 29(2), 211–229.
- Gerardi, K., Goette, L., & Meier, S. (2013). Numerical ability predicts mortgage default. *Proceedings of the National Academy of Sciences*, 110(28), 11267–11271. <https://doi.org/10.1073/pnas.1220568110>
- Gesel, S. A., LeJeune, L. M., Chow, J. C., Sinclair, A. C., & Lemons, C. J. (2021). A meta-analysis of the impact of professional development on teachers' knowledge, skill, and self-efficacy in data-based decision-making. *Journal of Learning Disabilities*, 54(4), 269–283. <https://doi.org/10.1177/0022219420970196>

- Gess-Newsome, J., Taylor, J. A., Carlson, J., Gardner, A. L., Wilson, C. D., & Stuhsatz, M. A. (2017). Teacher pedagogical content knowledge, practice, and student achievement. *International Journal of Science Education*, 41(7), 944–963. <https://dx.doi.org/10.1080/09500693.2016.1265158>
- Ghanizadeh, A., Bahredar, M. J., & Moeini, S. R. (2006). Knowledge and attitudes towards attention deficit hyperactivity disorder among elementary school teachers. *Patient Education and Counseling*, 63(1-2), 84–88. <https://doi.org/10.1016/j.pec.2005.09.002>
- Ginsburg, G. S., Pella, J. E., Ogle, R. R., DeVito, A., Raguin, K., & Chan, G. (2021). Teacher knowledge of anxiety and use of anxiety reduction strategies in the classroom. *Journal of Psychologists and Counsellors in Schools*, 1–11. <https://doi.org/10.1017/jgc.2021.26>
- Göransson, K., & Nilholm, C. (2014). Conceptual diversities and empirical shortcomings – a critical analysis of research on inclusive education. *European Journal of Special Needs Education*, 29, 265–280. <https://doi.org/10.1080/08856257.2014.933545>
- Grosche, M. (2015). Was ist Inklusion? Ein Diskussions- und Positionsartikel zur Definition von Inklusion aus Sicht der empirischen Bildungsforschung. In P. Kuhl, P. Stanat, B. Lütje-Klose, C. Gresch, H. A. Pant & M. Prenzel (Eds.), *Inklusion von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf in Schulleistungserhebungen* (pp. 17–39). Springer.
- Guignard, J. H., Bacro, F., & Guimard, P. (2021). School life satisfaction and peer connectedness of intellectually gifted adolescents in France: Is there a labeling effect?. *New Directions for Child and Adolescent Development*, 2021(179), 59–74. <https://doi.org/10.1002/cad.20448>
- Haimour, A. I., & Obaidat, Y. F. (2013). School teachers' knowledge about autism in Saudi Arabia. *World Journal of Education*, 3(5), 45–56. <https://doi.org/10.5430/wje.v3n5p45>
- Hannover, B., & Wolter, I. (2021). Schule und Geschlecht. In T. Hascher, T.-S. Idel, & W. Helsper (Eds.), *Handbuch Schulforschung* (pp. 1–21). Springer. https://doi.org/10.1007/978-3-658-24734-8_38-1
- Hart, J. E., & More, C. M. (2013). Investigating the impact of technology on pre-service teacher knowledge of autism spectrum disorder. *Education and Training in Autism and Developmental Disabilities*, 48(4), 504–513.
- Hartig, J., Frey, A., & Jude, N. (2020). Validität von Testwertinterpretationen. In H. Moosbrugger & A. Kelava (Eds.), *Testtheorie und Fragebogenkonstruktion* (pp. 529–545). Springer. https://doi.org/10.1007/978-3-662-61532-4_21
- Headley, C. (2013). Teachers' knowledge of anxiety and identification of excessive anxiety in children. *Australian Journal of Teacher Education*, 38(5), 48–66. <https://doi.org/10.14221/ajte.2013v38n5.2>
- Hehir, T., Grindal, T., Freeman, B., Lamoreau, R., Borquaye, Y., & Burke, S. (2016). *A summary of the evidence on inclusive education*. Abt Associates.

- Heller, K. A., Reimann, R., & Senfter, A. (2005). *Hochbegabung im Grundschulalter: Erkennen und Fördern*. Lit Verlag.
- Henke, T., Bosse, S., Lambrecht, J., Jäntschi, C., Jaeuthe, J., & Spörer, N. (2017). Mittendrin oder nur dabei? Zum Zusammenhang zwischen sonderpädagogischem Förderbedarf und sozialer Partizipation von Grundschülerinnen und Grundschulern. *Zeitschrift für Pädagogische Psychologie*, 31(2), 111–123. <https://doi.org/10.1024/1010-0652/a000196>
- Hess, D., Ruland, M., Meyer, M., & Steinwede, J. (2019). *Schulische Inklusion. Untersuchung zu Einstellungen zu schulischer Inklusion und Wirkungen im Bildungsverlauf*. Institut für angewandte Sozialwissenschaften GmbH. https://delivery-aktion-mensch.stylelabs.cloud/api/public/content/Studie_Schulische_Inklusion_Langfassung_barrierefrei.pdf?v=81f934bb
- Heyder, A., Bergold, S., & Steinmayr, R. (2018). Teachers' knowledge about intellectual giftedness: A first look at levels and correlates. *Psychology Learning & Teaching*, 17(1), 27–44. <https://doi.org/10.1177/1475725717725493>
- Heyder, A., Südkamp, A., & Steinmayr, R. (2020). How are teachers' attitudes toward inclusion related to the social-emotional school experiences of students with and without special educational needs?. *Learning and Individual Differences*, 77, 101776. <https://doi.org/10.1016/j.lindif.2019.101776>
- Heyder, A., van Hek, M., & Van Houtte, M. (2021). When gender stereotypes get male adolescents into trouble: A longitudinal study on gender conformity pressure as a predictor of school misconduct. *Sex Roles*, 84, 61–75. <https://doi.org/10.1007/s11199-020-01147-9>
- Heyder, A., Vaskova, A., Hußmann, A., & Steinmayr, R. (2018). Wissen von angehenden Lehrkräften zu Diversität im Kontext schulischer Inklusion – Die Entwicklung eines Wissenstests und erste Ergebnisse. In S. Hußmann & B. Welzel (Eds.), *DoProfiL – Das Dortmunder Profil für inklusionsorientierte Lehrerinnen- und Lehrerbildung* (pp. 267–278). Waxmann.
- Hohenstein, F., Kleickmann, T., Zimmermann, F., Köller, O., & Möller, J. (2017). Erfassung von pädagogischem und psychologischem Wissen in der Lehramtsausbildung: Entwicklung eines Messinstruments. *Zeitschrift für Pädagogik*, 63, 91–112. <https://doi.org/10.1007/s11618-014-0563-9>
- Horn, W., Lukesch, H., Mayrhofer, S., & Kormann, A. (2004). *PSB-R 6-13, Prüfbildungssystem für Schul- und Bildungsberatung für 6. bis 13. Klassen (rev. ed.)*. Hogrefe.
- Horne, P. E., & Timmons, V. (2009). Making it work: teachers' perspectives on inclusion. *International Journal of Inclusive Education*, 13(3), 273–286. <https://doi.org/10.1080/13603110701433964>

- Huber, C., & Wilbert, J. (2012). Soziale Ausgrenzung von Schülern mit sonderpädagogischem Förderbedarf und niedrigen Schulleistungen im gemeinsamen Unterricht. *Empirische Sonderpädagogik*, 4(2), 147–165.
- Institut der deutschen Wirtschaft. (2022, Mai). *MINT-Frühjahrsreport 2022. Demografie, Dekarbonisierung und Digitalisierung erhöhen MINT-Bedarf – Zuwanderung stärkt MINT-Fachkräfteangebot und Innovationskraft.* https://www.iwkoeln.de/fileadmin/user_upload/Studien/Gutachten/PDF/2022/MINT-Fr%C3%BChjahrsreport_2022.pdf
- Jones, T. W. (2013). Equally cursed and blessed: Do gifted and talented children experience poorer mental health and psychological well-being?. *Educational and Child Psychology*, 30(2), 44–66. <https://doi.org/10.53841/bpsecp.2013.30.2.44>
- Kalka, D., & Lockiewicz, M. (2018). Happiness, life satisfaction, resiliency and social support in students with dyslexia. *International Journal of Disability, Development and Education*, 65(5), 493–508. <https://doi.org/10.1080/1034912x.2017.1411582>
- Kart, A., & Kart, M. (2021). Academic and social effects of inclusion on students without disabilities: A review of the literature. *Education Sciences*, 11(1), 16. <https://doi.org/10.3390/educsci11010016>
- Keller, L., Preckel, F., Eccles, J. S., & Brunner, M. (2021). Top-performing math students in 82 countries: An integrative data analysis of gender differences in achievement, achievement profiles, and achievement motivation. *Journal of Educational Psychology*, 114(5), 966–991. <https://doi.org/10.1037/edu0000685>
- Keller, M. M., Neumann, K., & Fischer, H. E. (2017). The impact of physics teachers' pedagogical content knowledge and motivation on students' achievement and interest. *Journal of Research in Science Teaching*, 54(5), 586–614. <https://doi.org/10.1002/tea.21378>
- Klusmann, U., Kunter, M., Voss, T., & Baumert, J. (2012). Berufliche Beanspruchung angehender Lehrkräfte: Die Effekte von Persönlichkeit, pädagogischer Vorerfahrung und professioneller Kompetenz. *Zeitschrift für Pädagogische Psychologie*, 26(4), 275–290. <https://doi.org/10.1024/1010-0652/a000078>
- Klusmann, U., Richter, D., & Lüdtke, O. (2016). Teachers' emotional exhaustion is negatively related to students' achievement: Evidence from a large-scale assessment study. *Journal of Educational Psychology*, 108(8), 1193–1203. <https://dx.doi.org/10.1037/edu0000125>
- König, J., & Blömeke, S. (2009). Pädagogisches Wissen von angehenden Lehrkräften. *Zeitschrift für Erziehungswissenschaft*, 12(3), 499–527. <https://doi.org/10.1007/s11618-009-0085-z>
- König, J., & Blömeke, S. (2010). *Pädagogisches Unterrichtswissen (PUW) – Dokumentation der Kurzfassung des TEDS-M Testinstruments zur Kompetenzmessung in der ersten*

Phase der Lehrerausbildung. Humboldt-Universität Berlin.

- König, J., Gerhard, K., Melzer, C., Rühl, A.-M., Zenner, J., & Kaspar, K. (2017). Erfassung von pädagogischem Wissen für inklusiven Unterricht bei angehenden Lehrkräften: Testkonstruktion und Validierung. *Unterrichtswissenschaft*, 45(4), 223–242. <https://doi.org/10.3726/pr012019.0004>
- König, J., & Pflanzl, B. (2016). Is teacher knowledge associated with performance? On the relationship between teachers' general pedagogical knowledge and instructional quality. *European Journal of Teacher Education*, 39(4), 419–436. <https://doi.org/10.1080/02619768.2016.1214128>
- Krämer, P., Przibilla, B., & Grosche, M. (2016). *Indikatorenschema zur operationalen Definition schulischer Inklusion.* Lizenziert unter CC-BY. <https://doi.org/10.6084/m9.figshare.3384670>
- Krämer, S., Möller, J., & Zimmermann, F. (2021). Inclusive education of students with general learning difficulties: A meta-analysis. *Review of Educational Research*, 91(3), 432–478. <https://doi.org/10.3102/0034654321998072>
- Krawinkel, S., Südkamp, A. S., & Tröster, H. (2017). Soziale Partizipation in inklusiven Grundschulklassen. Bedeutung von Klassen- und Lehrkraftmerkmalen. *Empirische Sonderpädagogik*, 3, 277–295.
- Krüger-Potratz, M. (2011). Intersektionalität. In H. Faulstich-Wieland (Eds.), *Umgang mit Heterogenität und Differenz* (pp. 183–200). Schneider. <https://doi.org/10.14361/9783839429617-003>
- Krull, J., Wilbert, J., & Hennemann, T. (2014). Soziale Ausgrenzung von Erstklässlerinnen und Erstklässlern mit sonderpädagogischem Förderbedarf im Gemeinsamen Unterricht. *Empirische Sonderpädagogik*, 6(1), 59–75.
- Kunina-Habenicht, O., Lohse-Bossenz, H., Kunter, M., Dicke, T., Förster, D., Gößling, J., Schulze, F., Baumert, J., Leutner, D., & Terhart, E. (2012). Welche bildungswissenschaftlichen Inhalte sind wichtig in der Lehrerbildung? *Zeitschrift für Erziehungswissenschaft*, 15, 649–682. <https://doi.org/10.1007/s11618-012-0324-6>
- Kunina-Habenicht, O., Maurer, C., Wolf, K., Holzberger, D., Schmidt, M., Dicke, T., Teuber, Z., Koc-Januchta, M., Lohse-Bossenz, H., Leutner, D., Seidel, T., & Kunter, M. (2020). Der BilWiss-2.0-Test – Ein revidierter Test zur Erfassung des bildungswissenschaftlichen Wissens von (angehenden) Lehrkräften. *Diagnostica*, 66(2), 80–92. <https://doi.org/10.1026/0012-1924/a000238>
- Kunter, M., Baumert, J., Blum, W., Klusmann, U., Krauss, S. & Neubrand, M. (Eds.) (2013). Cognitive activation in the mathematics classroom and professional competence of teachers: Results from the COACTIV project. Springer. <https://doi.org/10.1007/978-1-4614-5149-5>
- Kunter, M., Kleickmann, T., Klusmann, U., & Richter, D. (2013). In M. Kunter, J. Baumert,

- W. Blum, U. Klusmann, S. Krauss & M. Neubrand (Eds.), *The development of teachers' professional competence* (pp. 63–77). Springer. <https://doi.org/10.1007/978-1-4614-5149-5>
- Kunter, M., Klusmann, U., Baumert, J., Richter, D., Voss, T., & Hachfeld, A. (2013). Professional competence of teachers: effects on instructional quality and student development. *Journal of Educational Psychology*, 105(3), 805–820. <https://doi.org/10.1037/a0032583>
- Kunz, A., Luder, R., & Moretti, M. (2010). Die Messung von Einstellungen zur Integration (EZI). *Empirische Sonderpädagogik*, 2(3), 83–94.
- Kuyini, A. B., & Desai, I. (2007). Principals' and teachers' attitudes and knowledge of inclusive education as predictors of effective teaching practices in Ghana. *Journal of Research in Special Educational Needs*, 7(2), 104–113. <https://doi.org/10.1111/j.1471-3802.2007.00086.x>
- Lacruz-Pérez, I., Sanz-Cervera, P., & Tárraga-Mínguez, R. (2021). Teachers' attitudes toward educational inclusion in Spain: A systematic review. *Education Sciences*, 11(2), 58. <https://doi.org/10.3390/educsci11020058>
- Landesregierung Nordrhein-Westfalen. (2020). *Lese-Rechtschreib-Schwäche und Dyskalkulie an den Schulen in NRW. Antwort der Landesregierung auf die Kleine Anfrage 4529 vom 7. Oktober 2020.* https://www.landtag.nrw.de%2Fportal%2FWWW%2Fdokumentenarchiv%2FDokument%2FMMD17-11816.pdf&usg=AOvVaw1-5_kcQTBWako3RNhyStmy
- Lang, E., Grittner, F., Rehle, C., & Hartinger, A. (2010). Das Heterogenitätsverständnis von Lehrkräften im jahrgangsgemischtem Unterricht der Grundschule. In J. Hagedorn, V. Schurt, C. Steber & W. Waburg (Eds.), *Ethnizität, Geschlecht, Familie und Schule* (pp. 315–331). VS Verlag für Sozialwissenschaften.
- Lauermann, F., & König, J. (2016). Teachers' professional competence and wellbeing – Understanding the links between general pedagogical knowledge, self-efficacy and burnout. *Learning and Instruction*, 45, 9–19. <http://doi.org/10.1016/j.learninstruc.2016.06.006>
- Laursen, B., Richmond, A., Kiuru, N., Lerkkanen, M. K., & Poikkeus, A. M. (2021). Off on the wrong foot: Task avoidance at the outset of primary school anticipates academic difficulties and declining peer acceptance. *European Journal of Developmental Psychology*, 1–15. <https://doi.org/10.1080/17405629.2021.1936491>
- Lautenbach, F., & Heyder, A. (2019). Changing attitudes to inclusion in preservice teacher education: A systematic review. *Educational Research*, 61(2), 231–253. <https://doi.org/10.1080/00131881.2019.1596035>
- Lenske, G., Thillmann, H., Wirth, J., Dicke, T., & Leutner, D. (2015). Pädagogisch-psychologisches Professionswissen von Lehrkräften: Evaluation des ProwiN-Tests.

- Zeitschrift für Erziehungswissenschaft*, 18(2), 225–245. <https://doi.org/10.1007/s11618-015-0627-5>
- Leonardi, M. M., Di Blasi, F. D., Savelli, E., & Buono, S. (2021). Reading and spelling disorders in a school-based population screening in Sicily (Italy). *Dyslexia*, 27(4), 452–467. <https://doi.org/10.1002/dys.1694>
- Lu, M., Zou, Y., Chen, X., Chen, J., He, W., & Pang, F. (2020). Knowledge, attitude and professional self-efficacy of Chinese mainstream primary school teachers regarding children with autism spectrum disorder. *Research in Autism Spectrum Disorders*, 72, 101513. <https://doi.org/10.1016/j.rasd.2020.101513>
- Malva, L., Leijen, Ä., & Baucal, A. (2020). Towards measuring teachers' general pedagogical knowledge – A mixed method investigation of a pilot test. *Studies in Educational Evaluation*, 64, 100815. <https://doi.org/10.1016/j.stueduc.2019.100815>
- McGeown, S., Goodwin, H., Henderson, N., & Wright, P. (2012). Gender differences in reading motivation: Does sex or gender identity provide a better account?. *Journal of Research in Reading*, 35(3), 328–336. <https://doi.org/10.1111/j.1467-9817.2010.01481.x>
- Ministerium für Schule und Bildung des Landes Nordrhein-Westfalen. (2016). *Verordnung über die sonderpädagogische Förderung, den Hausunterricht und die Schule für Kranke*. https://www.duesseldorf.de/fileadmin/Amt40/PDF/A-Z/AO_SF.pdf
- Ministerium für Schule und Bildung des Landes Nordrhein-Westfalen. (2022). *Verwaltungsvorschriften zur Verordnung zur Ausführung des § 93 Abs. 2 Schulgesetz (AVO-Richtlinien 2022/2023 - AVO-RL)*. <https://bass.schulwelt.de/pdf/6218.pdf?20230126143627>
- Mintz, J., Hick, P., Solomon, Y., Matziari, A., Ó'Murchú, F., Hall, K., Cahill, K., Curtin, C., Anders, J., & Margariti, D. (2020). The reality of reality shock for inclusion: How does teacher attitude, perceived knowledge and self-efficacy in relation to effective inclusion in the classroom change from the pre-service to novice teacher year?. *Teaching and Teacher Education*, 91, 103042. <https://doi.org/10.1016/j.tate.2020.103042>
- Monei, T., & Pedro, A. (2017). A systematic review of interventions for children presenting with dyscalculia in primary schools. *Educational Psychology in Practice*, 33(3), 277–293. <https://doi.org/10.1080/02667363.2017.1289076>
- Moosbrugger, H., & Kelava, A. (2020). Qualitätsanforderungen an Tests und Fragebogen („Gütekriterien“). In H. Moosbrugger & A. Kelava (Eds.), *Testtheorie und Fragebogenkonstruktion* (pp. 13–38). Springer. https://doi.org/10.1007/978-3-662-61532-4_2
- Moosbrugger, H., Schermelleh-Engel, K., Gädde, J. C., & Kelava, A. (2020). Testtheorien im Überblick. In H. Moosbrugger & A. Kelava (Eds.), *Testtheorie und Fragebogenkonstruktion* (pp. 251–273). Springer. https://doi.org/10.1007/978-3-662-61532-4_12

- Morrisette, V., Jesme, S., & Hunter, C. (2017). Teacher and administrator perceptions of gender in the classroom. *Educational Studies*, 44(3), 295–312. <https://doi.org/10.1080/03055698.2017.1373631>
- Mostafa, T., & Schwabe, M. (2019). *Programme for international student assessment (PISA) – Ländernotiz: PISA 2018 Ergebnisse*. Organisation for Economic Co-operation and Development. https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&ved=2ahUKEwjLyfuzsvT8AhVnSfEDHZwUA6IQFnoECCIQAQ&url=https%3A%2F%2Fwww.oecd.org%2Fpisa%2Fpublications%2FPISA2018_CN_DEU_German.pdf&usg=AOvVaw3I1Mi-t1a_N1cy6YXaScBM
- Muntoni, F., Dunekacke, S., Heinze, A., & Retelsdorf, J. (2019). Geschlechtsspezifische Erwartungseffekte in Mathematik. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 51(2), 84–96. <https://doi.org/10.1026/0049-8637/a000212>
- Nelson, J. R., Benner, G. J., Lane, K., & Smith, B. W. (2004). Academic achievement of K-12 students with emotional and behavioral disorders. *Exceptional Children*, 71(1), 59–73. <https://doi.org/10.1177/001440290407100104>
- Nespor, J. (1987). The role of beliefs in the practice of teaching. *Journal of Curriculum Studies*, 19, 317–328. <https://doi.org/10.1080/0022027870190403>
- Ness, M. K., & Southall, G. (2010). Preservice Teachers' Knowledge of and Beliefs About Dyslexia. *Journal of Reading Education*, 36(1), 36–43.
- Novita, S. (2016). Secondary symptoms of dyslexia: A comparison of self-esteem and anxiety profiles of children with and without dyslexia. *European Journal of Special Needs Education*, 31(2), 279–288. <https://doi.org/10.1080/08856257.2015.1125694>
- Ogurlu, U. (2021). A meta-analytic review of emotional intelligence in gifted individuals: A multilevel analysis. *Personality and Individual Differences*, 171, 110503. <https://doi.org/10.1016/j.paid.2020.110503>
- Pajares, M. F. (1992). Teachers' beliefs and educational research: Cleaning up a messy construct. *Review of Educational Research*, 62(3), 307–332. <https://doi.org/10.3102/00346543062003307>
- Parsons, S., & Bynner, J. (2005). *Does numeracy matter more?*. National Research and Development Centre for Adult Literacy and Numeracy.
- Peries, W. N. N., Indrarathne, B., Jayamanne, B. D. W., Wickramasekara, T. D., Alwis, K. A. C., & Jayatilleke, A. U. (2021). Primary school teachers' readiness in identifying children with dyslexia: A national survey in Sri Lanka. *Dyslexia*, 27(4), 486–509. <https://doi.org/10.1002/dys.1696>
- Perold, M., Louw, C., & Kleynhans, S. (2010). Primary school teachers' knowledge and misperceptions of attention deficit hyperactivity disorder (ADHD). *South African Journal of Education*, 30, 457–473. <https://doi.org/10.15700/saje.v30n3a364>

- Pflanzl, B., Thomas, A., & Matischek-Jauk, M. (2013). Pädagogisches Wissen und pädagogische Handlungskompetenz. *Erziehung und Unterricht*, 163(1/2), 1–8.
- Preckel, F., & Vock, M. (2021). *Hochbegabung: Ein Lehrbuch zu Grundlagen, Diagnostik und Fördermöglichkeiten* (2nd rev. ed.). Hogrefe. <https://doi.org/10.1026/2191-9186/a000223>
- Ramberg, J., Lénárt, A., & Watkins, A. (Eds.) (2018). *European Agency Statistics on Inclusive Education: 2016 Dataset Cross-Country Report*. European Agency for Special Needs and Inclusive Education. <https://www.european-agency.org/resources/publications/european-agency-statistics-inclusive-education-2016-dataset-cross-country>
- Riddick, B., Sterling, C., Farmer, M., & Morgan, S. (1999). Self-esteem and anxiety in the educational histories of adult dyslexic students. *Dyslexia*, 5(4), 227–248. [https://doi.org/10.1002/\(sici\)1099-0909\(199912\)5:4<227::aid-dys146>3.0.co;2-6](https://doi.org/10.1002/(sici)1099-0909(199912)5:4<227::aid-dys146>3.0.co;2-6)
- Rix, J., Hall, K., Nind, M., Sheehy, K., & Wearmouth, J. (2009). What pedagogical approaches can effectively include children with special educational needs in mainstream classrooms? A systematic literature review. *Support for Learning*, 24(2), 86–94. <https://doi.org/10.1111/j.1467-9604.2009.01404.x>
- Rost, D. H. (2000). *Hochbegabte und hochleistende Jugendliche*. Waxmann.
- Rost, D. H., & Hanses, P. (1997). Wer nichts leistet, ist nicht begabt? Zur Identifikation. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 29(2), 167–177.
- Rowan, L., Bourke, T., L'Estrange, L., Lunn Brownlee, J., Ryan, M., Walker, S., & Churchward, P. (2021). How does initial teacher education research frame the challenge of preparing future teachers for student diversity in schools? A systematic review of literature. *Review of Educational Research*, 91(1), 112–158. <https://doi.org/10.3102/0034654320979171>
- Ruberg, C., & Porsch, R. (2017). Einstellungen von Lehramtsstudierenden und Lehrkräften zur schulischen Inklusion. Ein systematisches Review deutschsprachiger Forschungsarbeiten. *Zeitschrift für Pädagogik*, 63(4), 393–415.
- Ruijs, N. M., & Peetsma, T. T. (2009). Effects of inclusion on students with and without special educational needs reviewed. *Educational Research Review*, 4(2), 67–79. <https://doi.org/10.1016/j.edurev.2009.02.002>
- Sainio, P. J., Eklund, K. M., Ahonen, T. P., & Kiuru, N. H. (2019). The role of learning difficulties in adolescents' academic emotions and academic achievement. *Journal of Learning Disabilities*, 1–12. <https://doi.org/10.1177/0022219419841567>
- Sanz-Cervera, P., Fernández-Andrés, M. I., Pastor-Cerezuela, G., & Tárraga-Mínguez, R. (2017). Pre-service teachers' knowledge, misconceptions and gaps about autism spectrum disorder. *Teacher Education and Special Education*, 40(3), 212–224. <https://doi.org/10.1177/0888406417700963>

- Scharenberg, K. (2013). Heterogenität in der Schule. Definitionen, Forschungsbefunde, Konzeptionen und Perspektiven für die empirische Bildungsforschung. In N. McElvany, M. M. Gebauer, W. Bos, & H.-G. Holtappels (Eds.), *Jahrbuch der Schulentwicklung Band 17* (pp. 10–49). Beltz Juventa.
- Schermelleh-Engel, K., & Gåde, J. C. (2020). Modellbasierte Methoden der Reliabilitätsschätzung. In H. Moosbrugger & A. Kelava (Eds.), *Testtheorie und Fragebogenkonstruktion* (pp. 335–368). Springer. https://doi.org/10.1007/978-3-662-61532-4_15
- Schmiedeler, S. (2013). Wissen und Fehlannahmen von deutschen Lehrkräften über die Aufmerksamkeitsdefizit-/Hyperaktivitätsstörung (ADHS). *Psychologie in Erziehung und Unterricht*, 60, 143–153. <https://doi.org/10.2378/peu2013.art12d>
- Schwab, S., Hellmich, F., & Görel, G. (2017). Self-efficacy of prospective Austrian and German primary school teachers regarding the implementation of inclusive education. *Journal of Research in Special Educational Needs*, 17(3), 205–217. <https://doi.org/10.1111/1471-3802.12379>
- Schwab, S., Rossmann, P., Tanzer, N., Hagn, J., Oitzinger, S., Thurner, V., & Wimberger, T. (2015). Schulisches Wohlbefinden von SchülerInnen mit und ohne sonderpädagogischen Förderbedarf. *Zeitschrift für Kinder-und Jugendpsychiatrie und Psychotherapie*, 43(4), 265–274. <https://doi.org/10.1024/1422-4917/a000363>
- Sciutto, M. J., Terjesen, M. D., & Bender Frank, A. S. (2000). Teachers' knowledge and misperceptions of attention-deficit/hyperactivity disorder. *Psychology in the Schools*, 37(2), 115–122. [https://doi.org/10.1002/\(sici\)1520-6807\(200003\)37:2<115::aid-pits3>3.0.co;2-5](https://doi.org/10.1002/(sici)1520-6807(200003)37:2<115::aid-pits3>3.0.co;2-5)
- Sciutto, M. J., Terjesen, M. D., Kučerová, A., Michalová, Z., Schmiedeler, S., Antonopoulou, K., Shaker, N. Z., Lee, J.-Y., Alkahtani, K., & Rossouw, J. (2016). Cross-national comparisons of teachers' knowledge and misconceptions of ADHD. *International Perspectives in Psychology*, 5(1), 34–50. <https://doi.org/10.1037/ipp0000045>
- Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (1991, July). *Förderung von Schülerinnen und Schülern bei besonderen Schwierigkeiten im Erlernen des Lesens und Rechtschreibens (LRS)*. <https://bass.schulwelt.de/pdf/280.pdf?20221217083453>
- Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2007, November). Grundsätze zur Förderung von Schülerinnen und Schülern mit besonderen Schwierigkeiten im Lesen und Rechtschreiben oder im Rechnen. https://www.kmk.org%2Ffileadmin%2Fveroeffentlichungen_beschluesse%2F2003%2F2003_12_04-Lese-Rechtschreibschwaeche.pdf&usg=AOvVaw0ZRK_n-IxpHDS-6qKnB-e
- Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik

- Deutschland. (2014, June). *Standards für die Lehrerbildung: Bildungswissenschaften*. http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2004/2004_12_16-Standards-Lehrerbildung-Bildungswissenschaften.pdf
- Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2016, December). *Sonderpädagogische Förderung in allgemeinen Schulen (ohne Förderschulen) 2015/2016*. https://www.kmk.org/fileadmin/Dateien/pdf/Statistik/Dokumentationen/Aus_SoPae_Int_2015.pdf
- Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2022, January). *Sonderpädagogische Förderung in Schulen 2011 bis 2020*. https://www.kmk.org/fileadmin/Dateien/pdf/Statistik/Dokumentationen/Dok231_SoPaeFoe_2020.pdf
- Semeraro, C., Coppola, G., Taurino, A., & Cassibba, R. (2020). Understanding the impact of diagnosis: Emotional well-being, peers and teachers. In D. Lucangeli (Ed.), *Understanding Dyscalculia* (pp. 94–119). Routledge. <https://doi.org/10.4324/9780429423581-6>
- Shalev, R. S., Manor, O., Auerbach, J., & Gross-Tsur, V. (1998). Persistence of developmental dyscalculia: What counts?: Results from a 3-year prospective follow-up study. *The Journal of Pediatrics*, 133(3), 358–362. [https://doi.org/10.1016/s0022-3476\(98\)70269-0](https://doi.org/10.1016/s0022-3476(98)70269-0)
- Sheppard, M. E., & Wieman, R. (2020). What do teachers need? Math and special education teacher educators' perceptions of essential teacher knowledge and experience. *The Journal of Mathematical Behavior*, 59, 100798. <https://doi.org/10.1016/j.jmathb.2020.100798>
- Shulman, L. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review*, 57(1), 1–23.
- Soriano-Ferrer, M., & Echegaray-Bengoa, J. A. (2014). A scale of knowledge and beliefs about developmental dyslexia: Scale development and validation. *Procedia-Social and Behavioral Sciences*, 132, 203–208. <https://doi.org/10.1016/j.sbspro.2014.04.299>
- Sousa, P., Dias, P. C., & Cadime, I. (2017). Predictors of primary school teachers' knowledge about developmental dyscalculia. *European Journal of Special Needs Education*, 32(2), 204–220. <https://doi.org/10.1080/08856257.2016.1216635>
- Sparfeldt, J. R., & Rost, D. H. (2012). Underachievement: Diskrepanz von Leistungspotential und schulischer Leistung. *Erziehung & Unterricht*, 162, 435–441.
- Spörer, N., Maaz, K., Vock, M., Schründer-Lenzen, A., Luka, T., Bosse, S., Vogel, J., & Jäntsich, C. (2015). Lernen in der inklusiven Grundschule. Zusammenhänge zwischen fachlichen Kompetenzen, Sozialklima und Facetten des Selbstkonzepts. *Unterrichtswissenschaft*, 43(1), 22–35. https://doi.org/10.1007/978-3-658-06955-1_45

- Steinmayr, R., Heyder, A., & Tometten, L. (2022). DiWi. Test zum Wissen über verschiedene Diversitätsbereiche von (angehenden) Lehrkräften. Verfahrensdokumentation, Fragebogen Lang- und Kurzversion, Auswertungssyntax. In Leibniz-Institut für Psychologie (ZPID) (Ed.), *Open Test Archive*. ZPID. <https://doi.org/10.23668/psycharchives.5136>
- Steinmayr, R., & Kessels, U. (2017). Good at school= successful on the job? Explaining gender differences in scholastic and vocational success. *Personality and Individual Differences*, *105*, 107–115. <https://dx.doi.org/10.1016/j.paid.2016.09.032>
- Steinmayr, R., Weidinger, A. F., Heyder, A., & Bergold, S. (2019). Warum schätzen Mädchen ihre mathematischen Kompetenzen geringer ein als Jungen?. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, *51*(2), 71–83. <https://doi.org/10.1026/0049-8637/a000213>
- Szumski, G., Smogorzewska, J., & Karwowski, M. (2017). Academic achievement of students without special educational needs in inclusive classrooms: A meta-analysis. *Educational Research Review*, *21*, 33–54. <https://doi.org/10.1016/j.edurev.2017.02.004>
- Tasdemir, M. (2010). A comparison of multiple-choice tests and true false tests. *Journal of Instructional Psychology*, *37*(3), 258–266.
- Terman L. M. (1925). *Genetic studies of genius: Volume I. Mental and physical traits of a thousand gifted children*. Stanford University Press. <https://doi.org/10.1126/science.64.1650.160>
- Thomas, K., Schulte-Körne, G. & Hasselhorn, M. (2015). Stichwort – Entwicklungsstörungen schulischer Fertigkeiten. *Zeitschrift für Erziehungswissenschaft*, *18*, 431–451. <https://doi.org/10.1007/s11618-015-0642-6>
- Tristani, L., & Bassett-Gunter, R. (2020). Making the grade: Teacher training for inclusive education: A systematic review. *Journal of Research in Special Educational Needs*, *20*(3), 246–264. <https://doi.org/10.1111/1471-3802.12483>
- United Nations Educational, Scientific and Cultural Organization. (1994). *The Salamanca statement and framework for action on special needs education*. <https://unesdoc.unesco.org/ark:/48223/pf0000098427>
- Voss, T., Kunina-Habenicht, O., Hoehne, V., & Kunter, M. (2015). Stichwort Pädagogisches Wissen von Lehrkräften: Empirische Zugänge und Befunde. *Zeitschrift für Erziehungswissenschaft*, *18*(2), 187–223. <https://doi.org/10.1007/s11618-015-0626-6>
- Voss, T., Kunter, M., & Baumert, J. (2011). Assessing teacher candidates' general pedagogical/psychological knowledge: Test construction and validation. *Journal of Educational Psychology*, *103*(4), 952–969. <https://doi.org/10.1037/a0025125>
- Voss, T., Zachrich, L., Fauth, B., & Wittwer, J. (2022). The same yet different? Teaching quality differs across a teacher's classes, but teachers with higher knowledge make teaching quality more similar. *Learning and Instruction*, *80*, 101614.

- <https://doi.org/10.1016/j.learninstruc.2022.101614>
- Voyer, D., & Voyer, S. D. (2014). Gender differences in scholastic achievement: a meta-analysis. *Psychological Bulletin*, 140(4), 1174–1204. <https://dx.doi.org/10.1037/a0036620>
- Washburn, E. K., Mulcahy, C. A., Musante, G., & Joshi, R. (2017). Novice Teachers' Knowledge of Reading-Related Disabilities and Dyslexia. *Learning Disabilities: A Contemporary Journal*, 15(2), 169–191.
- Wildbrett, J., Lohse-Bossenz, H., & Dörfler, T. (2022). Empirische Arbeit: Vignettenbasierte Messung pädagogisch-psychologischen Professionswissens bei Lehramtsstudierenden. *Psychologie in Erziehung und Unterricht*, 69(4), 278–291. <https://doi.org/10.2378/peu2022.art20d>
- Wischer, B. (2007). Wie sollen LehrerInnen mit Heterogenität umgehen? Über 'programmatische Fallen' im aktuellen Reformdiskurs. *Die deutsche Schule*, 99(4), 422–433.
- Wocken, H. (2009). Inklusion & Integration. Ein Versuch, die Integration vor der Abwertung und die Inklusion vor Träumereien zu bewahren. In A.-D. Stein, I. Niediek, & S. Krach (Eds.), *Integration und Inklusion auf dem Wege ins Gemeinwesen* (pp. 204–234). Klinkhardt. https://doi.org/10.1007/978-3-531-91187-8_11
- Wolf, K. (2017). Prädiktive Validität von Schulnoten. In K. Wolf (Hrsg.), *Prädiktoren für den multikriterialen Berufserfolg von Lehrkräften* (S. 9 – 26). Springer. http://doi.org/10.1007%2F978-3-658-16819-3_3
- Wong, T. T. Y., Ho, C. S. H., & Tang, J. (2017). Defective number sense or impaired access? Differential impairments in different subgroups of children with mathematics difficulties. *Journal of Learning Disabilities*, 50(1), 49–61. <https://doi.org/10.1177/0022219415588851>
- Workman, J., & Heyder, A. (2020). Gender achievement gaps: the role of social costs to trying hard in high school. *Social Psychology of Education*, 23(6), 1407–1427. <https://doi.org/10.1007/s11218-020-09588-6>
- World Health Organization. (2016). *International Statistical Classification of Diseases and Related Health Problems* (10th ed.). <https://icd.who.int/browse10/2016/en>
- Wray, E., Sharma, U., & Subban, P. (2022). Factors influencing teacher self-efficacy for inclusive education: A systematic literature review. *Teaching and Teacher Education*, 117, 103800. <https://doi.org/10.1016/j.tate.2022.103800>
- Yada, A., Leskinen, M., Savolainen, H., & Schwab, S. (2022). Meta-analysis of the relationship between teachers' self-efficacy and attitudes toward inclusive education. *Teaching and Teacher Education*, 109, 103521. <https://doi.org/10.1016/j.tate.2021.103521>

- Yin, L., Joshi, R. M., & Yan, H. (2019). Knowledge about dyslexia among early literacy teachers in China. *Dyslexia*, 1–19. <https://doi.org/10.1002/dys.1635>
- Ziesig Deng, C. (2016). *Developmental trends in peer acceptance among struggling mathematics learners*. Doktorarbeit.
- Zulliger, S. & Tanner, S. (2013). Der Begriff Heterogenität in empirischen Studien. *Schweizerische Zeitschrift für Bildungswissenschaften*, 35, 37–52. <https://doi.org/10.24452/sjer.35.1.4900>

5 Studien

5.1 Studie 1: Die Erfassung von Wissen über verschiedene Diversitätsbereiche im Kontext einer inklusionsorientierten Lehrkräftebildung

Anmerkung: Dies ist die Version des Erstautors einer Arbeit, die in der Zeitschrift *Psychologie in Erziehung und Unterricht* veröffentlicht wurde. Dieser Artikel entspricht nicht genau der endgültigen Version, die in der Zeitschrift veröffentlicht wurde, er ist keine Kopie des ursprünglich veröffentlichten Artikels und ist nicht zum Zitieren geeignet.

Tometten, L., Heyder, A., & Steinmayr, R. (2022). Empirische Arbeit: Die Erfassung von Wissen über verschiedene Diversitätsbereiche im Kontext einer inklusionsorientierten Lehrkräftebildung. *Psychologie in Erziehung und Unterricht*, 69(4), 292–304. <https://doi.org/10.2378/peu2022.art23d>

Das diesem Artikel zugrunde liegende Vorhaben wurde im Rahmen der gemeinsamen „Qualitätsoffensive Lehrerbildung“ von Bund und Ländern mit Mitteln des Bundesministeriums für Bildung und Forschung unter dem Förderkennzeichen 01JA1630 gefördert. Die Verantwortung für den Inhalt dieser Veröffentlichung liegt bei den Autorinnen.

Zusammenfassung

Für das Wissen über Diversität, einer Facette pädagogisch-psychologischen Wissens von Lehrkräften, liegen bislang nur wenige deutschsprachige Messinstrumente vor. Deshalb wurde ein Test zum deklarativen Faktenwissen in einer Lang- (DiWi) und Kurzform (DiWi-K) mit den Subskalen Förderschwerpunkte emotionale und soziale Entwicklung und Lernen, Teilleistungsstörungen, Geschlecht und kognitive Hochbegabung entwickelt und in drei unabhängigen Lehramtsstudierendenstichproben evaluiert. Studie 1 bestätigte die 5-faktorielle Struktur des DiWi. Das objektive Wissen korrelierte mit der Abiturnote, dem selbsteingeschätzten Wissen und den behandelten Diversitätsbereichen im Studium. Studierende des Lehramts für sonderpädagogische Förderung wussten mehr als Studierende anderer Lehrämter. Studie 2 bestätigte diese Ergebnisse für den DiWi-K und ergänzte sie um die konvergenten Konstrukte Vorerfahrungen, Einstellungen und Selbstwirksamkeitserwartungen bezüglich schulischer Inklusion. Studie 3 zeigte einen positiven Zusammenhang zum pädagogischen Unterrichtswissen. Diese erwartungskonformen Ergebnisse weisen auf die Eignung des Tests zur Erfassung des Diversitätswissens hin. Sie werden hinsichtlich ihrer Limitationen und Implikationen diskutiert.

Schlüsselbegriffe: Pädagogisch-psychologisches Wissen von Lehrkräften, Diversität, sonderpädagogischer Förderbedarf, Geschlecht, kognitive Hochbegabung

The assessment of knowledge about diversity in the context of an inclusion-oriented teacher training

Summary

There are so far just a few German instruments available measuring knowledge about diversity, one facet of teachers' pedagogical-psychological knowledge. Therefore, a test in a long (DiWi) and short form (DiWi-K) was developed and evaluated in three independent samples of teacher students. It measures declarative factual knowledge with the subscales special educational needs in emotional and social development and learning, partial achievement disorders, gender and cognitive giftedness. In Study 1, the 5-factorial structure of the DiWi was supported. Students' objective knowledge correlated with their GPA, self-assessed knowledge, and diversity aspects covered in their university courses. Students for special needs education had higher knowledge than students of other teaching professions. These results were confirmed in Study 2 with the DiWi-K and expanded by the convergent constructs experiences, attitudes, and self-efficacy in the context of inclusive education. In Study 3, a positive correlation with pedagogical teaching knowledge was found. The results of the studies indicate the test versions' ability to capture knowledge about diversity. The studies are discussed in terms of their limitations and implications.

Keywords: Teachers' pedagogical-psychological knowledge, diversity, special educational needs, gender differences, cognitive giftedness

5.2 Studie 2: Teachers' knowledge about diversity: Correlations with attitudes, self-efficacy, and instructions

Note: This is the first author's version of a work that was published on *OSF Preprints*. This article does not exactly replicate the final version, it is not a copy of the original published article, and it is not suitable for citation.

Tometten, L., Heyder, A., & Steinmayr, R. (2023, May 4). Teachers' knowledge about diversity: Correlations with attitudes, self-efficacy, and instructions. *OSF Preprints*. <https://doi.org/10.31219/osf.io/qtfn7>

The project this article was funded by the German Federal Ministry of Education and Research (BMBF) within the joint program "Qualitätsoffensive Lehrerbildung" of the Federal Government and the Länder under the funding code 01JA1630. The authors are responsible for the content of this publication.

Abstract

Background: Learners differ in many characteristics. In order to meet an important goal of inclusive education, the appreciation and academic support of all students, concepts are needed to meet this diversity. Knowledge in general is an important prerequisite for teachers in the implementation of inclusion and this seems especially applicable to knowledge about learners' diversity.

Purpose: Primary goal of this study is to explore in-service teachers' knowledge about diversity operationalised by five different diversity dimensions, and its relation to other constructs relevant for teaching. For this purpose, we select the variables learning opportunities, attitudes towards inclusion, self-efficacy for teaching in inclusive education and instructional practices. In a second step, misconceptions, i.e. what teachers mistakenly believe they know, are evaluated as well.

Method: We report on 139 in-service teachers' knowledge and misconceptions about diversity measured by an objective, standardised test with the subscales special educational needs due to (1) emotional and behavioural and (2) learning difficulties, (3) partial achievement disorders, (4) gender, and (5) intellectual giftedness. The teachers answered all further items in self-report within a questionnaire in a correlative design.

Findings: Teachers answered 51% of the items correctly (knowledge) and 12% incorrectly (misconceptions), with more knowledge and misconceptions than student comparison samples. Special education teachers knew more than mainstream education teachers. The higher teachers' knowledge, the more positive their attitudes towards inclusion were. Knowledge was positively related to teaching self-efficacy and mastery-oriented instructional practices.

Conclusion: The results highlight the importance of knowledge about diversity as part of teachers' professional competence in inclusive education. The correlation to mastery-

oriented practices is particularly noteworthy, as the promotion of learning processes and adaptive teaching practices are important competences in dealing with diversity.

Keywords: Teachers' pedagogical-psychological knowledge, diversity, special educational needs, gender, intellectual giftedness

1. Introduction

An important goal defined by the United Nations which is to be achieved by 2030 is that children and youth acquire knowledge and skills to participate equally in society regardless of their individual background (United Nations General Assembly, 2015). In doing so, they build on an understanding of inclusion that goes beyond the placement of students with special educational needs (SEN) in mainstream schools and focuses on the appreciation and academic support of all students (Grosche, 2015). Besides the benefits that inclusive education has to offer, it may create challenges for those who put it into practice. Teaching increasingly diverse classes is a job-specific stressor for teachers which they should be prepared for (Mintz et al., 2020; Rowan et al., 2021).

To convey knowledge as a basis for teaching practice is an evident approach in teacher qualification. Knowledge about dimensions of learners' diversity is an essential component of teachers' professional competence (Baumert & Kunter, 2013; Voss et al., 2015). It is assigned to the subject-independent pedagogical-psychological knowledge (PPK, also called general pedagogical knowledge) which has already been shown to be highly relevant for teachers' job success: In studies with pre- and in-service teachers, different facets of PPK were negatively related to indicators of burnout on the one hand (Klusmann et al., 2012; Lauermann & König, 2016), and positively related to teacher- and student-reported dimensions of instructional quality on the other hand (Depaepe & König, 2018; König & Pflanzl, 2016; Voss et al., 2011). Inclusion-relevant knowledge in special has been less explored so far but studies already provided first insights. Correlations of perceived knowledge about available support services with instructional quality (Kuyini & Desai, 2007) and of knowledge about different SEN with students' social participation and academic achievement (Tometten et al., 2021) have been demonstrated. However, not only what teachers know, but also what they mistakenly *think* they know, that is, their misconceptions

(Sciutto et al., 2000), matter in everyday school life: Misconceptions can be of particular importance in practice because they may prevent information seeking and instead lead to incorrect recommendations or interventions (Sciutto et al., 2000). Furthermore, research in the field of intellectual giftedness has shown that misconceptions correlate negatively with teachers' attitudes towards supporting gifted students (Heyder et al., 2018).

Summing up, empirical evidence on teachers' knowledge and misconceptions about diversity remains scarce: What do teachers actually know and what do they mistakenly think they know about different dimensions of student diversity? How are both related to other teaching-relevant variables? This study explores in-service teachers' declarative knowledge and misconceptions about selected dimensions of diversity (i.e., emotional/behavioural difficulties [EBD], learning difficulties [LD], partial achievement disorder [PAD], gender, and giftedness) measured by an objective, standardised and validated German test (Tometten et al., 2020) and their correlates. In doing so, we join previous research suggesting that basic demographic facts and definitions of different dimensions of student diversity as well as pedagogical approaches to dealing with these students are an important knowledge base for teachers (for an overview, see Rowan et al., 2021). Thereby, we would like to pave the way to better prepare teachers to teach inclusively.

1.1. Teachers' knowledge and misconceptions about diversity

Knowledge is an important part of teachers' professional competence (Baumert & Kunter, 2013; Shulman, 1987) and can be captured through knowledge tests that clearly define correct and incorrect answers (e.g., Voss et al., 2011). In contrast to the construct of beliefs, it must be internally consistent and meet a group or research consensus (Baumert & Kunter, 2013; Voss et al., 2011; for a more detailed discussion of how the two constructs are differentiated, see Pajares, 1992). Knowledge of relevant student characteristics is an important basis for being a good teacher. Among others it is necessary in order to provide

flexible and adapted teaching depending on the individual prerequisites and needs of the students (Shulman, 1987).

Knowledge about different diversity dimensions (e.g., giftedness, Heyder et al., 2018; autism spectrum disorder, Sanz-Cervera et al., 2017; dyslexia, Yin et al., 2019; attention deficit hyperactivity disorder [ADHD], Sciutto et al., 2000; SEN EBD, SEN LD, PAD, gender, giftedness, Tometten et al., 2021, 2022) has been studied in pre- and in-service teachers. The proportion of correct answers given in the respective knowledge tests thereby ranged between 15% and 74%, indicating substantial variation between the tests. Teachers' misconceptions have been less studied so far. In the superordinate area of PPK, to our knowledge, there are tests that contain a subscale on diversity but do not allow or report any insights into teachers' misconceptions (e.g., König & Blömeke, 2009; Voss et al., 2011). There are also studies in which the misconceptions were not analysed, although it was possible (Sousa et al., 2017; Tometten et al., 2021). Concerning specific dimensions of diversity, the proportion of incorrect answers given in the respective knowledge tests ranged between 7% and 35% (Heyder et al., 2018; Sanz-Cervera et al., 2017; Sciutto et al., 2016; Tometten et al., 2022; Yin et al., 2019). Even if these numbers show us that knowledge tended to be higher than misconceptions across the different tests, they also suggest that an expansion of evidence-based knowledge and a reduction of common misconceptions for practice should be strived for.

1.2. Correlates of teachers' knowledge and misconceptions about diversity

Theoretical models (Baumert & Kunter, 2013; Voss et al., 2015) as well as empirical research (e.g., Ghanizadeh et al., 2006; Yin et al., 2019) proposes correlations between objectively measured PPK such as knowledge about diversity and variables on teacher level such as learning opportunities, attitudes, and self-efficacy. As far as we know, there is no basic theoretical model that explicitly addresses teachers' misconceptions. But even if they

have been less researched so far, there are few empirical hints that they are also related to teaching aspects and teachers' characteristics (Heyder et al., 2018; Sanz-Cervera et al., 2017). In the following, we outline studies on both teacher's knowledge and misconception.

1.2.1. Learning opportunities

According to the COACTIV model of teachers' professional competence (Baumert & Kunter, 2013), theoretical and practical learning opportunities have a positive impact on teachers' professional knowledge. For instance, the dimensions of diversity covered in pre-service teachers' university courses correlated positively with their knowledge about diversity (Tometten et al., 2022). Likewise, the field (special vs. mainstream education) and duration of study were both associated with higher knowledge of different diversity dimensions (Sanz-Cervera et al., 2017; Sousa et al., 2017; Tometten et al., 2022) and PPK (for inclusive education, König et al., 2017; in general, Wildbrett et al., 2022). Furthermore, the more training on ADHD (Schmiedeler, 2013), autism spectrum disorder (Sanz-Cervera et al., 2017), and dyslexia (Yin et al., 2019) (student) teachers attended, the more knowledge they had about the specific topic. Job-related and general contact to affected children correlated with teachers' knowledge about ADHD (Sciutto et al., 2000) and dyslexia (Yin et al., 2019). Overall, the body of research on objectively measured knowledge about diversity supports the assumption that it increases with formal learning opportunities in this field.

Only one study tested an association between learning opportunities and misconceptions (Sanz-Cervera et al., 2017). Special education students had fewer misconceptions than students trained to be teachers at mainstream schools, but special training to teach these children with autism spectrum disorder was not associated with fewer misconceptions. Surprisingly, fourth-year students held more misconceptions than first-year students about autism spectrum disorder. Summarising, we assume a positive relationship of learning opportunities with knowledge manifesting also in higher knowledge of in- vs.

preservice teachers, whereas current research does not allow clear conclusions on the relationship with misconceptions.

1.2.2. Attitudes and self-efficacy

Attitudes are not objective facts but personal, evaluative beliefs (de Boer et al., 2011) that are related to teachers' objective knowledge about various inclusion-specific dimensions: The higher pre-service teachers' overall knowledge about different diversity dimensions, the more positive their attitudes towards inclusion were (Tometten et al., 2022). The same results were found for in-service teachers' PPK for inclusive education (König et al., 2017) and knowledge about autism spectrum disorder (Lu et al., 2020), and ADHD (Ghanizadeh et al., 2006) and their attitudes towards children with SEN, autism spectrum disorder, and ADHD in the classroom. In contrast, teachers' attitudes towards supporting gifted students correlated with self-reported but not objectively measured knowledge about giftedness (Heyder et al., 2018).

To our knowledge, there is only one study that analysed the relationship between misconceptions and attitudes. Here, a negative correlation between misconceptions about giftedness and attitudes towards supporting gifted students was found (Heyder et al., 2018).

Self-efficacy (= optimistic competence expectations; Schwarzer & Jerusalem, 1999) is also conceptually assigned to beliefs and thus based more on personal judgments and less on objective facts compared to knowledge (Pajares, 1992). Despite the theoretical assumption that knowledge and different types of beliefs (including self-efficacy beliefs) are both important aspects of teachers' professional competence (Baumert & Kunter, 2013), mixed findings on the relationship between the former and self-efficacy beliefs emerged in empirical studies. For example, in-service teachers' PPK was positively associated with their teaching self-efficacy (Laueremann & König, 2016), but pre-service teachers' PPK and their self-efficacy regarding instructional strategy, classroom management, and student

engagement were not (Depaepe & König, 2018). Concerning inclusion-specific knowledge, in-service teachers' knowledge about ADHD was positively correlated with self-efficacy in teaching these students (Sciutto et al., 2000), and pre-service teachers' overall knowledge about diversity was associated with their self-efficacy in the planning of inclusive lessons (Tometten et al., 2022).

To our knowledge, there are no studies measuring the relationship between teachers' misconceptions and their self-efficacy.

In conclusion, the research suggests rather positive correlations of knowledge with attitudes and self-efficacy in different samples using different measures. Again, the relation between both variables with misconceptions is not clear yet.

1.2.3. Instructional practices

According to theoretical models of teachers' professional competence (e.g., Baumert & Kunter, 2013), teachers' knowledge manifests in their instructional practices in class. Also empirically, positive correlations between teachers' PPK and student- and self-rated practices have already been shown (e.g., cognitive activation, Voss et al., 2011; classroom management, König & Pflanzl, 2016; management of student heterogeneity, Depaepe & König, 2018) suggesting potential benefits of high knowledge about diversity also for the practical implementation of inclusive education. Moreover, perceived knowledge about roles and partnerships in inclusion processes correlated positively with observed adaptive instructions recommended for inclusive classrooms (Kuyini & Desai, 2007).

In order to extend research findings on instructional practices in inclusive education, we suggest to study mastery- and performance-oriented instructional practices (e.g., Retelsdorf et al., 2010). The first stands for the promotion of learning processes and the adaptation of tasks to the individual needs of students and should be aimed at in education in schools, whereas the latter focuses on highlighting good results: While mastery-oriented

instructional practices are associated with higher student subject interest (Schiefele, 2017), performance-oriented instructional practices are rather related to lower student academic achievement (Anderman et al., 2001). Since mastery-orientation includes adaptive instructions to meet the individual needs of all students and support their individual development (cf. Retelsdorf et al., 2010), such practices seem particularly valuable and desirable in the context of inclusive education. We thus expect a positive relation between teachers' knowledge about diversity and their use of mastery-oriented instructional practices focussing on individual student needs, and a negative correlation with performance-oriented instructional practices focussing on good results as a key construct.

To our knowledge, there are no studies measuring the relationship between teachers' misconceptions and their instructional practices. We therefore do not make any assumptions about misconceptions.

1.3. The present study

The purpose of this study was to explore in-service teachers' knowledge and misconceptions about diversity and their relations with other variables that are theoretically relevant for teaching (cf. Baumert & Kunter, 2013). We see both as a basis for future classroom-based research and the adaptation of teacher training to the requirements of inclusive education. Previous studies have only captured perceived knowledge (e.g., Mintz et al., 2020) or used highly diverse tests that were either very specific (e.g., Sanz-Cervera et al., 2017; Sciutto et al., 2016) or very global (e.g., Voss et al., 2011). Moreover, many studies focused on pre-service teachers (e.g., Tometten et al., 2022; Depaepe & König, 2018; Sanz-Cervera et al., 2017), thus insights into in-service teachers' knowledge are limited. Also, only a few studies on misconceptions are available so far (e.g., Heyder et al., 2018; Schmiedeler, 2013) and the empirical basis should be considered fragmentary. To address these research gaps, we used a test that captures objectively measured knowledge of in-service teachers with

the five subscales of SEN EBD, SEN LD, PAD, gender, and giftedness, thus covering important dimensions of learners' diversity (DiWi; Tometten et al., 2022; available in open access via <https://doi.org/10.23668/psycharchives.5136>). Consistent with previous tests, it allows a differentiation between knowledge and misconceptions (Sanz-Cervera et al., 2017; Sciutto et al., 2000; Yin et al., 2019).

We ground our hypotheses on models of teachers' professional competence (Baumert & Kunter, 2013; Voss et al., 2015) and empirical study results in (student) teacher samples on different learning opportunities (König et al., 2017; Sousa et al., 2017), attitudes (Ghanizadeh et al., 2006; Lu et al., 2020), self-efficacy (Sciutto et al., 2000; Tometten et al., 2022), and instructional practices (Depaepe & König, 2018; Kuyini & Desai, 2007).

Consequently, we expected:

H₁: In-service teachers have higher knowledge about diversity in comparison to bachelor and master students.

H₂: Higher knowledge about diversity is shown by (a) in-service teachers for special needs education compared to those for mainstream education and (b) teachers who attended a special training on inclusion compared to those who did not.

H₃: In-service teachers' knowledge about diversity is positively correlated with their job-related contact with students with different diversity backgrounds, attitudes towards inclusion, self-efficacy in teaching in an inclusive school, and mastery-oriented instructional practices and negatively correlated with performance-oriented instructional practices.

Besides, we want to explore whether there are correlations with and group differences in misconceptions.

2. Method

2.1. Participants

A total of 139 in-service teachers from schools in Germany took part in the survey.

They were on average 40.41 years old ($SD = 11.29$) and worked in school for 12.52 years on average ($SD = 10.86$). There were more women (68.3%) than men (30.9%; 1 person without indication of gender). Of all participating teachers, 21.6% were qualified for special needs education and 31.9% had participated in at least one further training on inclusive education. The participants taught in primary schools (19.4%), secondary schools (28.8%), grammar schools (25.9%), vocational colleges (11.5%), segregated special schools (12.2%) and others (2.2%). In Germany, all students normally go to primary school first and then enter secondary schools for vocational training or grammar schools for being prepared to study at a university afterwards. Students with special educational needs can be prepared for various qualifications (mostly vocational training) in mainstream schools (called “inclusive education” in Germany) or separately from students without these needs in segregated special schools. The distribution of teachers across gender and school types was close to official statistics (Information and Technology North Rhine-Westphalia, 2022; Statista, 2022).

2.2. Measures

2.2.1. Knowledge and misconceptions about diversity

Teacher knowledge and misconceptions were assessed with the DiWi (Tometten et al., 2022) consisting of 142 items ($\alpha = .95$) assessing basic declarative knowledge about SEN EBD (35 items; $\alpha = .86$), SEN LD (28 items; $\alpha = .86$), PAD (dyscalculia, dyslexia; 27 items; $\alpha = .82$), gender (26 items, two of them [GES6 & GES8] were slightly reworded for this survey; $\alpha = .82$), and giftedness (26 items; $\alpha = .83$). These were clearly true or false statements and teachers were instructed to indicate whether each statement was true, false, or whether they did not know. Correct answers were coded with 1 and incorrect/don't-know answers with 0 to assess knowledge. The opposite coding procedure was used to assess misconceptions ($1 = incorrect\ answer, 0 = correct/don't-know\ answer$). Reliability and factorial validity have been shown before (Tometten et al., 2022). The complete test is

available in open access via <https://doi.org/10.23668/psycharchives.5136>. Exemplary items are shown in Table 1.

2.2.2. *Job-related contact with diversity*

Teachers answered the question 'How much job-related contact do you have with pupils with SEN EBD/SEN LD/PAD/giftedness?'. The mean value of the answers ($1 = none$, $5 = continuous$) across diversity dimensions was used as an indicator of contact with diversity.

2.2.3. *Attitudes towards inclusion*

Teachers' attitudes towards inclusive education were assessed by 11 items from the EZI (Einstellungen zu Integration [Attitudes towards integration]; Kunz et al., 2010), slightly adjusted for inclusive classes ($\alpha = .83$; e.g., 'The quality of regular classroom instruction is better when children with special educational needs are included'). The participants indicated their agreement on a 5-point Likert scale ($1 = not\ at\ all$, $5 = fully$).

2.2.4. *Self-efficacy for teaching in inclusive education*

Self-efficacy was assessed using four items of the German scale of general self-efficacy (Skala zur Allgemeinen Selbstwirksamkeitserwartung [SWE]; Schwarzer & Jerusalem, 1999). The items ($\alpha = .90$) were slightly adapted to teaching in an inclusive school (e.g., 'Whatever happens in my work as a teacher in an inclusive school, I will be able to cope'). The participants indicated their agreement on a 5-point Likert scale ($1 = not\ at\ all$, $5 = fully$).

2.2.5. *Instructional practices*

To measure the type of instruction in class, two subscales from the German version of the patterns of adaptive learning scales (PALS; Retelsdorf et al., 2010) were used. Six items assessed mastery-oriented instructional practices ($\alpha = .84$; e.g., 'In my class, the individual development of my students is important to me') and five items assessed performance-

oriented instructional practices ($\alpha = .79$; e.g., 'In my class, I present the work of the best student as an example'). The answers were given on a 5-point Likert scale ($1 = \textit{not applicable at all}$, $5 = \textit{fully applicable}$).

2.2.6. Professional background

In addition to demographic information, teachers reported their teaching profession (= being trained to teach in special needs vs. mainstream education), and attendance of a special training on inclusion.

2.3. Procedure

The survey was available in an online (46%) and a paper-pencil version (54%). Participants were recruited through school administrations, teacher training institutions, personal inquiries and social media. Participation lasted approximately 30–45 min, was voluntary, took place on participants' private time, and was not paid. Teachers could optionally participate in a voucher lottery.

2.4. Statistical analysis

The analyses were performed using the IBM SPSS Statistics 26. We calculated bivariate correlations, univariate analyses of variance (ANOVA) in case of two-group-comparisons with one dependent variable, and multivariate analyses of variance (MANOVA) in case of two-group-comparisons with more than one dependent variable. Partial Eta² (η^2_{part}) was used to estimate effect sizes, whereby the effects are estimated to be small from .01, medium from .06 and large from .14 onwards (Cohen, 1988).

In order to address our first hypothesis on knowledge differences between teachers and students, we compared the present in-service teacher data with available data of bachelor and master students (Tometten et al., 2022). Comparison Sample 1 consisted of 395 bachelor students (284 female, 98 male, 13 persons with no indication of gender; age: $M = 21.45$, $SD = 3.10$), of whom 21.5% studied special education. These bachelor students also filled in the

DiWi. Comparison Sample 2 consisted of 181 master students (126 female, 39 male, 16 persons without indication of gender; age: $M = 25.56$, $SD = 3.36$). Here, diversity knowledge and misconceptions were assessed with the DiWi-K, a validated short version consisting of 36 items. Since these items are also part of the long version used in the present in-service teacher sample, we formed the short version for the teachers as well and compared teachers' and master students' knowledge as well as misconceptions, all measured by the DiWi-K (for details on the samples, see Tometten et al., 2022).

3. Results

3.1. Descriptive statistics

The mean values and standard deviations of all variables are shown in Table 2, subscale-specific values are presented in the text (for subscale-specific statistical analyses see the supplemental online material). Teachers answered 51% of the items of the knowledge test correctly, with knowledge being above 50% in the areas of SEN EBD ($M = .60$, $SD = .18$), SEN LD ($M = .56$, $SD = .21$) and PAD ($M = .57$, $SD = .20$) and below 50% in the areas of gender ($M = .41$, $SD = .20$) and giftedness ($M = .39$, $SD = .19$). There were 12% answers indicating misconceptions, with slightly fewer in the areas of SEN EBD ($M = .09$, $SD = .07$), SEN LD ($M = .11$, $SD = .09$), and PAD ($M = .11$, $SD = .09$) than in the areas of gender ($M = .13$, $SD = .09$) and giftedness ($M = .15$, $SD = .11$). Teachers were uncertain about 37% of the items. Exemplary items of the DiWi with a low, medium and high proportion of knowledge including the indication of the proportion of misconceptions and uncertainty are shown in Table 1.

Teachers reported a moderately high level of job-related contact with children from different diversity backgrounds (see Table 2). Compared to SEN EBD ($M = 3.34$, $SD = 1.46$), SEN LD ($M = 3.43$, $SD = 1.37$) and PAD ($M = 3.48$, $SD = 1.28$), teachers reported relatively low contact with children with giftedness ($M = 2.06$, $SD = 0.97$). Teachers' attitudes were in

the middle range of the scale with a minimal tendency towards inclusive schools compared to segregated special schools. Self-efficacy was slightly higher at the upper half of the scale. The teachers assessed their teaching instructions as more mastery-oriented and less performance-oriented.

3.2. Group differences (H_1 & H_2)

For testing H_1 , we compared the results from this study with two earlier collected student data sets. The teachers had significantly higher knowledge ($M = .51$, $SD = .16$) compared to the bachelor students ($M = .29$, $SD = .17$; $F[532] = 181.01$, $p < .001$, $\eta^2_{\text{part}} = .25$) supporting our first hypothesis. Moreover, they also had slightly more misconceptions ($M = .12$, $SD = .06$) than the bachelor students ($M = .08$, $SD = .08$; $F[532] = 26.45$, $p < .001$, $\eta^2_{\text{part}} = .05$). The same pattern was found for the comparison between in-service teachers and master students, also supporting our first hypothesis: Teachers had higher knowledge ($M = .53$, $SD = .17$) compared to the master students ($M = .41$, $SD = .21$; $F[318] = 30.13$, $p < .001$, $\eta^2_{\text{part}} = .09$). They also had more misconceptions ($M = .15$, $SD = .08$) than the master students ($M = .10$, $SD = .08$; $F[318] = 28.10$, $p < .001$, $\eta^2_{\text{part}} = .08$).

As hypothesised (H_{2a}), group comparisons within the teacher sample revealed higher knowledge for teachers for special needs education ($M = .62$, $SD = .13$) compared to those for mainstream education ($M = .48$, $SD = .15$; $F[137] = 18.74$, $p < .001$, $\eta^2_{\text{part}} = .12$). This difference also remained after controlling for teachers' age, gender, teaching experience in years, and job-related contact. Teachers for special needs education did not differ in their overall misconceptions ($M = .11$, $SD = .04$) from teachers for mainstream education ($M = .12$, $SD = .07$; $F[137] = 0.16$, $p = .69$, $\eta^2_{\text{part}} = .001$).

Unlike what we expected (H_{2b}), there were no differences in knowledge between teachers who participated in a further training or certificate course focusing on inclusion ($M = .54$, $SD = .13$) and those who did not ($M = .50$, $SD = .17$; $F[136] = 1.43$, $p = .23$, $\eta^2_{\text{part}} = .01$).

There were also no differences in misconceptions of teachers with ($M = .12$, $SD = .05$) and without further training ($M = .12$, $SD = .07$; $F[136] = 0.03$, $p = .87$, $\eta^2_{\text{part}} < .001$).

3.3. Bivariate correlations (H_3)

As expected (H_3), overall knowledge correlated significantly positive with job-related contact ($p < .001$), attitudes towards inclusion ($p = .003$), self-efficacy in teaching in an inclusive school ($p < .001$), and mastery-oriented instructional practices ($p < .001$; see Table 2). The negative correlation between knowledge and performance-oriented instructional practices barely missed significance ($p = .07$). There was only one positive correlation of misconceptions with overall job-related contact ($p = .02$).

4. Discussion

The aim of this study was to explore teachers' knowledge and misconceptions about diversity in the context of inclusive education and their relationship to other teaching-relevant variables (here learning opportunities, attitudes, self-efficacy, and instructional practices). These findings serve to deepen the understanding of teachers' knowledge and misconceptions in order to better prepare pre- and in-service teachers for teaching diverse groups of students. An objective knowledge test with five subscales on different diversity dimensions was used (see Tometten et al., 2022) in a sample of in-service teachers to capture knowledge more extensively and practice-related than in previous studies (e.g., König et al., 2017; Sanz-Cervera et al., 2017).

In summary, the analyses revealed higher knowledge among in-service teachers compared to bachelor and master students. Knowledge was positively related to job-related contact, attitudes, self-efficacy, and mastery-oriented instructional practices, whereas the negative correlation with performance-oriented practices just missed significance. Special education teachers knew more than those for mainstream education. However, teachers who had attended at least one vs. no training on inclusion did not differ in their knowledge.

4.1. Interpretation of the results

Teachers in our sample answered approximately half of the items correctly and only one-tenth incorrectly. They tended to have descriptively higher knowledge and fewer misconceptions about SEN EBD, SEN LD, and PAD than about giftedness and gender, although both giftedness and gender are included in German teacher education standards (Conference of the Ministers of Education and Cultural Affairs, 2014). We can only speculate whether these topics were underrepresented at universities at the time when the teachers in our sample were studying, as it has been reported, for example, by primary school teachers on knowledge about giftedness (Heller et al., 2005). It might also be that due to the general discussion on inclusive education, teachers consider gender and giftedness to be less relevant as more attention is paid to special educational needs, for example, in research, media, or teacher training (Grosche, 2015; Kurniawati et al., 2014). The call for promoting knowledge and skills based on evidence-based knowledge about gender and giftedness (e.g., Heyder et al., 2018; Kollmayer et al., 2016) is thus corroborated by our findings.

4.1.1. Teachers' knowledge about diversity (H₁-H₃)

The comparison between teachers and students revealed that in-service teachers had higher than students in the bachelor and master samples (H₁), underscoring the importance of both theoretical and practical learning opportunities for gaining professional knowledge (Baumert & Kunter, 2013; Voss et al., 2015). In this vein, teachers of special needs education had higher knowledge than teachers of mainstream education (H_{2a}) as it has been shown in pre-service teacher samples before (Tometten et al., 2022). Teachers who reported to have attended in-service training on inclusion did not have higher knowledge in the subscales and the overall test (H_{2b}). Given that teacher training is likely to vary in content (e.g., attitude/knowledge/skills or a combinations of these topics) and learning activities (e.g., school visits, guest lectures; Kurniawati et al., 2014), this result appears in line with earlier,

partly inconsistent findings of previous studies (Schmiedeler, 2013; Yin et al., 2019) and can further be explored in future studies.

Teachers' knowledge correlated positively (H_3) with their job-related contact with students with different diversity backgrounds, attitudes towards inclusion, self-efficacy in teaching in an inclusive school, and mastery-oriented instructional practices as has been shown in previous studies with slightly different knowledge tests or samples (e.g., Depaepe & König, 2018; Kuyini & Desai, 2007; Lu et al., 2020; Sciutto et al., 2000; Tometten et al., 2022). The correlations between instructional practices and knowledge about diversity, in particular, are noteworthy, as they have not been studied in this way before. Mastery-oriented instructional practices and knowledge seem relevant for providing the best individual support to each student in the context of inclusive education. However, different explanations for the positive correlation are plausible. More knowledge might enable mastery-oriented practices, or the desire for individual support might lead to more knowledge acquisition, or teachers for special needs education learn more about both than others during their studies. A negative correlation between knowledge and performance-oriented instructional practices was found in the expected direction. However, the effect was small and not significant in this sample.

4.1.2. Teachers' misconceptions about diversity

In sum, analyses revealed few correlations and group differences. Our sample was thus in the range of previous studies (e.g., Heyder et al., 2018; Sciutto et al., 2000), with 12% answers indicating misconceptions overall and 9%–15% per subscale.

Counterintuitively, misconceptions appeared to be positively correlated with knowledge which can be explained by the idea that teachers with high knowledge may also be particularly self-confident about this topic and tend to guess more frequently (cf. Sherman, 1976). Therefore, other formats including the assessment of response confidence (Hasan et al., 1999) could be tested in future studies in addition to the repeatedly used true/false/don't-

know answer format (see Heyder et al., 2018; Sanz-Cervera et al., 2017; Sciutto et al., 2000; Yin et al., 2019). Another explanation might be that high knowledge teachers are more open to further information and actively seek for it, as suggested, for example, by the Openness-Fluid-Crystallized-Intelligence model (Ziegler et al., 2012) on the positive relationship between openness and knowledge. Since popular media are an important source of information for teachers (e.g., Heyder et al., 2018), actively information seeking teachers might also gain information classified as misconceptions from such sources. This could also explain why in-service teachers had more misconceptions than bachelor and master students who might rely less on popular media and more on teacher education.

Misconceptions also correlated positively with job-related contact with students with different diversity backgrounds. This correlation could also be explained by the idea that teachers with high contact feel more confident in their answers and are thus more likely to guess. Another explanation could be a bias in the reporting of job-related contact. It might be that teachers only report contact when the corresponding students confirm their subjective assumptions or image of students with this diagnosis or characteristics, and report no or less contact if both are not aligned. In most cases, these assumptions should contain both knowledge and misconceptions, resulting in positive correlations between misconceptions and job contact, as well as knowledge and job contact. Therefore, the evaluation of misconceptions can also generate new insights as well as research questions concerning teachers' knowledge.

4.2. Limitations

Self-evidently, this is a cross-sectional design that does not allow statements on causality. Various mechanisms that underlie our correlational findings are conceivable. Accordingly, directed associations remain an open question for future longitudinal research, for example if knowledge about diversity predicts instructional practices. It is also worth

clarifying if the results can be replicated in a larger sample.

Moreover, we cannot exclude that the teachers who participated online searched for the correctness of the statements in parallel. However, since there was a total of 142 items that would require a long time to research, this can be considered unlikely. That is supported by additional analyses showing that the mean values of teachers who participated online vs. paper-pencil did not differ significantly from each other. One should also keep in mind that teachers actively chose to participate in this study and completed the questionnaires in their free time, whereas the comparison student samples were acquired in university lectures, setting the threshold for participation lower for the latter.

Furthermore, all variables except for teachers' knowledge and misconceptions were recorded by teacher self-reports. To understand what actually happens in the classroom, teaching-related variables could be assessed in future research through class-room observations (cf. Kuyini & Desai, 2007).

4.3. Implications for research and practice

The results indicate an association between objectively measured knowledge about diversity and different other variables related to teaching in inclusive education as we hypothesised based on the COACTIV model of teachers' professional competence (Baumert & Kunter, 2013). They thus highlight the importance of evidence-based knowledge also for inclusive education to be promoted in teacher training. Misconceptions, on the other hand, should be reduced in order to avoid, for example, wrong recommendations and interventions (Sciutto et al., 2000).

The implementation of German standards for teacher training seems to vary widely (Hohenstein et al., 2014). The importance of gender was, for example, only implemented in somewhat more than half of the curricula studied. As knowledge about gender and giftedness might be needed by all teachers, we consider a critical examination of the curricula and

specific lectures for correct representation of what is currently known from research to be important. This also applies to in-service teacher training. Teachers' contact with students with different diversity backgrounds, their individual prior knowledge (when rather low), or other teaching-related factors might increase the need to convey knowledge about specific diversity dimensions. In-service training for teachers on these dimensions could complement current trainings on inclusion in general, such as those offered for example in North Rhine-Westphalia (Ministry for Schools and Education of North Rhine-Westphalia, 2014).

There is further potential to develop and expand our test. Diversity does not only refer to the five dimensions presented but encompasses more (e.g., Grosche, 2015; Rowan et al., 2021). Thus, further scales may be included, such as motivation to learn, German as a second language, socio-economic status, and cultural differences.

In a cross-national comparison study, German teachers' knowledge about ADHD was above average, whereas their misconceptions were below (Sciutto et al., 2016). Research across countries may shed light on the extent to which teachers in different countries differ in their knowledge and misconceptions about different diversity dimensions.

Especially with regard to the misconceptions, further research is needed: Are there correlations between misconceptions and other teaching-related variables? How do misconceptions perform when measured with other instruments? It is also unclear which misconceptions have high or low practical relevance. Those that are highly relevant in practice can be addressed, for example, with the help of refutation texts in which the false belief is explicitly addressed (Zengilowski et al., 2021).

4.4. Conclusion

In summary, important implications for research and practice can be derived. Our findings support the assumption that formal learning opportunities, whether theoretical or practical, such as being practical in-service teacher, having job-related contact with students

with different diversity backgrounds, having attended special training on inclusion, and being special education teacher are related to teachers' professional knowledge (Baumert & Kunter, 2013). Also, correlations with other teaching-related variables such as attitudes towards inclusion, self-efficacy in teaching in an inclusive school, and mastery-oriented instructional practices emerged, highlighting the importance of the examined knowledge content for inclusive education. Therefore, the development of diversity-specific curricula in teacher education for all types of schools seems particularly important to us. Especially the two areas of giftedness and gender might be underrepresented so far because less knowledge and more misconceptions were shown here. Our study also provides a basis for future research on capturing and managing teachers' misconceptions. This includes researching the mechanisms in the emergence of misconceptions and addressing them in teacher education so that they can be uncovered and corrected.

References

- Anderman, E. M., Eccles, J. S., Yoon, K. S., Roeser, R., Wigfield, A., & Blumenfeld, P. (2001). Learning to value mathematics and reading: Relations to mastery and performance-oriented instructional practices. *Contemporary Educational Psychology*, 26(1), 76–95. <https://doi.org/10.1006/ceps.1999.1043>
- Baumert, J., & Kunter, M. (2013). The COACTIV model of teachers' professional competence. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss, & M. Neubrand (Eds.), *Cognitive activation in the mathematics classroom and professional competence of teachers. Results from the COACTIV project* (pp. 25–48). Springer. <https://doi.org/10.1007/978-1-4614-5149-5>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Taylor and Francis.
- Conference of the Ministers of Education and Cultural Affairs. (2014, June). *Standards für die Lehrerbildung: Bildungswissenschaften [Standards for teacher education: Educational sciences]*. http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2004/2004_12_16-Standards-Lehrerbildung-Bildungswissenschaften.pdf
- de Boer, A., Pijl, S. J., & Minnaert, A. (2011). Regular primary schoolteachers' attitudes towards inclusive education: A review of the literature. *International Journal of Inclusive Education*, 15(3), 331–353. <https://doi.org/10.1080/13603110903030089>
- Depaepe, F., & König, J. (2018). General pedagogical knowledge, self-efficacy and instructional practice: Disentangling their relationship in pre-service teacher education. *Teaching and Teacher Education*, 69, 177–190. <https://doi.org/10.1016/j.tate.2017.10.003>
- Ghanizadeh, A., Bahredar, M. J., & Moeini, S. R. (2006). Knowledge and attitudes towards attention deficit hyperactivity disorder among elementary school teachers. *Patient Education and Counseling*, 63, 84–88. <https://doi.org/10.1016/j.pec.2005.09.002>
- Grosche, M. (2015). Was ist Inklusion? Ein Diskussions- und Positionsartikel zur Definition von Inklusion aus Sicht der empirischen Bildungsforschung [What is inclusion? A discussion and position paper on the definition of inclusion from the perspective of empirical educational research]. In P. Kuhl, P. Stanat, B. Lütje-Klose, C. Gresch, H. A. Pant, & M. Prenzel (Eds.), *Inklusion von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf in Schulleistungserhebungen [Inclusion of pupils with special educational needs in school achievement surveys]* (pp. 17–39). Springer VS. <https://doi.org/10.1007/978-3-658-06604-8>
- Hasan, S., Bagayoko, D., & Kelley, E. L. (1999). Misconceptions and the certainty of response index (CRI). *Physics Education*, 34(5), 294–299. <https://doi.org/10.1088/0031->

[9120/34/5/304](#)

- Heller, K. A., Reimann, R., & Senfter, A. (2005). Hochbegabung im Grundschulalter: Erkennen und Fördern [Giftedness in primary-school-age children: Identification and promotion]. LIT.
- Heyder, A., Bergold, S. & Steinmayr, R. (2018). Teachers' knowledge about intellectual giftedness: A first look at levels and correlates. *Psychology Learning & Teaching*, 17(1), 27–44. <https://doi.org/10.1177/1475725717725493>
- Hohenstein, F., Zimmermann, F., Kleickmann, T., Köller, O., & Möller, J. (2014). Sind die bildungswissenschaftlichen Standards für die Lehramtsausbildung in den Curricula der Hochschulen angekommen? [Have the education standards for teacher training programmes arrived in the university curriculum?]. *Zeitschrift für Erziehungswissenschaft*, 17(3), 497–507. <https://doi.org/10.1007/s11618-014-0563-9>
- Information and Technology North Rhine-Westphalia (2022, March). *Schulen, Klassen, Schülerinnen, Schüler und Lehrkräfte an allgemeinbildenden und beruflichen Schulen nach Schulformen [Schools, classes, students and teachers in mainstream and vocational schools by type of school]*. <https://www.it.nrw/node/1142/pdf>
- Klusmann, U., Kunter, M., Voss, T., & Baumert, J. (2012). Berufliche Beanspruchung angehender Lehrkräfte: Die Effekte von Persönlichkeit, pädagogischer Vorerfahrung und professioneller Kompetenz [Emotional exhaustion and job satisfaction of beginning teachers: The role of personality, educational experience and professional competence]. *Zeitschrift für Pädagogische Psychologie*, 26(4), 275–290. <https://doi.org/10.1024/1010-0652/a000078>
- König, J., & Blömeke, S. (2009). Pädagogisches Wissen von angehenden Lehrkräften – Erfassung und Struktur von Ergebnissen der fachübergreifenden Lehrerausbildung [Pedagogic knowledge of future teachers – Capturing and structuring the results of cross-subject teacher education]. *Zeitschrift für Erziehungswissenschaft*, 12(3), 499–527. <https://doi.org/10.1007/s11618-009-0085-z>
- König, J., Gerhard, K., Melzer, C., Rühl, A.-M., Zenner, J., & Kaspar, K. (2017). Erfassung von pädagogischem Wissen für inklusiven Unterricht bei angehenden Lehrkräften: Testkonstruktion und Validierung [Assessing General Pedagogical Knowledge for Inclusive Teaching (GPK-IT) Among Pre-Service Teachers: Test Construction and Validation]. *Unterrichtswissenschaft*, 45(4), 223–242. <https://doi.org/10.3726/pr012019.0004>
- König, J., & Pflanzl, B. (2016). Is teacher knowledge associated with performance? On the relationship between teachers' general pedagogical knowledge and instructional quality. *European Journal of Teacher Education*, 39(4), 419–436. <https://doi.org/10.1080/02619768.2016.1214128>

- Kollmayer, M., Schober, B., & Spiel, C. (2016). Gender stereotypes in education: Development, consequences, and interventions. *European Journal of Developmental Psychology, 15*(4), 361–377. <https://doi.org/10.1080/17405629.2016.1193483>
- Kunz, A., Luder, R., & Moretti, M. (2010). Die Messung von Einstellungen zur Integration (EZI) [The measurement of attitudes towards integration (ATI)]. *Empirische Sonderpädagogik, 2*(3), 83–94.
- Kurniawati, F., de Boer, A. A., Minnaert, A. E. M. G., & Mangunsong, F. (2014). Characteristics of primary teacher training programmes on inclusion: A literature focus. *Educational Research, 56*(3), 310–326. <https://doi.org/10.1080/00131881.2014.934555>
- Kuyini, A. B., & Desai, I. (2007). Principals' and teachers' attitudes and knowledge of inclusive education as predictors of effective teaching practices in Ghana. *Journal of Research in Special Educational Needs, 7*(2), 104–113. <https://doi.org/10.1111/j.1471-3802.2007.00086.x>
- Lauermann, F., & König, J. (2016). Teachers' professional competence and wellbeing: Understanding the links between general pedagogical knowledge, self-efficacy and burnout. *Learning and Instruction, 45*, 9–19. <https://doi.org/10.1016/j.learninstruc.2016.06.006>
- Lu, M., Zou, Y., Chen, X., Chen, J., He, W., & Pang, F. (2020). Knowledge, attitude and professional self-efficacy of Chinese mainstream primary school teachers regarding children with autism spectrum disorder. *Research in Autism Spectrum Disorders, 72*, 101513. <https://doi.org/10.1016/j.rasd.2020.101513>
- Ministry for Schools and Education of North Rhine-Westphalia. (2014, April). Fort- und Weiterbildung: Strukturen und Inhalte der Fort- und Weiterbildung für das Schulpersonal [In-service training: Structures and contents of in-service training for school staff]. <https://bass.schul-welt.de/pdf/14149.pdf?20221012084448>
- Mintz, J., Hick, P., Solomon, Y., Matziari, A., Ó'Murchú, F., Hall, K., Cahill, K., Curtin, C., Anders, J., & Margariti, D. (2020). The reality of reality shock for inclusion: How does teacher attitude, perceived knowledge and self-efficacy in relation to effective inclusion in the classroom change from the pre-service to novice teacher year? *Teaching and Teacher Education, 91*, 103042. <https://doi.org/10.1016/j.tate.2020.103042>
- Pajares, M. F. (1992). Teachers' beliefs and educational research: Cleaning up a messy construct. *Review of Educational Research, 62*(3), 307–332. <https://doi.org/10.3102/00346543062003307>
- Retelsdorf, J., Butler, R., Streblov, L., & Schiefele, U. (2010). Teachers' goal orientations for teaching: Associations with instructional practices, interest in teaching, and burnout. *Learning and Instruction, 20*, 30–46. <https://doi.org/10.1016/j.learninstruc.2009.01.001>
- Rowan, L., Bourke, T., L'Estrange, L., Lunn Brownlee, J., Ryan, M., Walker, S., &

- Churchward, P. (2021). How does initial teacher education research frame the challenge of preparing future teachers for student diversity in schools? A systematic review of literature. *Review of Educational Research*, 91(1), 112–158. <https://doi.org/10.3102/0034654320979171>
- Sanz-Cervera, P., Fernández-Andrés, M. I., Pastor-Cerezuela, G., & Tárraga-Mínguez, R. (2017). Pre-service teachers' knowledge, misconceptions and gaps about autism spectrum disorder. *Teacher Education and Special Education*, 40(3), 212–224. <https://doi.org/10.1177/0888406417700963>
- Schiefele, U. (2017). Classroom management and mastery-oriented instruction as mediators of the effects of teacher motivation on student motivation. *Teaching and Teacher Education*, 64, 115–126. <https://doi.org/10.1016/j.tate.2017.02.004>
- Schmiedeler, S. (2013). Wissen und Fehlannahmen von deutschen Lehrkräften über die Aufmerksamkeitsdefizit-/Hyperaktivitätsstörung (ADHS) [German teachers' knowledge and misconceptions about aAttention-Deficit-/Hyperactivity Disorder (ADHD)]. *Psychologie in Erziehung und Unterricht*, 60, 143–153. <https://doi.org/10.2378/peu2013.art12d>
- Schwarzer, R., & Jerusalem, M. (1999). Skalen zur Erfassung von Lehrer- und Schülermerkmalen. Dokumentation der psychometrischen Verfahren im Rahmen der Wissenschaftlichen Begleitung des Modellversuchs Selbstwirksame Schulen [Scales for measuring teacher and student characteristics. Documentation of the psychometric procedures within the framework of the scientific monitoring of the pilot project Self-effective Schools]. Berlin: Freie Universität Berlin.
- Sciotto, M. J., Terjesen, M. D., & Bender Frank, A. S. (2000). Teachers' knowledge and misperceptions of Attention-Deficit/Hyperactivity Disorder. *Psychology in the Schools*, 37(2), 115–122. [https://doi.org/10.1002/\(sici\)1520-6807\(200003\)37:2<115::aid-pits3>3.0.co;2-5](https://doi.org/10.1002/(sici)1520-6807(200003)37:2<115::aid-pits3>3.0.co;2-5)
- Sciotto, M. J., Terjesen, M. D., Kučerová, A., Michalová, Z., Schmiedeler, S., Antonopoulou, K., Shaker, N. Z., Lee, J.-Y., Drake, B., & Rossouw, J. (2016). Cross-national comparisons of teachers' knowledge and misconceptions of ADHD. *International Perspectives in Psychology*, 5(1), 34–50. <https://doi.org/10.1037/ipp0000045>
- Sherman, S. W. (1976, April). *Multiple choice test bias uncovered by use of an "I don't know" alternative*. [Paper presentation]. Annual Meeting of the American Educational Research Association, San Francisco.
- Shulman, L. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review*, 57(1), 1–23.
- Sousa, P., Dias, P. C., & Cadime, I. (2017). Predictors of primary school teachers' knowledge about developmental dyscalculia. *European Journal of Special Needs Education*, 32(2),

- 204–220. <https://doi.org/10.1080/08856257.2016.1216635>
- Statista. (2022, November). Anteil der weiblichen Lehrkräfte an allgemeinbildenden Schulen in Deutschland im Schuljahr 2020/2021 nach Schulart [Proportion of female teachers at schools in Germany in 2020/2021 by type of school]. <https://de.statista.com/statistik/daten/studie/1129852/umfrage/frauenanteil-unter-den-lehrkraeften-in-deutschland-nach-schulart/>
- Tometten, L., Heyder, A., & Steinmayr, R. (2021). Links between teachers' knowledge about special educational needs and students' social participation and academic achievement in mainstream classes. *Contemporary Educational Psychology*, *67*, 102022. <https://doi.org/10.1016/j.cedpsych.2021.102022>
- Tometten, L., Heyder, A., & Steinmayr, R. (2022). Die Erfassung von Wissen zu Diversität im Kontext einer inklusionsorientierten, psychologischen Lehrkräftebildung [The assessment of knowledge on diversity in the context of an inclusion-oriented, psychological teacher training]. *Psychologie in Erziehung und Unterricht*, *69*(4), 292–304. <https://doi.org/10.2378/peu2022.art23d>
- United Nations General Assembly. (2015, October). *Transforming our world: The 2030 agenda for sustainable development*. A/RES/70/1, United Nations. https://www.unfpa.org/sites/default/files/resource-pdf/Resolution_A_RES_70_1_EN.pdf
- Voss, T., Kunina-Habenicht, O., Hoehne, V., & Kunter, M. (2015). Stichwort Pädagogisches Wissen von Lehrkräften: Empirische Zugänge und Befunde [Teachers' pedagogical knowledge: Empirical approaches and findings]. *Zeitschrift für Erziehungswissenschaft*, *18*(2), 187–223. <https://doi.org/10.1007/s11618-015-0626-6>
- Voss, T., Kunter, M., & Baumert, J. (2011). Assessing teacher candidates' general pedagogical/psychological knowledge: Test construction and validation. *Journal of Educational Psychology*, *103*(4), 952–969. <https://doi.org/10.1037/a0025125>
- Wildbrett, J., Lohse-Bossenz, H., & Dörfler, T. (2022). Vignettenbasierte Messung pädagogisch-psychologischen Professionswissens bei Lehramtsstudierenden [Vignette based assessment of pedagogical psychological knowledge of teacher students]. *Psychologie in Erziehung und Unterricht*, *69*(4), 278–291. <https://doi.org/10.2378/peu2022.art20d>
- Yin, L., Joshi, R. M., & Yan, H. (2019). Knowledge about dyslexia among early literacy teachers in China. *Dyslexia*, 1–19. <https://doi.org/10.1002/dys.1635>
- Zengilowski, A., Schuetze, B. A., Nash, B. L., & Schallert, D. L. (2021). A critical review of the refutation text literature: Methodological confounds, theoretical problems, and possible solutions. *Educational Psychologist*, *56*(3), 175–195. <https://doi.org/10.1080/00461520.2020.1861948>
- Ziegler, M., Danay, E., Heene, M., Asendorpf, J., & Bühner, M. (2012). Openness, fluid

intelligence, and crystallized intelligence: Toward an integrative model. *Journal of Research in Personality*, 46(2), 173–183. <https://doi.org/10.1016/j.jrp.2012.01.002>

Table 1. Exemplary items of the DiWi with a low, medium and high proportion of knowledge.

Subscale	No.	Item (<i>correct answer</i>)	Proportion in percent (%)		
			Knowledge	Misconceptions	Uncertainty
SEN EBD	2	Children and adolescents with special educational needs with regard to EBD are at particular risk for drug abuse. (<i>true</i>)	18.0	30.2	51.8
	33	Tokens are objects that are specifically given to reinforce desired behaviors. (<i>true</i>)	51.1	2.2	46.8
	27	In most cases, ADHD is associated with a lower IQ than average. (<i>false</i>)	82.7	3.6	13.7
SEN LD	2	Learning disability is defined as a disorder in the ICD 10. (<i>false</i>)	5.8	26.8	67.4
	11	The probability of a learning impairment increases with decreasing socio-economic status. (<i>true</i>)	48.6	13.0	38.4
	24	Support goals should be positive and realistic in order to achieve a positive learning development. (<i>true</i>)	91.4	0.0	8.6
PAD	26	The likelihood of occurrence of dyscalculia also depends on the teaching instruction that is offered (e.g., specific remedial instruction in mathematics). (<i>true</i>)	14.4	42.5	43.2
	12	Frequent stuttering is a common symptom of a reading disorder. (<i>false</i>)	48.9	14.6	36.5
	23	Dyscalculia is characterized by a circumscribed and significant impairment of arithmetic skills. (<i>true</i>)	89.2	2.2	8.6

	17	On average, boys perform better when they are taught by a male than a female teacher. (<i>false</i>)	18.3	15.3	66.4
Gender	15	On average, girls rate their verbal skills higher than boys. (<i>true</i>)	46.0	8.8	45.3
	4	The percentage of girls who graduate from high school has increased over the past 30 years. (<i>true</i>)	68.6	2.2	29.2
	19	Intellectually gifted children and adolescents often have more problems compared to their non-gifted peers. (<i>false</i>)	7.2	61.2	31.7
Giftedness	26	Skipping a class is a useful support strategy even for intellectually gifted children with a low level of motivation. (<i>false</i>)	43.9	11.5	44.6
	3	To diagnose intellectual giftedness, the intelligence of children and adolescents is assessed. (<i>true</i>)	76.6	2.9	20.4

Note. $N = 137-139$ in-service teachers. Missings are not listed. SEN EBD = special educational needs with regard to emotional-behavioral difficulties, SEN LD = special educational needs with regard to learning difficulties, PAD = partial achievement disorders.

Table 2. Mean (M), standard deviation (SD), and bivariate correlations of all measured variables.

Variables	Descriptive statistics		Bivariate correlations						
	<i>M</i>	<i>SD</i>	(1)	(2)	(3)	(4)	(5)	(6)	(7)
(1) Job-related contact	3.08	1.02	-						
(2) Attitudes	2.82	0.65	.03	-					
(3) Self-efficacy	3.20	0.97	.34***	.45***	-				
(4) Mastery-oriented teaching practices	3.99	0.67	.37***	.21*	.45***	-			
(5) Performance-oriented teaching practices	2.31	0.80	-.13	-.22**	-.20*	-.16	-		
(6) Knowledge DiWi	.51	.16	.33***	.25**	.39***	.34***	-.15	-	
(7) Misconceptions DiWi	.12	.06	.20*	.02	.10	.14	-.04	.48***	-
(8) Uncertainty DiWi	.37	.20	-.33***	-.21*	-.35***	-.32***	.14	-.96***	-.71***

Note. $N = 138-139$. The DiWi values indicate the percentage of answers coded as correct (knowledge), incorrect (misconceptions), and do not-know answers. All other variables were measured on a scale from 1 to 5.

* $p < .05$, ** $p < .01$, *** $p < .001$.

Supplemental Online Material

Contact with students of different diversity dimensions and knowledge about diversity were assessed also by subscales. Additional statistical analyses of these variables, including a comparison of teachers of different school types, are reported below.

Group differences (in line with H_1 & H_2)

The teachers had significantly higher knowledge in all five subscales compared to the bachelor students ($F[5,526] = 56.37, p < .001, \eta^2_{\text{part}} = .35$; see Table 3). Moreover, they also had slightly more misconceptions in all five subscales than the bachelor students ($F[5,526] = 6.86, p < .001, \eta^2_{\text{part}} = .06$; see Table 3). Almost the same pattern was found for the comparison between in-service teachers and master students: Teachers had higher knowledge in the subscales except for gender compared to the master students ($F[5,312] = 10.83, p < .001, \eta^2_{\text{part}} = .15$; see Table 4). They also had more misconceptions in the subscales except for SEN EBD than the master students ($F[5,312] = 14.85, p < .001, \eta^2_{\text{part}} = .19$; see Table 4).

We found higher knowledge for teachers for special needs education compared to those for mainstream education for the subscales SEN EBD, SEN LD and PAD, but not for gender and giftedness ($F[5,132] = 15.16, p < .001, \eta^2_{\text{part}} = .37$; see Table 5). Teachers for special needs education did not differ in their misconceptions in the subscales ($F[5,132] = 1.04, p = .40, \eta^2_{\text{part}} = .04$).

There were no differences in subscale-specific knowledge between teachers who participated in a further training or certificate course focusing on inclusion and those who did not ($F[5,131] = 1.95, p = .09, \eta^2_{\text{part}} = .07$). Also, there were no differences in misconceptions in the subscales ($F[5,131] = 0.86, p = .51, \eta^2_{\text{part}} = .03$).

Bivariate correlations (in line with H_3)

The correlations of teachers' contact to students of different diversity dimensions with knowledge were consistently positive (SEN EBD: $r = .35, p < .001$; SEN LD: $r = .44, p <$

.001; PAD: $r = .31, p < .001$; giftedness: $r = .18, p = .04$). The correlations with misconceptions were also positive but smaller (SEN EBD: $r = .07, p = .45$; SEN LD: $r = .19, p = .03$; PAD: $r = .13, p = .12$; giftedness: $r = .21, p = .01$).

Comparing teachers from different school types

An overview of teachers' knowledge and misconceptions in the subscales and the overall test, sorted by teachers' current school type, is shown in Figure 1 and 2. Since children in Germany are, after being taught together in primary school, distributed to different types of schools based on their academic achievement, diversity is, in addition to schools specialized on students with SEN (in what follows, special schools), particularly high in primary schools. We thus compared teachers in primary and special schools with teachers of other schools in their overall knowledge and misconceptions.

Teachers in primary schools had higher overall knowledge ($M = .57, SD = .12$) than teachers in grammar schools ($M = .47, SD = .18; F[61] = 5.90, p = .02, \eta^2_{\text{part}} = .09$) and vocational colleges ($M = .45, SD = .17; F[41] = 6.80, p = .01, \eta^2_{\text{part}} = .14$). However, the differences disappeared after adding the control variables age, gender, teaching experience in years, job-related contact, and having studied special needs education. There were no significant differences in overall knowledge compared to teachers in secondary schools ($M = .51, SD = .15; F[65] = 2.57, p = .11, \eta^2_{\text{part}} = .04$). Also, there were no significant differences in overall misconceptions of teachers in primary schools ($M = .12, SD = .05$) compared to teachers in secondary schools ($M = .12, SD = .07; F[65] = 0.01, p = .94, \eta^2_{\text{part}} < .001$), grammar schools ($M = .11, SD = .07; F[61] = 0.43, p = .52, \eta^2_{\text{part}} = .01$), and vocational colleges ($M = .11, SD = .08; F[41] = 0.20, p = .66, \eta^2_{\text{part}} = .01$).

Teachers in special schools ($M = .55, SD = .15$) did not differ from teachers in secondary schools ($M = .51, SD = .15; F[55] = 0.71, p = .40, \eta^2_{\text{part}} = .01$), grammar schools ($M = .47, SD = .18; F[51] = 2.41, p = .13, \eta^2_{\text{part}} = .05$), and vocational colleges ($M = .45, SD =$

.17; $F[31] = 3.04, p = .09, \eta^2_{\text{part}} = .09$) in their overall knowledge. Also, misconceptions did not differ between teachers in special schools ($M = .11, SD = .05$) and teachers in secondary schools ($M = .12, SD = .075; F[55] = 0.32, p = .58, \eta^2_{\text{part}} = .01$), grammar schools ($M = .11, SD = .07; F[51] = 0.01, p = .93, \eta^2_{\text{part}} < .001$), and vocational colleges ($M = .11, SD = .08; F[31] < 0.001, p = .99, \eta^2_{\text{part}} < .001$).

Table 3. Group comparisons between teachers and bachelor students in their knowledge and misconceptions in the DiWi subscales.

Dependent variable		Teachers (<i>n</i> = 138)		Bachelor students (<i>n</i> = 394)		MANOVA			
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>F</i>	df	<i>p</i>	η^2_{part}
Knowledge	SEN EBD	.60	.18	.36	.23	122.88	1, 530	< .001	.19
	SEN LD	.56	.20	.31	.23	129.15	1, 530	< .001	.20
	PAD	.57	.19	.24	.22	246.33	1, 530	< .001	.32
	Gender	.41	.20	.28	.18	56.06	1, 530	< .001	.10
	Giftedness	.40	.19	.25	.19	63.53	1, 530	< .001	.11
Misconceptions	SEN EBD	.09	.07	.07	.08	11.00	1, 530	< .001	.02
	SEN LD	.11	.08	.07	.09	29.45	1, 530	< .001	.05
	PAD	.12	.09	.08	.11	11.34	1, 530	< .001	.02
	Gender	.13	.09	.09	.10	15.6	1, 530	< .001	.03
	Giftedness	.15	.11	.10	.12	19.30	1, 530	< .001	.04

Note. SEN EBD = special educational needs with regard to emotional-behavioral difficulties, SEN LD = special educational needs with regard to learning difficulties, PAD = partial achievement disorders.

Table 4. Group comparisons between teachers and master students in their knowledge and misconceptions in the DiWi-K subscales.

Dependent variable		Teachers (<i>n</i> = 138)		Master students (<i>n</i> = 180)		MANOVA			
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>F</i>	df	<i>p</i>	η^2_{part}
Knowledge	SEN EBD	.67	.24	.46	.29	47.00	1, 316	< .001	.13
	SEN LD	.53	.24	.38	.27	27.81	1, 316	< .001	.08
	PAD	.56	.24	.43	.29	16.86	1, 316	< .001	.05
	Gender	.40	.23	.36	.23	2.47	1, 316	.12	.01
	Giftedness	.50	.19	.40	.27	13.38	1, 316	< .001	.04
Misconceptions	SEN EBD	.09	.11	.08	.11	0.42	1, 316	.52	< .001
	SEN LD	.16	.15	.12	.15	6.29	1, 316	.01	.02
	PAD	.13	.14	.09	.12	9.19	1, 316	.003	.03
	Gender	.17	.13	.09	.14	22.03	1, 316	< .001	.07
	Giftedness	.23	.16	.15	.16	20.98	1, 316	< .001	.06

Note. SEN EBD = special educational needs with regard to emotional-behavioral difficulties, SEN LD = special educational needs with regard to learning difficulties, PAD = partial achievement disorders.

Table 5. Group comparisons between teachers of special needs education and mainstream education in their knowledge in the DiWi subscales.

Dependent variable	Special needs education (<i>n</i> = 30)		Mainstream education (<i>n</i> = 108)		MANOVA				
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>F</i>	df	<i>p</i>	η^2_{part}	
Knowledge	SEN EBD	.76	.13	.55	.17	39.24	1, 136	< .001	.22
	SEN LD	.74	.12	.51	.20	34.45	1, 136	< .001	.20
	PAD	.69	.13	.54	.19	15.92	1, 136	< .001	.11
	Gender	.41	.22	.41	.19	0.02	1, 136	.90	.00
	Giftedness	.43	.19	.39	.19	1.28	1, 136	.26	.01

Note. SEN EBD = special educational needs with regard to emotional-behavioral difficulties, SEN LD = special educational needs with regard to learning difficulties, PAD = partial achievement disorders.

Figure 1. Mean knowledge about SEN EBD, SEN LD, PAD, gender and giftedness sorted by teachers' current school type.

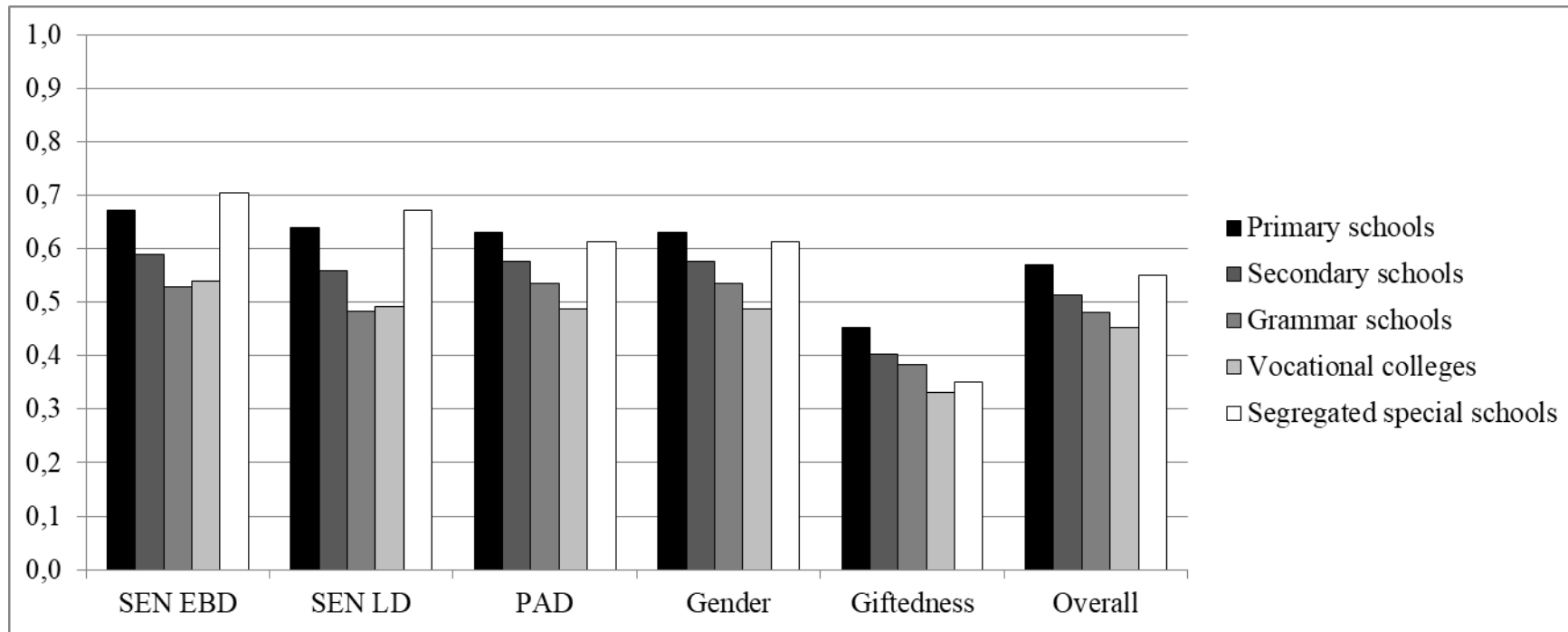
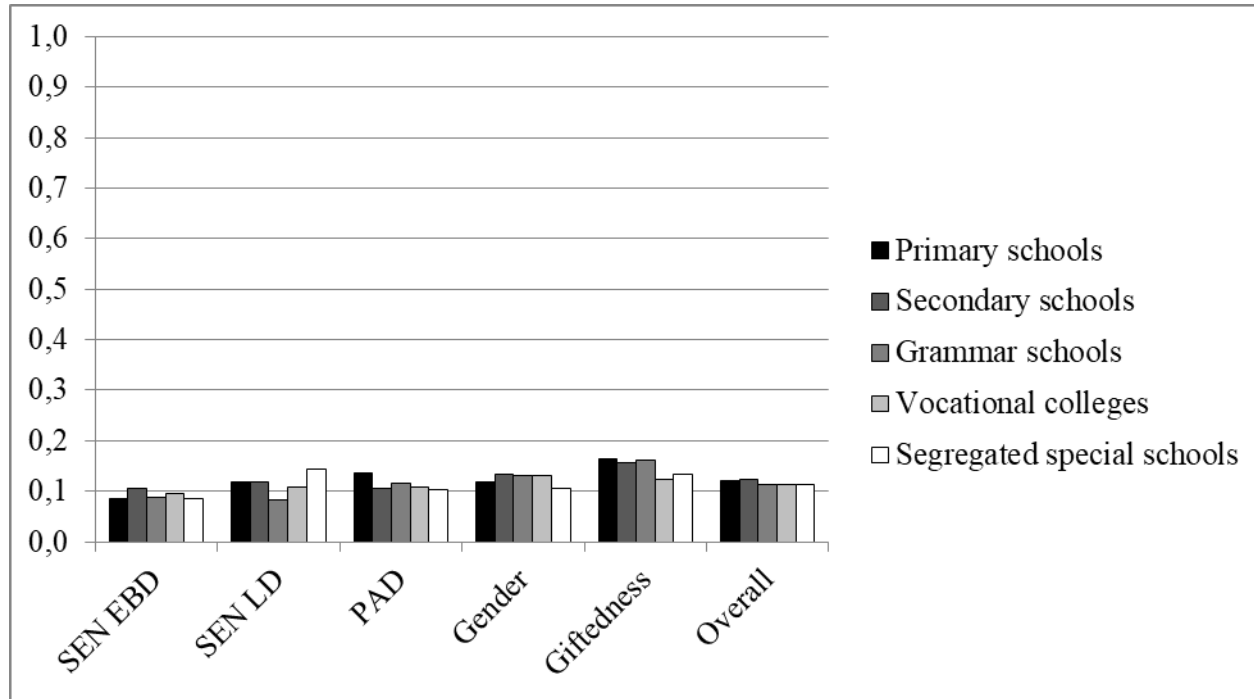


Figure 2. Mean misconceptions about SEN EBD, SEN LD, PAD, gender and giftedness sorted by teachers' current school type.



5.3 Studie 3: Links between teachers' knowledge about special educational needs and students' social participation and academic achievement in mainstream classes

Note: This is the first author's version of a work that was published in the journal *Contemporary Educational Psychology*. This article does not exactly replicate the final version that is published in the journal, it is not a copy of the original published article, and it is not suitable for citation.

Tometten, L., Heyder, A., & Steinmayr, R. (2021). Links between teachers' knowledge about special educational needs and students' social participation and academic achievement in mainstream classes. *Contemporary Educational Psychology*, 67, 102022. <https://doi.org/10.1016/j.cedpsych.2021.102022>

Acknowledgement: We thank Caroline Kirchner and Mark Fischer, without whom this study would not have been possible, for their engagement in data collection.

Abstract

Teachers' pedagogical knowledge is considered a prerequisite for effective teaching and is also expected to be relevant in highly diverse, inclusive classes. This study examines the social participation and academic achievement of children with and without special educational needs (SEN) due to emotional and behavioural (EBD) and learning (LD) difficulties and the importance of teachers' knowledge of these SEN. In 25 primary school classes with at least one child with SEN EBD ($N_1 = 421$) and 40 classes with at least one child with SEN LD ($N_2 = 715$), multilevel analyses revealed poorer performance of students with SEN. Students with SEN EBD also felt less integrated than their peers without SEN. Teachers' EBD knowledge was positively related to every student's social participation and academic achievement, but teachers' LD knowledge was unrelated. The results underscore the importance of teacher knowledge for student outcomes in classes with students with SEN.

Keywords: Teachers' pedagogical knowledge, special educational needs, social participation, academic achievement, inclusive education

Highlights

- Study on students with special educational needs (SEN) in mainstream classes.
- Focus on SEN in learning (LD) and emotional and behavioral difficulties (EBD).
- Students with SEN (LD and partly EBD) show poorer performance than their peers.
- Students with SEN EBD but not with SEN LD feel less socially integrated.
- Teacher knowledge is related to student social participation and academic performance.

1. Introduction

Education in schools plays an important role in the development of children and young adults. Central educational goals include an increase in performance for all students and a reduction of differences in academic areas resulting from different challenges students have to face (Helmke, 1988; Kyriakides, Creemers, & Charalambous, 2018). Furthermore, students should feel socially valued and accepted and be part of the class community. These goals also apply to children with special educational needs (SEN) whose disadvantages regarding their achievement and social participation should be reduced through inclusive education (Kocaj et al., 2015). Approximately 25 years after the Salamanca Statement (UNESCO, 1994), there is an increasing trend towards educating students with SEN in mainstream schools (Ramberg et al., 2018). In Europe, 4.4% of all students had confirmed SEN, and almost a third of them were educated together with their peers without SEN.

But the simple placement of students with SEN in mainstream schools defines only the initial baseline in the process of achieving inclusive education, with the longer term goal of creating communities in which no differentiation is made based on the category SEN (e.g., Göransson & Nilholm, 2014). An important step on the way to achieve this goal is the promotion of good social-emotional school experiences and the acquisition of competencies for all children in the class (Florian et al., 2016; Grosche, 2015). However, studies indicate that this is not always the case and students with SEN often score lower than their peers on performance and social-emotional measures (e.g., Crede et al., 2019; Nelson et al., 2004).

Teachers are considered crucial for fostering students' learning (e.g., Baumert & Kunter, 2013) and social-emotional school experiences (Audley-Piotrowski et al., 2015; Farmer et al., 2011) and play an important role in the implementation of inclusion. Their general pedagogical knowledge, which – unlike pedagogical content knowledge – is independent of the specific subject, has been found to be associated with effective classroom

management (König & Pflanzl, 2016) and instructional quality (Voss et al., 2011) in regular classes. Little empirical research has focused on the role of teachers' knowledge in the context of inclusive education although its theoretical relevance has often been highlighted (Engsig & Johnstone, 2015; Kracke, 2014; Paju et al., 2016).

The present study pursues two goals. First, we want to shed more light on the academic achievement and social participation of students with SEN due to emotional and behavioral difficulties (EBD), learning difficulties (LD), and of their classmates in mainstream classes. Students with these types of SEN represent the two largest groups of SEN in inclusive education in Germany (Conference of the Ministers of Education and Cultural Affairs, 2018). Second, we want to empirically explore the role of teachers' knowledge about both SEN for the social participation and academic achievement of students with and without SEN, as teacher knowledge might be one factor influencing the successful inclusion of students with SEN (Engsig & Johnstone, 2015). In order to have valid indicators of students' achievement as well as teachers' knowledge, we applied standardized tests to assess them. Please note that by studying the relation between teachers' knowledge and their students' achievement and social integration in mainstream classes, we reject looking at students with SEN from a deficit perspective but rather like to point towards teachers as agents of change towards the successful implementation of inclusive education.

1.1. Students with emotional and behavioral difficulties in mainstream classes

Here, we used the abbreviation EBD to describe students with special educational needs due to emotional and behavioral difficulties (see also Banks et al., 2017). The terms emotional/behavioral disorders (E/BD; Chen, 2006; Nelson et al., 2004), social, emotional, and behavioral difficulties (SEBD; Smeets, 2009), and behavioral, emotional, and social difficulties (BESD; Goodman & Burton, 2010) have also been used in international literature to describe synonyms or similar phenomena.

Disorders such as hyperkinetic disorders (e.g., attention deficit hyperactivity disorder [ADHD]), social behavior disorders (e.g., dissocial and aggressive behavior), and childhood emotional disorders (e.g., phobias and social anxiety), which are described in ICD-10 (World Health Organization, 2016), may – but do not automatically – lead to a need for special educational support with a focus on EBD. But also students without an ICD-10 diagnosis may have SEN EBD: The Ministry for Schools and Education of North Rhine-Westphalia (2016) defines that there is a SEN with a focus on EBD if a student persistently refuses or resists education so that he or she cannot be sufficiently supported in the classroom. By definition, his or her development and that of the classmates may be disturbed or endangered. Students with SEN EBD are more likely to influence the lessons for all students in a class and this might explain why teachers are particularly negative about the inclusion of children with SEN EBD compared to other SEN or disabilities (for an overview see de Boer et al., 2011). In fact, mainstream teachers are more likely to change schools if they have to teach students with SEN EBD (Gilmour & Wehby, 2019). And the relationship between teachers and students with internalizing and externalizing behavioural problems is characterized by more conflicts than the relationship between teachers and students without those problems (Henricsson & Rydell, 2004).

Students with SEN EBD are at risk of poor outcomes (Nye et al., 2016). This holds for both school experiences and academic achievement. School experiences include students' social participation which can be characterized by four key concepts: relationships between students with and without SEN, positive interactions between them (e.g., working together on tasks), the self-perception of the students with SEN, and acceptance by their classmates (Koster et al., 2009). Perceived classroom climate and social integration can be seen as constructs indicating the extent of social participation (Crede et al., 2019; Henke et al., 2017; Heyder et al., 2020). Students with SEN EBD were more likely to have no or only one friend

and more peer problems than their classmates (Banks et al., 2017), which results in poor social integration (Schuck & Rauer, 2018). In a recent study, students with SEN EBD reported a poorer classroom climate and lower social integration than their control group formed by propensity score matching with students without SEN (Crede et al., 2019). Compared to all other students in the class without SEN, students with SEN EBD reported more negative emotional-social school experiences.

Furthermore, students with SEN EBD also performed more poorly on different measures of academic achievement such as grades (e.g., Sainio et al., 2019) or standardized tests in reading, writing and maths (e.g., Nelson et al., 2004). Antisocial behaviour correlated negatively with academic achievement (Dalton, 2010), and students with SEN EBD who were educated in an inclusive setting from kindergarten to Grade 12 showed deficits relatively to their classmates in all measured aspects of academic achievement including reading, writing, and maths (Nelson et al., 2004). The 2012 PISA survey also showed that German students with SEN EBD in mainstream classes achieved on average only about 85% of the points of their peers without SEN in reading and maths (Gebhardt et al., 2015). Overall, there is piecemeal evidence that students with SEN EBD in mainstream classes score lower than their peers on measures of social participation and academic achievement. The factors which moderate these effects and thus further promote successful inclusion have rarely been tested. As already been stated, the role of teachers seems especially important to us, as they are heavily involved in both students' social-emotional und academic development.

1.2. Students with learning difficulties in mainstream classes

Here, we used the abbreviation LD to describe SEN due to learning difficulties (see also Banks et al., 2017; Kavale & Forness, 1996; Vaughn et al., 2001). SEN LD are defined as severe, extensive, and long-lasting learning and performance deficits (Ministry for Schools

and Education of North Rhine-Westphalia, 2016). The 2012 PISA survey showed that German students with SEN LD in mainstream classes achieved on average about 80% of the points of their peers without SEN in reading and maths (Gebhardt et al., 2015). They were able to solve basic mathematical problems when all information was available, but they were not able to select and apply simple problem-solving strategies as well as their peers could. With regard to reading competencies, students with SEN LD understood the main idea of a text but were less able to recognize relations and distinguish important information from unimportant ones (OECD, 2014).

Furthermore, these students also had more problems in social areas: An older meta-analysis demonstrated that students with SEN LD showed fewer social competencies than their peers ($ES = 0.65$; Kavale & Forness, 1996). Results were comparable for teacher-, peer-, and self-ratings. Teachers considered academic problems to be the most serious, followed by the lack of social competencies of students with SEN LD. A further review demonstrated that students with SEN LD had lower social status than their peers without SEN LD, which was indicated, for example, by their classmates' lower acceptance (Vaughn et al., 2001). But unlike students with SEN EBD, they did not have an increased risk of having no or only one friend. Nevertheless, relationships with peers were characterized by more problems (Banks et al., 2017).

However, there is also research suggesting that students with SEN LD do not differ or only differ to a small extent in their social-emotional school experience from students without SEN. For instance, significant differences to the disadvantage of students with SEN LD were only found for the self-perceived academic self-concept and social integration but not for perceived classroom climate, school attitude, achievement motivation, pleasure of learning, and relation to the teachers (Schuck & Rauer, 2018). The results by Henke et al. (2017) support these findings even though these researchers did not distinguish between different

types of SEN. However, among students with SEN, those with LD formed the largest proportion with 47%. Here, students with SEN LD and without SEN did not differ in perceived classroom climate and social integration when the comparison group was formed by means of propensity score matching (Henke et al., 2017). Overall, SEN LD are characterized by performance deficits, which are consistently reported in empirical research. The role of social variables on the other hand remains unclear, which may be due to the fact that the groups which have been studied so far differed considerably from each other.

1.3. The importance of teacher knowledge in the context of inclusive education

Teachers are considered to be important actors in the education system in general and for implementing inclusion in particular (Loreman, 2007). Shifting the focus away from a deficit perspective to students with SEN, teachers have a high level of responsibility for these students' social-emotional school experiences and learning development. They can provide adequate support for them, act as an "early warning system" and enable early intervention (e.g., Smeets, 2009). Although many teachers have a positive attitude towards inclusion and consider diversity in the classroom to be an enrichment (Horne & Timmons, 2009), the inclusion of students with SEN creates many challenges for them in areas such as diagnostics, individual support, and working in a team with special needs teachers and therapists (Kracke, 2014). Teachers themselves see a need for additional training in inclusive practices (Horne & Timmons, 2009) and researchers also highlight teachers' need for adequate knowledge to teach inclusively (Engsig & Johnstone, 2015; Paju et al., 2016).

The importance of teachers' general pedagogical knowledge (GPK) has been demonstrated in several studies which have been investigated in regular but not in inclusive classes. GPK correlated positively with teachers' classroom management (König & Pflanzl, 2016) and instructional quality operationalized by cognitive activation and pace of instruction perceived by the students (Voss et al., 2011). In addition, GPK was associated with a lower

probability of experiencing burnout, directly and indirectly via the positive correlation with teachers' self-efficacy (Lauermann & König, 2016). But what do mainstream teachers know and what should they know supplementary to GPK when teaching inclusive classes?

In the context of inclusive education, most research has dealt with the importance of teachers' attitudes and self-efficacy rather than their knowledge (e.g., Heyder et al., 2020; Krischler & Pit-ten Carte, 2019). Teaching students with SEN provides the biggest change in teachers' daily school experiences elicited by the pursuit of inclusion (Ramberg et al., 2018; UNESCO, 1994). Therefore, we focus on teachers' knowledge about SEN here. In addition to more general knowledge about inclusive education, we consider target group-specific knowledge to be particularly important (see also Sousa et al., 2017). Please note however that this knowledge represents only one part of all the knowledge that might be considered as relevant for the successful implementation of inclusive education (e.g., Black-Hawkins & Florian, 2012, craft knowledge of teachers' inclusive practice; König et al., 2017; Klug et al., 2013, diagnosis and intervention; Kuyini & Desai, 2007, policy guidelines, specific teaching practices).

So far, little research has been conducted on mainstream teachers' knowledge about students with SEN. One obvious explanation is the lack of adequate instruments. Available knowledge tests about student diversity as one dimension of pedagogical knowledge showed an insufficient reliability for the specific subscale in contrast to the test as a whole (e.g., Voss et al., 2011), did not reflect any specific SEN (e.g., König & Blömeke, 2009; pedagogical knowledge), or reflected only a small aspect of SEN (e.g., Sousa et al., 2017; knowledge about developmental dyscalculia). We are aware of only one instrument that assesses teachers' knowledge about both SEN EBD and LD (see Tometten et al., 2022) in a long and short version. Prior research indicated that pre-service teachers' knowledge about diversity aspects (including SEN EBD and LD) was related to their self-efficacy and attitudes towards

inclusion, and pre-service teachers of special needs education knew more than teachers of other teaching professions (Tometten et al., 2022). The same pattern of results was demonstrated for other measures in studies with in-service teachers: Special education teachers knew more about intervention strategies in dyscalculia than other teachers (Sousa et al., 2017). Furthermore, experienced and student teachers had inadequate knowledge of ADHD and answered only a third of the knowledge items correctly (Akram et al., 2009).

Other studies assessed knowledge only by self-reports: A cross-country comparative study revealed partial differences in teachers' self-reported knowledge from Ghana, Germany, and Spain. Whilst teachers did not differ in their self-reported knowledge about strategies and characteristics of students with specific learning disabilities, teachers from Germany and Ghana reported higher knowledge about characteristics of students with SEN EBD (Mónico et al., 2018). In a further study, teachers rated their confidence in the teaching of SEN students higher than their knowledge about their students' specific problems (Paju et al., 2016). However, due to known limitations of self-reported competencies such as overestimating one's knowledge (e.g., von Stumm, 2014), these studies only give limited insights into teachers' knowledge about students with SEN.

Although theoretically plausible and highly relevant for teacher education, we are aware of only one study investigated the relation between teachers' knowledge and teaching quality in inclusive classrooms. Teachers' knowledge about the role of different actors in inclusion processes predicted the observed adaptive instructions during the lesson (Kuyini & Desai, 2007). However, there seems to be no study that assessed teachers' knowledge of different SEN by standardized tests and investigated the relationship with variables on the student level. In what follows, we elaborate on the theoretically expected relationship between teachers' knowledge about SEN EBD and LD and their students' social participation and academic achievement. As education should a) provide students with the best possible

support for their individual development while b) reducing existing differences (Kyriakides et al., 2018), we thereby refer to a) the quality and b) the equity in mainstream classes with students with SEN EBD and LD.

Teachers meet the challenge to meet the individual needs of students with SEN EBD while offering effective teaching practices for all students in the class (Goodman & Burton, 2010). From research in mainstream and inclusive classes and models of teachers' professional competence, we know that different knowledge domains predict teachers' use of effective teaching practices (König & Pflanzl, 2016; Kuyini & Desai, 2007) and are relevant for student outcomes (Baumert & Kunter, 2011; Depaepe et al., 2013; Voss et al., 2015). Self-evidently, knowledge does not automatically lead to better student outcomes, but it does increase the probability that teachers apply effective teaching strategies and that all students are guided toward their individual best. In mainstream classes with students with SEN EBD, teachers' knowledge about EBD should be positively related to the outcomes of all students in class (= quality; Kyriakides et al., 2018) because this knowledge should help teachers to create an appropriate – in this case less or undisturbed – learning environment that benefits all students. Furthermore, existing differences between students with and without SEN EBD in social participation and academic achievement (Crede et al., 2019; Nelson et al., 2004) should be reduced but not necessarily eliminated, the more the teacher knows (= equity; Kyriakides et al., 2018) because teachers' knowledge about SEN EBD should especially help to support the individual student with SEN EBD.

Compared to SEN EBD, SEN LD primarily concerns the learning process of the affected children resulting in lower achievement (Gebhardt et al., 2015). Thus, we theorize teachers' knowledge about SEN LD as especially helpful in teaching these students, resulting in smaller but not necessarily zero differences between students with and without SEN (= equity; Kyriakides et al., 2018).

1.4. Study overview and hypotheses

This study focuses on two central aims: First, we want to examine whether students with SEN EBD or LD differ from students without SEN. Second, we want to gain first insights into the relationship between teacher knowledge about SEN EBD and LD and students' social participation and academic achievement in inclusive education.

Even though distinctions are made between students with SEN EBD and those with SEN LD (Ministry for Schools and Education of North Rhine-Westphalia, 2016), many studies focusing on students' school experiences did not differentiate between these groups of students (e.g., Henke et al., 2017; Heyder et al., 2020). There is also a lack of studies including variables on classroom level in order to explain differences in students' school experiences and academic achievement between classes in inclusive education. Although teachers' knowledge of diversity is claimed to be relevant for student outcomes (Baumert & Kunter, 2011; Voss et al., 2015), their knowledge about EBD and LD, the two largest groups of SEN (Conference of the Ministers of Education and Cultural Affairs, 2018), for students in inclusive education has not yet been investigated. Many studies that are dealing with teacher knowledge in inclusive settings only focused on a small aspect of SEN (e.g., Sousa et al., 2017; dyscalculia) or relied on self-report measures (e.g., Paju et al., 2016).

To address these research gaps and overcome some of the previous studies' limitations, we set up a study with the following main characteristics: (1) Inclusion of teachers' knowledge about SEN EBD and LD as the central predictor in multilevel analyses to predict student outcomes, (2) measurement of teachers' knowledge by a standardized knowledge test, (3) social participation, operationalized by student-reported classroom climate and social integration, and academic achievement, operationalized by standardized reading comprehension and basic math tests, as outcome variables, and (4) separate analyses for students with SEN EBD and LD.

Based on prior research addressing students with SEN's social-emotional (e.g., Crede et al., 2019) and academic school outcomes (e.g., Nelson et al., 2004; Gebhardt et al., 2015) as well as teachers' pedagogical knowledge (König & Pflanzl, 2016; Kuyini & Desai, 2007), we expected the following results:

- H_1 : There are differences between students with and without SENEED in social participation and academic achievement disfavoring those with SEN.
- H_2 : In mainstream classes with students with SEN EED, teacher knowledge is positively related to the social participation and academic achievement of all students.
- H_3 : The more the teacher knows about SEN EED, the smaller the differences in social participation and academic achievement between students with and without SEN EED.
- H_4 : There are differences between students with and without SEN LD in academic achievement disfavoring those with SEN.
- H_5 : The more the teacher knows about SEN LD, the smaller the differences in academic achievement between students with and without SEN LD.

Since previous study results provided heterogeneous evidence regarding the social participation of students with SEN LD (e.g., Henke et al., 2017; Schwab, 2015), we refrained from formulating a directed hypothesis here and only explored the effects of SEN LD on students' social participation.

2. Method

2.1. Participants

The data stemmed from a cross-sectional study on students' emotional-social school experiences in mainstream classes. A total of 17 primary schools from North Rhine-Westphalia (Germany) participated in the study yielding data of 1091 third and fourth grade students and 51 teachers. In order to address the present study's hypotheses, we drew the two

analytic subsamples according to the following criteria: (1) Only classes with at least one student with SEN EBD (Subsample 1) or LD (Subsample 2) were included (eight students had both SEN and were therefore included in both subsamples). (2) Within these classes, all students with other SEN were excluded. The information on how many students with SEN were in each class was kept as a control variable (see below). Since there were classes in which there were both students with SEN EBD and LD, the analytic subsamples described below were partly overlapping.

Subsample 1 consisted of 421 students of whom 212 were female (50.3%), 204 were male (48.5%), and there was no gender information about five students. Students were on average 9.65 years old ($SD = 0.65$) and 35 of them had SEN EBD (80% boys, 54% with migration background). They were nested in 25 classes with an average of 23.08 children ($SD = 3.26$). Since not all children participated, the average class size was reduced to 20.83 children ($SD = 2.99$). Since two classes had the same main teacher and one teacher was ill, 23 teachers remained of which 22 were female (91.7%) and one was male (4.2%). They were on average 43.57 years old ($SD = 11.22$), had 15.41 years of teaching experience ($SD = 9.46$), and have worked at the current school for 11.26 years ($SD = 7.46$).

Subsample 2 consisted of 715 students, of whom 364 were female (50.9%), 342 were male (47.8%), and there was no gender information about nine students. They were on average 9.62 years old ($SD = 0.66$) and 76 of them had SEN LD (47% boys, 67% with migration background). They were nested in 40 classes with an average of 24.21 children ($SD = 3.17$). On average, 21.59 children per class ($SD = 4.12$) were surveyed. They were taught by 39 main teachers because the two classes with the same teacher were also included. There were 33 female (82.5%) and four male teachers (10.0%), three did not complete the questionnaires. They were on average 40.36 years old ($SD = 10.62$), had 12.43 years of teaching experience ($SD = 8.74$), and have worked at the current school for 8.24 years ($SD =$

6.44).

2.2. Procedure

Data was collected during regular class hours by trained research assistants from April to July 2018. In German primary schools, there is one teacher called the class teacher who is responsible for the class, spends most of the time with the students, teaches most of the subjects, and is responsible for organizational and social issues. We chose this teacher to answer our questionnaire. Following established ethical principles, participation was voluntary for both teachers and students. Parents' informed consent was obtained before student participation.

2.3. Measures

2.3.1. Level 1: Student variables

Social participation. Two scales of the German questionnaire for assessing students' social- emotional school experiences (Fragebogen zur Erfassung emotionaler und sozialer Schulerfahrungen; FEES 3–4; Rauer & Schuck, 2003) were used as indicators of students' social participation. Classroom climate described the cohesion within the class from the perspective of the surveyed student (e.g., “All children are allowed to play along”; $\alpha = .74/.77$ for third and fourth grade students, respectively). Social integration indicated the extent to which the student feels integrated into the class (e.g., “My classmates are nice to me”; $\alpha = .79/.84$ for third and fourth grade students, respectively). Both scales consisted of 11 items and the answers were coded from *1 = does not apply at all* to *5 = applies exactly*.

Academic achievement. Reading comprehension was measured using the ELFE II text comprehension subtest (Lenhard et al., 2017; r_{tt} [4 weeks] = .85³). The students were instructed to complete the 20 assignments as quickly and thoroughly as possible within seven

³ Since measures of internal consistency are not meaningful for speed tests, retest-reliability will be reported for both reading and math tests (Nunnally & Bernstein, 1994).

minutes. The tasks were short texts (e.g., “Tim is happy when the sun shines, then he can play football with his friends”) with four different answers, of which only one was correct (e.g., “Tim likes to play football”). The answers were coded with $0 = \textit{not edited/wrong}$ and $1 = \textit{correct}$. *Basic math competencies* were measured using the subscale arithmetic operations of the Heidelberger Rechentest (HRT 1–4; Haffner et al., 2005; r_{tt} [1–2 weeks] = .93). The scale consisted of six subtests (addition, subtraction, multiplication, division, completion tasks, and larger-smaller comparisons). Each had to be completed within two minutes. The answers were coded with $0 = \textit{not edited/wrong}$ and $1 = \textit{correct}$. The sum of the correct answers was used as the raw value per subtest. Only for the subtest larger-smaller comparisons, half of the errors were subtracted from the number of correct answers (Haffner et al., 2005).

SEN. Similar to other studies in Germany (e.g., Crede et al., 2019; Heyder et al., 2020), the teachers indicated whether the students had a need for special educational support (SEN) or not and if so, the specific kind of SEN. They indicated yes when SEN status had already been diagnosed in a corresponding procedure, when a procedure had been initiated but not yet completed, and when the child had received special support preventively without having an official diagnosis yet. A distinction in the type of SEN was made between EBD, LD, and others.

Control variables. The control variables were based on student indications. For this purpose, gender ($0 = \textit{female}$, $1 = \textit{male}$), age, and migration background ($0 = \textit{no migrant history}$, $1 = \textit{migrant history}$) were reported. A migrant background was coded if the child or at least one parent was born outside Germany or if the first language learned or the main language at home was not German. The approximate number of books at home was used as an indicator of socio-economic status (see Paulus, 2009). The scale ranged between $1 = 0\text{--}10$ books and $5 = \textit{over 200}$ books.

2.3.2. Level 2: Teacher and classroom variables

Teacher knowledge about EBD and LD. Teachers' declarative factual knowledge about the two SEN was measured through two scales of the DiWi-K (Tometten et al., 2022; short version EBD: $\alpha = .66$, LD: $\alpha = .70$). Each scale consisted of seven items (EBD: e.g., "Strongly fluctuating performance, daydreaming or motor clumsiness can indicate ADHD"; LD: e.g., "The probability of a learning impairment increases with decreasing socio-economic status"). The teachers evaluated whether the statement was correct, wrong, or whether they did not know. In order to quantify the knowledge, all answers were coded with $0 = do-not-know/wrong$ and $1 = correct$. Further information on the test can be found in the Appendix.

Control variables. Teachers' age, years at the current school, attitudes towards integration at mainstream schools, and self-efficacy in teaching students with SEN were used as control variables. The attitudes were measured by a slightly adapted version of the German questionnaire attitudes towards inclusion (Einstellungen zur Integration [EZI]; Kunz et al., 2010). The subscale educational support and encouragement consisted of seven items (e.g., "The quality of regular class teaching becomes better when children with SEN are integrated"; $\alpha = .90$). Four items measured the social integration subscale (e.g., "The more time children with SEN spend in a regular class, the more likely they are to be treated well by other classmates of their class"; $\alpha = .82$). Teachers' self-efficacy was assessed by the German self-efficacy scale (Skala zur Allgemeinen Selbstwirksamkeitserwartung [SWE]; Schwarzer & Jerusalem, 1999). We used four items and reformulated them for the context of teaching children with SEN EBD (e.g., "Whatever happens to children with SEN EBD in class, I will get along"; $\alpha = .89$) and LD (e.g., "I can find a solution for any problem with children with SEN LD in class"; $\alpha = .82$). The class size and number of students with SEN per class were used as additional class-level control variables.

2.4. Data analysis

We applied multilevel analyses in MPlus Version 6.12 (Muthén & Muthén, 1998-2017) with standard errors corrected for non-normality of variables (MLR) to test our hypotheses. Missing values on all metric variables were estimated with Full Information Maximum Likelihood (FIML). Taken together both subsamples, Little's MCAR test indicated missing completely at random only for the level-2- ($\chi^2 = 63.48, df = 50, p = .10$) but not for the level-1-variables ($\chi^2 = 8268.58, df = 7133, p < .01$). According to Newman (2003), FIML is the best option compared to other ad hoc procedures to handle missing data.

Step 1 of our analyses consisted of random-intercept-fixed-slope-models for each dependent variable with level-1-predictors. Step 2 consisted of random-intercept-random-slope-models with level-1-and-2-predictors. Random-intercept-fixed-slope-models with level-1-and-2-predictors were also calculated but there were no differences in the results compared to Step 2 and these models will not be reported. Step 3 should be the inclusion of cross-level interactions in order to test Hypotheses 3 and 5. As a prerequisite, we studied whether there was a significant slope variance in the models in Step 2 that could theoretically be traced back to a level 2 predictor. If no significant slope variance was found, we followed the "default position" (Aguinis et al., 2013) and did not continue with Step 3. Since teachers' knowledge and thus a variable at level 2 served as the central predictor in our hypotheses, all continuous independent variables were centered at the grand mean (see Enders & Tofighi, 2007).

We used control variables to determine the importance of students' SEN and teachers' knowledge above and beyond these variables to further ensure the robustness of our results. For this purpose, we have decided to use variables that have already been shown to be important in previous studies and, if possible, did not correlate with each other. Following the recommendation of Becker et al. (2016), the analyses were also calculated without control

variables. All differences found in the results will be reported. We used gender, age, socio-economic status, and migration background as control variables for the students (Crede et al., 2019; Karsten et al., 2001) and years of teaching experience, attitudes towards inclusion, and self-efficacy in teaching students with SEN for the teachers (Savolainen et al., 2012). We added class size (Blatchford & Webster, 2018), and number of children with SEN per class (Ruijs et al., 2010) as additional level 2-variables. Since eight students had both SEN EBD and LD, we controlled for the other SEN variable at student level and for the other knowledge variable at teacher level.

3. Results

3.1. Descriptive statistics

3.1.1. Students' characteristics

Descriptive statistics and bivariate correlations between all student variables are presented in Table 1. Boys were more likely to have SEN EBD than girls. Students with SEN LD on average were older, had a lower socio-economic status and more often a migration background. In both subsamples, SEN status correlated negatively with social integration, reading comprehension, and basic math competencies. Girls reported higher social integration (only Subsample 2) and boys achieved better results in the math test. Younger students reported a better class climate (both subsamples) and higher social integration (only Subsample 2). A higher socio-economic status correlated with higher values on all dependent variables. Students with migration background reported a more negative class climate and less social integration (the last only in Subsample 2) and achieved poorer results in the reading and math tests.

In both subsamples, values for classroom climate and social integration were in the upper range of the scale and students completed approximately half of the reading and math tests correctly. All dependent variables correlated positively with each other (see Table 1).

Intraclass correlation coefficients (ICCs) indicate how much of the variation in the dependent variables on level 2 can be attributed to between-class differences: 23% in classroom climate, 9% in social integration, 20% in reading comprehension, and 20% in basic math competencies in Subsample 1 and 31% in classroom climate, 13% in social integration, 22% in reading comprehension, and 18% in basic math competencies in Subsample 2.

3.1.2. Teachers' characteristics

Descriptive statistics and bivariate correlations between all teacher variables are presented in Table 2. None of the level 2-control variables correlated with teachers' knowledge. Teachers in Subsample 1 answered 73% of the knowledge items in the area of SEN EBD correctly. Teachers in Subsample 2 answered 60% of the knowledge items in the area of SEN LD correctly. These are high scores compared to subsamples of preservice teachers (Tometten et al., 2022).

3.2. Multilevel analyses

3.2.1. Results for subsample 1 (EBD)

In order to test our Hypotheses 1 to 3, we conducted multilevel analyses (see Table 3). There were differences between students with and without SEN EBD in social integration, and basic math competencies, disfavoring students with SEN EBD, but no differences in classroom climate and reading comprehension. Without the use of control variables, the difference in reading comprehension was significant ($\beta_1 = -0.08$, $p = .04$, $\beta_2 = -0.09$, $p = .03$; not presented in Table 3). For all differences found, students with SEN EBD had lower values than students without SEN EBD. Thus, Hypothesis 1 was partially supported.

Teachers' knowledge about EBD was positively related to the social participation (classroom climate and social integration) and academic achievement (reading comprehension and basic math competencies) of all students. Without including the control variables, however, regarding students' basic math competencies, it barely missed

significance ($\beta = 0.20$, $p = .07$; not presented in Table 3). Thus, Hypothesis 2 was mostly supported.

As there was no significant slope variance for SEN EBD, we decided against including cross-level interactions in the models.⁴ On the basis of these data, Hypothesis 3 must be rejected.

3.2.2. Results for subsample 2 (LD)

Results of multilevel analyses testing our Hypotheses 4 and 5 are presented in Table 4. There were significant relationships between SEN LD and academic achievement (reading comprehension and basic math competencies) in all models. Students with SEN LD performed lower than students without SEN. Hypothesis 4 is thus supported.

The exploratory analysis revealed no differences between students with SEN LD and students without SEN in classroom climate and social integration. Without including the control variables, there was a significant difference in social integration in both models ($\beta_{1/2} = -0.21$, $p = .01$). In terms of the level-2-predictors, there were no associations between LD knowledge and the student outcomes. Except for the math test, EBD knowledge was not related to the outcomes in the LD sample.⁵ However, teachers' attitudes toward the academic support of children with SEN in mainstream classes were positively related to students'

⁴ Even given the greater statistical power for testing cross-level interactions than slope variances (see Aguinis et al., 2013), the interaction term did not reach significance. In conclusion, no interaction effects were found. The same applies to the LD analyses.

⁵ The positive correlation between teachers' EBD knowledge and the students' math skills in LD classes could be due to the classes with both students with SEN EBD and students with SEN LD. In additional analyses, we checked whether this result was still found for LD classes if classes with students with SEN EBD were excluded. We found that here, teachers' EBD knowledge was not significantly related to any of the outcomes supporting our explanation. The correlation between teachers' LD knowledge and student outcomes did not change in these analyses.

classroom climate, social integration, and basic math competencies.

There was no significant slope variance for SEN LD and thus no cross-level interactions were specified. On the basis of these data, Hypothesis 5 must be rejected.

4. Discussion

This study focused on two student outcomes that represent central goals of inclusive education for all students: social participation (operationalized as perceived classroom climate and social integration) and academic achievement (operationalized as performance in two standardized reading and math tests). The first aim of this study was to examine whether students with SEN EBD or LD differ from students without SEN in these outcomes. Second, we wanted to gain first insights into the relationship between teacher knowledge about SEN EBD and LD and students' social participation and academic achievement in mainstream classes with students with SEN EBD and LD.

4.1. Students with SEN in mainstream classes

If there were differences between students with and without SEN in our data, they were all to the disadvantage of those with SEN as has also been shown in previous studies (Banks et al., 2017; Crede et al., 2019; Gebhardt et al., 2015; Nelson et al., 2004). Students with SEN EBD reported lower social integration and had poorer results in the reading and math tests. However, the difference in reading comprehension was only significant in the models without control variables suggesting a confounding of SEN EBD, socio-economic status and migration background (see also Crede et al., 2019; Strand & Lindsay, 2009). In contrast to the study by Crede et al. (2019), students with SEN EBD did not differ significantly from their peers without SEN in the perceived classroom climate. Possible explanations for the diverging results are the applied method and the students' age. Crede et al. (2019) applied propensity score matching (PSM). The control variables in our study (gender, age, socio-economic status, and migration) did not completely parallel those

variables used in Crede et al.'s study for the PSM (gender, SES, age, and student's class affiliation). Even though we controlled for clustered data, we did not only compare students in the same class due to the fact that this is not possible when running multi-level analyses. Crede et al. (2019) used different PSM procedures. The one which did not completely consider the classroom affiliation (Full Matching) resulted in group differences that were lower than those found for the other PSM procedures. Thus, it might well be that students with and without SEN EBD did not differ in the overall subsample but would have differed when only comparing students in the same class. Furthermore, compared to their study, students in our subsample were older and the class community had already existed for a longer time. Thus, the pattern might suggest that differences in the perceived classroom climate between students with SEN EBD and their classmates decrease in the course of elementary school, the more time the students have spent together. Furthermore, the different results in classroom climate and social integration indicate that primary school students can distinguish between the two constructs and are able to give a differentiated statement. The results also support the validity of teachers' assessment of SEN EBD because difficulties with social relationships and peer interaction are likely to reflect emotional and behavioral difficulties in students. This also applies to the partially poorer performance which can, among others, be affected by students' feeling of being poorly integrated (Bücker et al., 2018; Ministry for Schools and Education of North Rhine-Westphalia, 2016).

In addition, the results might also indicate that the inclusion of students with SEN EBD has still limitations with regard to their social integration which hinder the creation of communities as important goal of inclusion (Göransson & Nilholm, 2014). Encouragingly, the differences between students with and without SEN EBD are less pronounced in our

subsample⁶ in comparison to previous studies, in which differences up to $d = -1.5$ to the disadvantage of students with SEN EBD were found (e.g., Schuck & Rauer, 2018).

Students with SEN LD performed lower in the standardized academic achievement tests than their peers without LD, with stronger evidence compared to students with SEN EBD. This finding is consistent with previous research from the 1980s (Epstein & Cullinan, 1983) but has not been investigated in more recent studies. When interpreting this finding, one should keep in mind that poor performance is part of the diagnosis of SEN LD and using achievement tests with a time limit might have highlighted the differences even more. Our results support the validity of our data and should not generally be seen as a weakness in implementing inclusion even though we also see this as an indicator that efforts to achieve inclusion in schools should be further advanced in the future. All students should be encouraged and supported to their individual best (Kyriakides et al., 2018), which does not automatically mean that differences do not longer exist during the process of inclusion.

For social participation, there were no differences between students with and without SEN LD. Only without the use of control variables in the analyses, students with SEN LD reported lower social integration than their peers without SEN. The results suggest that SEN-status, gender, and socio-economic status are associated with each other (Strand & Lindsay, 2009), share common variance in the prediction of students' social integration, and potential differences between students with and without SEN LD cannot simply be explained by students' SEN status as such. The result is consistent with the finding that SEN LD is strongly dependent on students' background characteristics (e.g., Shifrer et al., 2011). Consequently, the heterogeneous findings of previous studies (e.g., Henke et al., 2017; Schuck & Rauer, 2018; Schwab, 2015) might be explained by different study designs and

⁶ The correlation between students' SEN EBD status and their social integration of $r = -.18$ (see Table 1) corresponds with $d = -0.37$.

analysis methods.

Overall, students with SEN EBD and LD showed different results, especially in terms of their social participation. However, previous studies often did not differentiate between these types of SEN (e.g., Henke et al., 2017; Schwab, 2015). In order to be able to empirically investigate inclusive education and to address the diversity of students, we therefore argue for a separate consideration of the different SEN in research. Please note that these groups are again heterogeneous in terms of their individual support needs and background variables.

4.2. Teacher knowledge in the context of inclusive education

Teachers' knowledge about EBD was positively related to all dependent variables in classes with students with SEN EBD while teachers' knowledge about LD was not, which underlines the specific nature of this knowledge. Thus, knowledge about EBD could explain variance in student outcomes over and above teachers' attitudes and self-efficacy, which have been investigated more intensively so far (e.g., Heyder et al., 2020; Krischler & Pit-ten Carte, 2019; Woodcock et al., 2019). What might explain the positive relation between teachers' knowledge and student outcomes? It seems plausible that teachers who know more about EBD can deal with these students more effectively during lessons. In these classes, children with SEN EBD therefore might have a lower risk to stand out negatively and teachers can both promote academic achievement and provide a basis for communication that strengthens the social integration of all students in the class. Note, however, that no conclusions about causality can be drawn based on the cross-sectional study design. Likewise, we can only speculate about the reasons for the differences between students with and without SEN due to the study design: These differences could, for example, have been caused or reinforced by specific expectations of teachers due to labelling a student as being a student with SEN (Shifrer, 2013).

Building on prior research indicating that both teachers and students benefit from teachers' general pedagogical knowledge (König & Pflanzl, 2016; Lauermann & König, 2016; Voss et al., 2011), we were able to show that specific knowledge about EBD is also positively related to student outcomes in inclusive settings. Although mainstream teachers are important actors in implementing school inclusion, they know less about diversity in general (Tometten et al., 2022) and, for example, dyscalculia in particular (Sousa et al., 2017) than teachers with expertise in special needs education. Training in inclusive practices as requested by the teachers themselves (Horne & Timmons, 2009) should therefore provide evidence-based knowledge (Engsig & Johnstone, 2015; Paju et al., 2016) at a theoretical level to complement practical components of teacher training.

In contrast to the findings in EBD classes, there were no effects of knowledge about LD. Since students with SEN LD do not necessarily have an effect on the whole class and the differences in academic achievement are a fundamental part of the diagnosis (Ministry for Schools and Education of North Rhine-Westphalia, 2016), it is likely that these large differences in academic achievement cannot be reduced by teachers' knowledge alone. In addition, these students usually bring a history of learning support, which was not successful over some years. Nevertheless, it is also possible that the use of items that are more behavior- and less facts-oriented would have led to different results. The knowledge test contained two items on teaching practices for supporting students with SEN EBD ("Appreciative remarks about desired behaviour lead to excessive demands on students with SEN EBD [wrong]" and "Students with SEN EBD are taught towards different goals than students without SEN EBD [wrong]") but no items about teaching students with SEN LD. Additional items with a stronger focus on specific teaching behaviors could provide further insights, especially in Subsample 2 with students with SEN LD.

Furthermore, differences between students with and without both types of SEN did

not vary significantly between classes suggesting that teachers' knowledge is unrelated to the gap. One possible explanation for this unexpected finding could be that competent teachers are more likely to be assigned to children with more severe SEN, so that teacher knowledge and SEN students' prior participation and achievement might be confounded.

4.3. Limitations and future research directions

We assessed teachers' knowledge about SEN and students' academic achievement and social participation in inclusive education. Due to its cross-sectional design, however, no causality could be tested. Our study provides a snapshot on these correlations and lays the ground for future studies considering the development of inclusive education in schools in a longitudinal design. These studies are needed in order to test whether initial differences in student outcomes (i.e., prior social participation and academic achievement) are reduced, unaffected, or increased by teachers' knowledge. Such longitudinal studies will also provide insights into the question whether differences between students with and without SEN simply support the validity of the SEN diagnosis or actually change over time, depending on the teacher. Regarding teachers' knowledge, the evaluation of misconceptions (= items that teachers mistakenly believe they can answer correctly) and their relevance for teaching behavior and student outcomes could provide new findings for the training of in-service but also of pre-service teachers (Schmiedeler, 2013; Thomm et al., 2021).

Furthermore, the present study did not include any information on the actual teaching behavior. We could thus not test whether teachers with a high level of knowledge were able to translate this knowledge into better teaching practices than those with a low level of knowledge (see also Florian & Black-Hawkins, 2011). In addition, no statement can be made about the specific implementation of inclusion (e.g., multiprofessional teams) compared to the simple placement of children with SEN (e.g., Grosche, 2015). Effective teaching in the context of inclusion might for example be characterized by prepared and spontaneous

adaptive instructions, academic progress monitoring, and characteristics of classroom management (Kuyini & Desai, 2007; Roy, Guay, & Valois, 2013). None of these variables could be tested here.

Moreover, with 35 students with SEN EBD in Subsample 1 and 76 students with SEN LD in Subsample 2, the sample size was relatively small, but still in the range of sample sizes in prior studies (e.g., Crede et al., 2019). It is therefore important to replicate the results in a larger sample in the future. A longitudinal design is also needed to separate teacher and class effects. As achievement test scores may not totally be comparable for students with and without SEN (Lane & Leventhal, 2015), the use of an adapted test may beyond that help to validate our findings.

Finally, there are some variables which have already been mentioned by other researchers and could be taken into account in future research addressing student outcomes in inclusive education. These include, for example, the intensity of the need for support at student level (Huber & Wilbert, 2012) and the encouragement of principals for inclusive program delivery at school level (Cobb, 2015). In this context, it is also interesting to consider how knowledge of SEN performs alongside more global knowledge measures such as GPK (Voss et al., 2011) or knowledge of inclusive education (Kuyini & Desai, 2007) in predicting classroom and student variables.

4.4. Conclusions

Implementing inclusive education is a highly important but challenging task for teachers. Despite many positive findings, the implementation of inclusion still seems to have some weaknesses. This is especially evident for students with SEN EBD, who, according to our analyses, are behind their classmates in terms of both social integration and school performance. Also, students with SEN LD showed worse school performance than their classmates without SEN, but similar levels of social participation. In addition, our results

support the importance of teachers' knowledge for student outcomes: Positive correlations were found between teachers' EBD knowledge, the social participation and academic achievement of all students in the class. The results support the call that all teachers should gain knowledge about SEN during pre-service teacher education or in-service teacher training programs in order to create a positive learning environment for all students.

References

- Aguinis, H., Gottfredson, R. K., & Culpepper, S. A. (2013). Best-practice recommendations for estimating cross-level interaction effects using multilevel modeling. *Journal of Management*, 39(6), 1490–1528. <https://doi.org/10.1177/0149206313478188>
- Akram, G., Thomson, A. H., Boyter, A. C., & McLarty, M. (2009). ADHD and the role of medication: Knowledge and perceptions of qualified and student teachers. *European Journal of Special Needs Education*, 24(4), 423–436. <https://doi.org/10.1080/08856250903223088>
- Audley-Piotrowski, S., Singer, A., & Patterson, M. (2015). The role of the teacher in children's peer relations: Making the invisible hand intentional. *Translational Issues in Psychological Science*, 1(2), 192–200. <https://doi.org/10.1037/tps0000038>
- Banks, J., McCoy, S., & Frawley, D. (2017). One of the gang? Peer relations among students with special educational needs in Irish mainstream primary schools. *European Journal of Special Needs Education*, 33(3), 396–411. <https://doi.org/10.1080/08856257.2017.1327397>
- Baumert, J., & Kunter, M. (2011). Das Kompetenzmodell von COACTIV [The COACTIV competence model]. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss, & M. Neubrand (Eds.), *Professionelle Kompetenz von Lehrkräften: Ergebnisse des Forschungsprogramms COACTIV [Teachers' professional competence: Results of the COACTIV research program]* (pp. 29–53). Waxmann. <https://doi.org/10.1007/s35834-011-0017-x>
- Baumert, J., & Kunter, M. (2013). The COACTIV model of teachers' professional competence. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss, & M. Neubrand (Eds.), *Cognitive activation in the mathematics classroom and professional competence of teachers. Results from the COACTIV project* (pp. 25–48). Springer. https://doi.org/10.1007/978-1-4614-5149-5_2.
- Becker, T. E., Atinc, G., Breaugh, J. A., Carlson, K. D., Edwards, J. R., & Spector, P. E. (2016). Statistical control in correlational studies: 10 essential recommendations for organizational researchers. *Journal of Organizational Behavior*, 37(2), 157–167. <https://doi.org/10.1002/job.2053>
- Black-Hawkins, K., & Florian, L. (2012). Classroom teachers' craft knowledge of their inclusive practice. *Teachers and Teaching*, 18(5), 567–584. <https://doi.org/10.1080/13540602.2012.709732>
- Blatchford, P., & Webster, R. (2018). Classroom contexts for learning at primary and secondary school: Class size, groupings, interactions and special educational needs. *British Educational Research Journal*, 44(4), 681–703.

- Bücker, S., Nuraydin, S., Simonsmeier, B. A., Schneider, M., & Luhmann, M. (2018). Subjective well-being and academic achievement: A meta-analysis. *Journal of Research in Personality*, 74, 83–94. <https://doi.org/10.1016/j.jrp.2018.02.007>
- Chen, K. (2006). Social skills intervention for students with emotional/behavioural disorders: A literature review from the American perspective. *Educational Research and Reviews*, 1(3), 143–149.
- Cobb, C. (2015). Principals play many parts: A review of the research on school principals as special education leaders 2001–2011. *International Journal of Inclusive Education*, 19(3), 213–234. <https://doi.org/10.1080/13603116.2014.916354>
- Conference of the Ministers of Education and Cultural Affairs. (2018). *Sonderpädagogische Förderung in allgemeinen Schulen (ohne Förderschulen) 2017/2018 [Special educational support in mainstream schools (excluding special schools) 2017/2018]*. https://www.kmk.org/fileadmin/Dateien/pdf/Statistik/Dokumentationen/Aus_SoPae_Int_2017.pdf
- Crede, J., Wirthwein, L., Steinmayr, R., & Bergold, S. (2019). Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf im Bereich emotionale und soziale Entwicklung und ihre Peers im inklusiven Unterricht [Students with special educational needs in the field of emotional and social development and their peers in inclusive classes]. *Zeitschrift für Pädagogische Psychologie*, 1–15. <https://doi.org/10.1024/1010-0652/a000244>
- Dalton, B. W. (2010). Antisocial and Prosocial Behavior. In J. A. Rosen, E. J. Glennie, B. W. Dalton, J. M. Lennon, & R. N. Bozick (Eds.), *Noncognitive skills in the classroom: New perspectives on educational research* (pp. 145–168). Research Triangle Park: RTI International. <https://doi.org/10.3768/rtipress.2010.bk.0000.1009>
- de Boer, A., Pijl, S. J., & Minnaert, A. (2011). Regular primary schoolteachers' attitudes towards inclusive education: A review of the literature. *International Journal of Inclusive Education*, 15(3), 331–353. <https://doi.org/10.1080/13603110903030089>
- Depaepe, F., Verschaffel, L., & Kelchtermans, G. (2013). Pedagogical content knowledge: A systematic review of the way in which the concept has pervaded mathematics educational research. *Teaching and Teacher Education*, 34, 12–25. <https://doi.org/10.1016/j.tate.2013.03.001>
- Enders, C. K., & Tofighi, D. (2007). Centering predictor variables in cross-sectional multilevel models: A new look at an old issue. *Psychological Methods*, 12(2), 121–138. <https://doi.org/10.1037/1082-989x.12.2.121>
- Engsig, T. T., & Johnstone, C. J. (2015). Is there something rotten in the state of Denmark? The paradoxical policies of inclusive education – lessons from Denmark. *International Journal of Inclusive Education*, 19(5), 469–486. <https://doi.org/10.1080/13603116.2014.940068>

- Epstein, M. H., & Cullinan, D. (1983). Academic performance of behaviorally disordered and learning-disabled pupils. *The Journal of Special Education*, 17(3), 303–307. <https://doi.org/10.1177/002246698301700306>
- Ramberg, J., Lénárt, A., & Watkins, A. (Eds.) (2018). *European Agency Statistics on Inclusive Education: 2016 Dataset Cross-Country Report*. European Agency for Special Needs and Inclusive Education. <https://www.european-agency.org/resources/publications/european-agency-statistics-inclusive-education-2016-dataset-cross-country>
- Farmer, T. W., McAuliffe Lines, M., & Hamm, J. V. (2011). Revealing the invisible hand: The role of teachers in children's peer experiences. *Journal of Applied Developmental Psychology*, 32(5), 247–256. <https://doi.org/10.1016/j.appdev.2011.04.006>
- Florian, L., & Black-Hawkins, K. (2011). Exploring inclusive pedagogy. *British Educational Research Journal*, 37(5), 813–828. <https://doi.org/10.1080/01411926.2010.501096>
- Florian, L., Black-Hawkins, K., & Rouse, M. (2016). *Achievement and inclusion in schools (2nd ed.)*. Routledge.
- Gebhardt, M., Sälzer, C., Mang, J., Müller, K., & Prenzel, M. (2015). Performance of students with special educational needs in Germany. Findings from PISA 2012. *Journal of Cognitive Education and Psychology*, 14(3), 343–356. <https://doi.org/10.1891/1945-8959.14.3.343>
- Gilmour, A. F., & Wehby, J. H. (2019). The association between teaching students with disabilities and teacher turnover. *Journal of Educational Psychology*, 1–19. <https://doi.org/10.1037/edu0000394>
- Göransson, K., & Nilholm, C. (2014). Conceptual diversities and empirical shortcomings—a critical analysis of research on inclusive education. *European Journal of Special Needs Education*, 29(3), 265–280. <https://doi.org/10.1080/08856257.2014.933545>
- Goodman, R. L., & Burton, D. M. (2010). The inclusion of students with BESD in mainstream schools: Teachers' experiences of and recommendations for creating a successful inclusive environment. *Emotional and Behavioural Difficulties*, 15(3), 223–237. <https://doi.org/10.1080/13632752.2010.497662>
- Grosche, M. (2015). Was ist Inklusion? Ein Diskussions- und Positionsartikel zur Definition von Inklusion aus Sicht der empirischen Bildungsforschung [What is inclusion? A discussion and position paper on the definition of inclusion from the perspective of empirical educational research]. In P. Kuhl, P. Stanat, B. Lütje-Klose, C. Gresch, H. A. Pant, & M. Prenzel (Eds.), *Inklusion von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf in Schulleistungserhebungen [Inclusion of pupils with special educational needs in school achievement surveys]* (pp. 17–39). Springer VS. <https://doi.org/10.1007/978-3-658-06604-8>

- Haffner, J., Baro, K., Parzer, P., & Resch, F. (2005). HRT 1-4: *Heidelberger Rechentest. Erfassung mathematischer Basiskompetenzen im Grundschulalter [HRT 1-4: Heidelberger Rechentest. Assessment of basic mathematical competencies at primary school age]*. Hogrefe.
- Helmke, A. (1988). Leistungssteigerung und Ausgleich von Leistungsunterschieden in Schulklassen: Unvereinbare Ziele? [Improving performance and compensating performance differences in school classes: Incompatible goals?]. *Zeitschrift für Entwicklungspsychologie und pädagogische Psychologie*, 20(1), 45–76. <https://doi.org/10.1026/0049-8637/a000062>
- Henke, T., Bosse, S., Lambrecht, J., J'antsch, C., Jaeuthe, J., & Spörer, N. (2017). Mittendrin oder nur dabei? Zum Zusammenhang zwischen sonderpädagogischem Förderbedarf und sozialer Partizipation von Grundschülerinnen und Grundschulern [In the middle or just thereby? On the relationship between special educational needs and social participation of primary school children]. *Zeitschrift für Pädagogische Psychologie*, 31(2), 111–123. <https://doi.org/10.1024/1010-0652/a000196>
- Henricsson, L., & Rydell, A. M. (2004). Elementary school children with behaviour problems: Teacher-child relations and self-perception. A prospective study. *Merrill-Palmer Quarterly*, 50(2), 111–138. <https://doi.org/10.1353/mpq.2004.0012>
- Heyder, A., Südkamp, A., & Steinmayr, R. (2020). How are teachers' attitudes toward inclusion related to the social-emotional school experiences of students with and without special educational needs? *Learning and Individual Differences*, 77, 101776. <https://doi.org/10.1016/j.lindif.2019.101776>
- Horne, P. E., & Timmons, V. (2009). Making it work: Teachers' perspectives on inclusion. *International Journal of Inclusive Education*, 13(3), 273–286. <https://doi.org/10.1080/13603110701433964>
- Huber, C., & Wilbert, J. (2012). Soziale Ausgrenzung von Schülern mit sonderpädagogischem Förderbedarf und niedrigen Schulleistungen im gemeinsamen Unterricht [Social exclusion of pupils with special educational needs and low school performance in inclusive classes]. *Empirische Sonderpädagogik*, 4(2), 147–165.
- Karsten, S., Peetsma, T., Roeleveld, J., & Vergeer, M. (2001). The Dutch policy of integration put to the test: Differences in academic and psychosocial development of pupils in special and mainstream education. *European Journal of Special Needs Education*, 16(3), 193–205. <https://doi.org/10.1080/08856250110074364>
- Kavale, K. A., & Forness, S. R. (1996). Social skill deficits and learning disabilities: A meta-analysis. *Journal of Learning Disabilities*, 29(3), 226–237. <https://doi.org/10.1177/002221949602900301>
- Klug, J., Bruder, S., Kelava, A., Spiel, C., & Schmitz, B. (2013). Diagnostic competence of

- teachers: A process model that accounts for diagnosing learning behavior tested by means of a case scenario. *Teaching and Teacher Education*, 30, 38–46. <https://doi.org/10.1016/j.tate.2012.10.004>
- Kocaj, A., Kuhl, P., Rjosk, C., Jansen, M., Pant, H. A., & Stanat, P. (2015). Der Zusammenhang zwischen Beschulungsart, Klassenkomposition und schulischen Kompetenzen von Kindern mit sonderpädagogischem Förderbedarf [The relationship between type of schooling, class composition and school competencies of children with special educational needs]. In P. Kuhl, P. Stanat, B. Lütje-Klose, C. Gresch, H. A. Pant, & M. Prenzel (Eds.), *Inklusion von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf in Schulleistungserhebungen [Inclusion of pupils with special educational needs in school achievement studies]* (pp. 335–370). Springer VS. https://doi.org/10.1007/978-3-658-06604-8_12
- König, J., & Blömeke, S. (2009). Pädagogisches Wissen von angehenden Lehrkräften [Pedagogical knowledge of pre-service teachers]. *Zeitschrift für Erziehungswissenschaft*, 12(3), 499–527. <https://doi.org/10.1007/s11618-009-0085-z>
- König, J., Gerhard, K., Melzer, C., Rühl, A. M., Zenner, J., & Kaspar, K. (2017). Erfassung von pädagogischem Wissen für inklusiven Unterricht bei angehenden Lehrkräften: Testkonstruktion und Validierung [Assessing general pedagogical knowledge for inclusive teaching (GPK-IT) among pre-service teachers: Test construction and validation. *Unterrichtswissenschaft*, 45(4), 223–242.
- König, J., & Pflanzl, B. (2016). Is teacher knowledge associated with performance? On the relationship between teachers' general pedagogical knowledge and instructional quality. *European Journal of Teacher Education*, 39(4), 419–436. <https://doi.org/10.1080/02619768.2016.1214128>
- Koster, M., Nakken, H., Pijl, S. J., & van Houten, E. (2009). Being part of the peer group: A literature study focusing on the social dimension of inclusion in education. *International Journal of Inclusive Education*, 13(2), 117–140. <https://doi.org/10.1080/13603110701284680>
- Kracke, B. (2014). Schulische Inklusion – Herausforderungen und Chancen [Inclusion in school – challenges and opportunities]. *Psychologische Rundschau*, 65(4), 237–240. <https://doi.org/10.1026/0033-3042/a000228>
- Krischler, M., & Pit-ten Cate, I. (2019). Pre-and in-service teachers' attitudes toward students with learning difficulties and challenging behavior. *Frontiers in Psychology*, 10, 327. <https://doi.org/10.3389/fpsyg.2019.00327>
- Kunz, A., Luder, R., & Moretti, M. (2010). Die Messung von Einstellungen zur Integration (EZI) [The measurement of attitudes towards integration (ATI)]. *Empirische Sonderpädagogik*, 2(3), 83–94.

- Kuyini, A. B., & Desai, I. (2007). Principals' and teachers' attitudes and knowledge of inclusive education as predictors of effective teaching practices in Ghana. *Journal of Research in Special Educational Needs*, 7(2), 104–113. <https://doi.org/10.1111/j.1471-3802.2007.00086.x>
- Kyriakides, L., Creemers, B., & Charalambous, E. (2018). *Equity and quality – dimensions in educational effectiveness*. Cham: Springer. https://doi.org/10.1007/978-3-319-72066-1_1
- Lauermann, F., & König, J. (2016). Teachers' professional competence and wellbeing: Understanding the links between general pedagogical knowledge, self-efficacy and burnout. *Learning and Instruction*, 45, 9–19. <https://doi.org/10.1016/j.learninstruc.2016.06.006>
- Lane, S., & Leventhal, B. (2015). Psychometric challenges in assessing English language learners and students with disabilities. *Review of Research in Education*, 39(1), 165–214. <https://doi.org/10.3102/0091732x14556073>
- Lenhard, W., Lenhard, A., & Schneider, W. (2017). *ELFE II: Ein Leseverständnistest für Erst- bis Siebtklässler – Version II [ELFE II: A reading comprehension test for first to 7th graders - Version II]*. Hogrefe.
- Loreman, T. (2007). Seven pillars of support for inclusive education. *International Journal of Whole Schooling*, 3, 22–38.
- Ministry for Schools and Education of North Rhine-Westphalia. (2016). *Verordnung über die sonderpädagogische Förderung, den Hausunterricht und die Schule für Kranke [Regulation on special educational support, home education and schools for ill persons]*. https://www.duesseldorf.de/fileadmin/Amt40/PDF/A-Z/AO_SF.pdf
- Mónico, P., Mensah, A. K., Grünke, M., Garcia, T., Fernández, E., & Rodríguez, C. (2018). Teacher knowledge and attitudes towards inclusion: A cross-cultural study in Ghana, Germany and Spain. *International Journal of Inclusive Education*, 22, 1–17. <https://doi.org/10.1080/13603116.2018.1471526>
- Muthén, L.K., & Muthén, B.O. (1998-2017). *Mplus User's Guide (8th ed.)*. Muthén & Muthén.
- Nelson, J. R., Benner, G. J., Lane, K., & Smith, B. W. (2004). Academic achievement of K-12 students with emotional and behavioral disorders. *Exceptional Children*, 71(1), 59–73. <https://doi.org/10.1177/001440290407100104>
- Newman, D. A. (2003). Longitudinal modeling with randomly and systematically missing data: A simulation of ad hoc, maximum likelihood, and multiple imputation techniques. *Organizational Research Methods*, 6(3), 328–362. <https://doi.org/10.1177/1094428103254673>
- Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory (3rd ed.)*. McGraw-Hill.

- Nye, E., Gardner, F., Hansford, L., Edwards, V., Hayes, R., & Ford, T. (2016). Classroom behaviour management strategies in response to problematic behaviours of primary school children with special educational needs: Views of special educational needs coordinators. *Emotional and Behavioural Difficulties*, 21(1), 43–60. <https://doi.org/10.1080/13632752.2015.1120048>
- OECD. (2014). *PISA 2012 results: What students know and can do – Student performance in mathematics, reading and science (Volume I, Revised edition)*. PISA, OECD Publishing. <https://doi.org/10.1787/efaa764e-en>
- Paju, B., Rätty, L., Pirttimaa, R., & Kontu, E. (2016). The school staff's perception of their ability to teach special educational needs pupils in inclusive settings in Finland. *International Journal of Inclusive Education*, 20(8), 801–815. <https://doi.org/10.1080/13603116.2015.1074731>
- Paulus, C. (2009). *Die „Bücheraufgabe“ zur Bestimmung des kulturellen Kapitals bei Grundschulern [The “book task” to determine the cultural capital of primary school children]*. http://bildungswissenschaften.uni-saarland.de/personal/paulus/homepage/files/BA_Artikel.pdf
- Rauer, W., & Schuck, K.-D. (2003). *FEESS 3-4: Fragebogen zur Erfassung emotionaler und sozialer Schulerfahrungen von Grundschulkindern dritter und vierter Klassen [FEESS 3-4: Questionnaire for assessing the emotional and social school experiences of primary school children in 3th and 4th grades]*. Hogrefe.
- Roy, A., Guay, F., & Valois, P. (2013). Teaching to address diverse learning needs: Development and validation of a differentiated instruction scale. *International Journal of Inclusive Education*, 17(11), 1186–1204. <https://doi.org/10.1080/13603116.2012.743604>
- Ruijs, N. M., Van der Veen, I., & Peetsma, T. T. (2010). Inclusive education and students without special educational needs. *Educational Research*, 52(4), 351–390.
- Sainio, P. J., Eklund, K. M., Ahonen, T. P., & Kiuru, N. H. (2019). The role of learning difficulties in adolescents' academic emotions and academic achievement. *Journal of Learning Disabilities*, 1–12. <https://doi.org/10.1177/0022219419841567>
- Savolainen, H., Engelbrecht, P., Nel, M., & Malinen, O. P. (2012). Understanding teachers' attitudes and self-efficacy in inclusive education: Implications for preservice and in-service teacher education. *European Journal of Special Needs Education*, 27(1), 51–68. <https://doi.org/10.1080/08856257.2011.613603>
- Schmiedeler, S. (2013). Wissen und Fehlannahmen von deutschen Lehrkräften über die Aufmerksamkeitsdefizit-/Hyperaktivitätsstörung (ADHS) [German teachers' knowledge and misconceptions about Attention-Deficit/Hyperactivity Disorder (ADHD)]. *Psychologie in Erziehung und Unterricht*, 60, 143–153.
- Schuck, K. D., & Rauer, W. (2018). II Teilprojekt 1: “Emotional-soziale Schulerfahrungen”

- der Grundschul Kinder [II Subproject 1: “Emotional-social school experiences” of primary school children]. In K. D. Schuck, W. Rauer, & D. Prinz (Eds.), *EiBiSch – Evaluation inklusiver Bildung in Hamburgs Schulen [EiBiSch – Evaluation of inclusive education in schools in Hamburg]* (pp. 26–228). Waxmann.
- Schwab, S. (2015). Social dimensions of inclusion in education of 4th and 7th grade pupils in inclusive and regular classes: Outcomes from Austria. *Research in Developmental Disabilities, 43*, 72–79. <https://doi.org/10.1016/j.ridd.2015.06.005>
- Schwarzer, R., & Jerusalem, M. (1999). *Skalen zur Erfassung von Lehrer- und Schülermerkmalen. Dokumentation der psychometrischen Verfahren im Rahmen der Wissenschaftlichen Begleitung des Modellversuchs Selbstwirksame Schulen [Scales for measuring teacher and student characteristics. Documentation of the psychometric procedures within the framework of the scientific monitoring of the pilot project Self-Effective Schools]*. Freie Universität Berlin.
- Shifrer, D. (2013). Stigma of a label: Educational expectations for high school students labeled with learning disabilities. *Journal of Health and Social Behavior, 54*, 462–480. <https://doi.org/10.1177/0022146513503346>
- Shifrer, D., Muller, C., & Callahan, R. (2011). Disproportionality and learning disabilities: Parsing apart race, socioeconomic status, and language. *Journal of Learning Disabilities, 44*(3), 246–257.
- Smeets, E. (2009). Managing social, emotional and behavioural difficulties in schools in the Netherlands. *The International Journal of Emotional Education, 1*(1), 50–63.
- Sousa, P., Dias, P. C., & Cadime, I. (2017). Predictors of primary school teachers’ knowledge about developmental dyscalculia. *European Journal of Special Needs Education, 32*(2), 204–220. <https://doi.org/10.1080/08856257.2016.1216635>
- Strand, S., & Lindsay, G. (2009). Evidence of ethnic disproportionality in special education in an English population. *The Journal of Special Education, 43*(3), 174–190. <https://doi.org/10.1177/0022466908320461>
- Thomm, E., Gold, B., Betsch, T., & Bauer, J. (2021). When preservice teachers’ prior beliefs contradict evidence from educational research. *British Journal of Educational Psychology, e12407*. <https://doi.org/10.1111/bjep.12407>
- Tometten, L., Heyder, A., & Steinmayr, R. (2022). Empirische Arbeit: Die Erfassung von Wissen über verschiedene Diversitätsbereiche im Kontext einer inklusionsorientierten Lehrkräftebildung [The assessment of knowledge about diversity in the context of an inclusion-oriented teacher training]. *Psychologie in Erziehung und Unterricht, 69*(4), 292–304. <https://doi.org/10.2378/peu2022.art23d>
- United Nations Educational, Scientific and Cultural Organization. (1994). *The Salamanca statement and framework for action on special needs education*.

<https://unesdoc.unesco.org/ark:/48223/pf0000098427>

- Vaughn, S., Elbaum, B., & Boardman, A. G. (2001). The social functioning of students with learning disabilities: Implications for inclusion. *Exceptionality*, 9(1–2), 47–65. https://doi.org/10.1207/s15327035ex091&2_5
- von Stumm, S. (2014). Intelligence, gender, and assessment method affect the accuracy of self-estimated intelligence. *British Journal of Psychology*, 105(2), 243–253. <https://doi.org/10.1111/bjop.12031>
- Voss, T., Kunina-Habenicht, O., Hoehne, V., & Kunter, M. (2015). Stichwort Pädagogisches Wissen von Lehrkräften: Empirische Zugänge und Befunde [Keyword teachers' pedagogical knowledge: Empirical approaches and findings]. *Zeitschrift für Erziehungswissenschaft*, 18(2), 187–223. <https://doi.org/10.1007/s11618-015-0626-6>
- Voss, T., Kunter, M., & Baumert, J. (2011). Assessing teacher candidates' general pedagogical/psychological knowledge: Test construction and validation. *Journal of educational psychology*, 103(4), 952–969. <https://doi.org/10.1037/a0025125>
- Woodcock, S., Hitches, E., & Jones, G. (2019). It's not you, it's me: Teachers' self-efficacy and attributional beliefs towards students with specific learning difficulties. *International Journal of Educational Research*, 97, 107–118. <https://doi.org/10.1016/j.ijer.2019.07.007>
- World Health Organization. (2016). *International Statistical Classification of Diseases and Related Health Problems* (10th rev. ed). <https://icd.who.int/browse10/2016/en>

Table 1. Mean (*M*), standard deviations (*SD*) and bivariate correlations for all level-1-variables.

Variables	Descriptive statistics				Correlations									
	Subsample 1		Subsample 2		Both Subsamples ($N_1 N_2$)									
	(EBD)		(LD)		1	2	3	4	5	6	7	8	9	10
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>										
1 SEN EBD ^a	0.08	0.28	0.11	0.31	-	.31***	.08*	.06	-.01	-.01	-.06	-.07	-.14***	-.11**
2 SEN LD ^a	0.02	0.14	0.01	0.11	.46***	-	.00	.16***	-.17***	.09*	-.04	-.10*	-.36***	-.40***
3 Gender ^a	0.49	0.50	0.48	0.50	.19***	.11*	-	.01	-.03	.03	-.04	-.07*	-.01	.17***
4 Age	9.65	0.65	9.62	0.66	.06	.07	-.05	-	-.13**	.17***	-.15***	-.13***	-.05	-.06
5 SES	3.29	1.26	3.28	1.28	.00	-.02	-.02	-.11*	-	-.29***	.13**	.16***	.34***	.22***
6 Migration ^a	0.59	0.49	0.54	0.50	-.08	-.02	-.06	.14**	-.25***	-	-.21***	-.12**	-.31***	-.10**
7 Class climate	3.50	0.68	3.65	0.67	-.07	-.06	-.02	-.17***	.15**	-.11*	-	.68***	.15***	.11**
8 Social integration	3.78	0.73	3.88	0.70	-.18***	-.07	-.10	-.07	.15**	.00	.67***	-	.20***	.24***
9 Reading comprehension	0.61	0.24	0.59	0.24	-.12*	-.19***	-.03	-.09	.35***	-.35***	.15**	.16**	-	.55***
10 Basic math competencies	0.52	0.15	0.50	0.15	-.17**	-.18***	.16**	-.02	.13**	-.10*	.18***	.26***	.46***	-

Note: $N_1 = 410-421$, $N_2 = 696-715$. SEN = Special educational needs, EBD = Emotional and behavioral difficulties, LD = Learning difficulties, SES = Socio-economic status. Correlations below the diagonal indicate the results in Subsample 1 and correlations above the diagonal indicate the results in Subsample 2. ^a These variables are dichotomous (coded with 0 and 1). The mean value indicates the proportion of students with SEN EBD/LD, boys and students with a migrant background.

* $p < .05$, ** $p < .01$, *** $p < .001$.

Table 2. Mean (*M*), standard deviations (*SD*) and bivariate correlations for all level-2-variables.

Variables	Descriptive statistics				Correlations							
	Subsample 1 (EBD)		Subsample 2 (LD)		Both Subsamples ($N_1 \setminus N_2$)							
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	1	2	3	4	5	6	7	8
Teacher variables												
1 Years of teaching experience	15.41	9.46	12.72	8.68	-	-.01	-.02	.32	.15	.24	-.45*	.03
2 ATI - Academic support	2.36	0.76	2.62	0.84	.16	-	.50**	.30	-.26	.04	.17	.20
3 ATI - Social integration	4.10	0.76	4.11	0.68	.22	.60**	-	.30	-.32	-.01	.30	.06
4 Self-efficacy	3.59	0.71	3.63	0.69	.18	.51*	.47*	-	.20	.19	.04	.17
5 Knowledge EBD	0.73	0.15	0.72	0.16	-.08	-.07	-.17	.14	-	.30	-.12	.05
6 Knowledge LD	0.60	0.16	0.60	0.17	.31	-.14	.06	.15	.28	-	-.02	.30
Class variables												
7 Class size	23.08	3.26	24.21	3.17	-.30	.21	.31	.04	-.12	-.01	-	.23
8 Number of students with SEN per class	5.46	2.75	5.64	2.23	.03	.09	.04	.40	.12	.18	.04	-

Note: $N_1 = 23$, $N_2 = 36$. ATI = Attitudes towards inclusion, EBD = Emotional and behavioral difficulties, LD = Learning difficulties, SEN = Special educational needs. Correlations below the diagonal indicate the results in Subsample 1 and correlations above the diagonal indicate the results in Subsample 2.

* $p < .05$, ** $p < .01$, *** $p < .001$.

Table 3. Results of multilevel models with standardized regression weights (β) predicting students' social participation and academic achievement in Subsample 1 (EBD).

Level and variable	Social participation				Academic achievement			
	Classroom climate		Social integration		Reading comprehension		Basic math competencies	
	Step 1	Step 2	Step 1	Step 2	Step 1	Step 2	Step 1	Step 2
Intercept	3.56***	3.55***	3.87***	3.86***	0.70***	0.70***	0.51***	0.51***
Level 1								
SEN EBD	-0.16	-0.15	-0.52**	-0.52**	-0.05	-0.05	-0.08**	-0.08**
SEN LD	-0.06	-.008	0.21	0.19	-0.29***	-0.32***	-0.11*	-0.12*
Gender	-0.05	-0.04	-0.12	-0.12	-0.01	-0.01	0.06***	0.06***
Age	-0.11**	-0.12**	-0.05	-0.06	-0.02	-0.01	-0.01	-0.01
SES	0.05*	0.05	0.09**	0.08**	0.05***	0.05***	0.01**	0.01**
Migration	-0.01	0.00	0.03	0.05	-0.13***	-0.13***	-0.02	-0.02
Level 2								
Knowledge EBD	-	0.92*	-	0.76**	-	0.23*	-	0.16*
Knowledge LD	-	-.11	-	0.29	-	0.12	-	0.17
Years of teaching experience	-	0.00	-	0.00	-	0.00	-	0.00
ATI - Academic support	-	-0.09	-	0.10	-	0.01	-	0.02

ATI - Social integration	-	0.13	-	-0.09	-	0.04	-	0.03
Self-efficacy	-	0.07	-	0.06	-	-0.06*	-	-0.03
Class size	-	-0.04	-	0.01	-	-0.01	-	-0.01
Number of students with SEN per class	-	-0.04	-	-0.01	-	0.00	-	0.00
Variance Slope SEN EBD	-	0.02	-	0.20	-	0.02	-	0.00

Note: $N = 414$. Step 1 = Random-intercept-fixed-slope-models with level-1-predictors. Step 2 = Random-intercept-random-slope-models with level-1-and-2-predictors. SEN = Special educational needs, EBD = Emotional and behavioral difficulties, LD = Learning difficulties, SES = Socio-economic status, ATI = Attitudes towards inclusion. If other results appear when the control variables are omitted, the corresponding values are highlighted in gray.

* $p < .05$, ** $p < .01$, *** $p < .001$.

Table 4. Results of multilevel models with standardized regression weights (β) predicting students' social participation and academic achievement in Subsample 2 (LD).

Level and variable	Social participation				Academic achievement			
	Classroom climate		Social integration		Reading comprehension		Basic math competencies	
	Step 1	Step 2	Step 1	Step 2	Step 1	Step 2	Step 1	Step 2
Intercept	3.72***	3.72***	3.97***	3.98***	0.67***	0.67***	0.49***	0.49***
Level 1								
SEN LD	-0.03	-0.03	-0.14	-0.13	-0.24***	-0.24***	-0.19***	-0.19***
SEN EBD	-0.24	-0.23	-0.20	-0.18	-0.10*	-0.10	0.00	0.00
Gender	-0.08	-0.08	-0.12*	-0.13*	0.00	0.00	0.05***	0.05***
Age	-0.01	-0.01	-0.05	-0.6	0.01	0.01	-0.01	-0.01
SES	0.02	0.03	0.06*	0.06*	0.03***	0.03***	0.01**	0.01**
Migration	-0.05	-0.05	-0.03	-0.03	-0.10***	-0.09***	0.00	0.00
Level 2								
Knowledge LD	-	0.11	-	0.13	-	0.05	-	0.01
Knowledge EBD	-	0.63	-	0.22	-	0.17	-	0.17*
Years of teaching experience	-	0.00	-	-0.01	-	0.00	-	0.00
ATI - Academic support	-	0.18**	-	0.14*	-	0.02	-	0.05**

ATI - Social integration	-	-0.19	-	-0.13	-	0.03	-	0.02
Self-efficacy	-	0.03	-	0.05	-	-0.07*	-	-0.03
Class size	-	0.02	-	0.01	-	0.00	-	-0.01
Number of students with SEN per class	-	-0.03	-	0.00	-	0.00	-	0.00
Variance Slope SEN LD	-	0.01	-	0.08	-	0.00	-	0.00

Note: $N = 703$. Step 1 = Random-intercept-fixed-slope-models with level-1-predictors. Step 2 = Random-intercept-random-slope-models with level-1-and-2-predictors. SEN = Special educational needs, LD = Learning difficulties, EBD = Emotional and behavioral difficulties, SES = Socio-economic status, ATI = Attitudes towards inclusion. If other results appear when the control variables are omitted, the corresponding values are highlighted in gray.

* $p < .05$, ** $p < .01$, *** $p < .001$.

Appendix

The test DiWi-K measures declarative factual knowledge about different diversity domains. So far, a short and a long version have been evaluated in three independent preservice teacher samples (Tometten et al., 2022). Various reliability and validity indicators were provided: As expected, knowledge was positively related to high school GPA, self-reported knowledge, diversity aspects covered in university courses, experiences with diversity, attitudes, and self-efficacy in the context of inclusive education. The overall test of both versions had good split-half reliability, retest reliability, and internal consistency scores. The two short version subscales EBD and LD had slightly lower internal consistency but correlated strongly ($r = .85/.89$) with the two respective subscales of the long version.

Items of the EBD subscale:

- Children with special educational needs in the area of EBD have at least one diagnosed mental disorder. (*wrong*)
- Appreciative remarks about desired behavior lead to excessive demands on students with SEN EBD. (*wrong*)
- Anxiety requires treatment if it impairs the development and everyday life of the child or adolescent. (*right*)
- Strongly varying performance, daydreaming or physical clumsiness can indicate ADHD. (*right*)
- In most cases, ADHD is accompanied by a below-average IQ. (*wrong*)
- The cause of the deterioration in concentration and impulse control in ADHD is an imbalance in the neurotransmitters dopamine and norepinephrine. (*right*)
- Students with SEN EBD are taught towards different goals than students without SEN EBD. (*wrong*)

Items of the LD subscale:

- A learning disability is present throughout the entire school life. (*wrong*)
- In order for special educational needs in the area of LD to be identified, there must be serious, extensive and long-term learning and performance deficits. (*right*)
- All students with dyslexia or dyscalculia are considered to have special educational needs in the area of LD. (*wrong*)
- Children with learning disabilities are more often affected by psychosocial problems than average. (*right*)
- The probability of a learning impairment increases with decreasing socio-economic status. (*right*)
- Students with educational needs in the area of LD are below average intelligence as defined by school standards. (*wrong*)
- Students with educational needs in the area of LD are uniformly weak in all academic areas. (*wrong*)

6 Gesamtdiskussion

Übergeordnetes Ziel dieser Arbeit war die Evaluation eines Tests zum Wissen über verschiedene Diversitätsbereiche von (angehenden) Lehrkräften und damit einhergehend die Erforschung des Diversitätswissens im Kontext schulischer Inklusion. Aus dem COACTIV-Modell zur professionellen Kompetenz von Lehrkräften (Baumert & Kunter, 2013a; Kunter, Kleickmann et al., 2013), bisherigen Tests zum pädagogisch-psychologischen Wissen bzw. Wissen über einzelne Diversitätsbereiche (Heyder, Bergold & Steinmayr, 2018; Kunina-Habenicht et al., 2020) und empirischen Zusammenhängen zwischen dem Wissen und Lerngelegenheiten (Sanz-Cervera et al., 2017), Überzeugungen (Lu et al., 2020), der Unterrichtsperformanz (Gess-Newsome et al., 2017) und dem Berufserfolg von Lehrkräften (Dicke et al., 2015) wurden die folgenden drei Fragestellungen abgeleitet und im Rahmen dieser Arbeit untersucht:

- 1) Welches Modell (1-Faktor-Modell vs. hierarchisches 5-Faktoren-Modell mit Generalfaktor) kann die Struktur des Wissens über die fünf Diversitätsbereiche am besten abbilden?
- 2) Wie hängen formale Lerngelegenheiten, Aspekte professioneller Kompetenz (berufliche Überzeugungen, Facetten des Professionswissens) und die Unterrichtsperformanz mit dem Diversitätswissen zusammen?
- 3) Wie hängt das Diversitätswissen mit dem Unterrichtserfolg auf Ebene der Schüler:innen zusammen?

Die zentralen Ergebnisse der drei Studien werden vor dem Hintergrund dieser Fragestellungen in Kapitel 6.1 zunächst zusammengefasst. Es folgen die Interpretation der Ergebnisse zum Wissen und zu den Fehlannahmen in Bezug auf die professionelle Kompetenz von Lehrkräften und den hier vorgestellten Test (Kapitel 6.2), eine Darlegung der

Stärken und Grenzen der Arbeit (Kapitel 6.3) und Implikationen für Forschung und Praxis (Kapitel 6.4). Die Arbeit schließt mit einem studienübergreifenden Fazit (Kapitel 6.5).

6.1 Zusammenfassung der Studienergebnisse

In Studie 1 (*Die Erfassung von Wissen über verschiedene Diversitätsbereiche im Kontext einer inklusionsorientierten Lehrkräftebildung*) wurde der Wissenstest in einer Langversion (DiWi) und einer Kurzversion (DiWi-K) vorgestellt und in drei unabhängigen Studierendenstichproben evaluiert. In Anlehnung an bisherige Wissenstests (Horn et al., 2004; Kunina-Habenicht et al., 2020) wurde eine 5-faktorielle Struktur anhand der inhaltlichen Bereiche SPF ESE, SPF Lernen, Teilleistungsstörungen, Geschlecht und Hochbegabung mit einem Generalfaktor, dem Diversitätswissen, angenommen. Als Reliabilitätsschätzer wurden McDonalds Omega und die Retest-Reliabilität über einen Zeitraum von 3 Monaten bestimmt. Für die Sammlung von Hinweisen auf die Testwertinterpretationen im Sinne eines nomologischen Netzes (Hartig et al., 2020) wurde das COACTIV-Modell zur professionellen Kompetenz von Lehrkräften (Baumert & Kunter, 2013a; Kunter, Kleickmann et al., 2013) herangezogen. In beiden Testversionen zeigte das angenommene hierarchische 5-Faktoren-Modell im Gegensatz zum 1-Faktor-Modell einen guten Modellfit und erlaubt somit eine Berechnung von Subskalen- und Gesamtwerten (Forschungsfrage 1). McDonalds Omega war gut bis sehr gut, und es zeigte sich eine mittlere Retest-Reliabilität. Es ließ sich ein nomologisches Netz aus Lerngelegenheiten (Abiturnote, Diversität in Lehrveranstaltungen, Vorerfahrungen mit Diversität, Studienfach, Studienabschnitt), beruflichen Überzeugungen (Einstellungen, Selbstwirksamkeitserwartungen) und weiteren Facetten des Professionswissens (selbsteingeschätztes Wissen, pädagogischen Unterrichtswissen) zeichnen (Forschungsfrage 2). Diese Studie bietet die Basis, damit der Test zukünftig in Forschung und Praxis eingesetzt werden kann.

Da die Ergebnisse nicht automatisch auf in der Praxis tätige Lehrkräfte übertragen werden können, wurde der DiWi in Studie 2 (*Teachers' knowledge about diversity: Correlations with attitudes, self-efficacy, and instructions*) in einer Stichprobe aus Lehrkräften erprobt. In dieser Studie konnte das in Studie 1 aufgestellte nomologische Netz bestätigt werden (Forschungsfrage 2). Erfasst wurden dazu der berufliche Kontakt zu Schüler:innen verschiedener Diversitätsmerkmale, das Studienfach, die Einstellungen und die Selbstwirksamkeitserwartungen. Außerdem wurden neue Variablen hinzugenommen: Das Wissen der Lehrkräfte korrelierte positiv mit den lernorientierten Unterrichtsinstruktionen, die sich zum Beispiel durch die Anpassung von Aufgaben an die individuellen Bedürfnisse der Schüler:innen auszeichnen (Retelsdorf et al., 2010). Außerdem wiesen die Lehrkräfte ein höheres Wissen auf als die hinzugezogenen Studierenden-Vergleichsstichproben (aus Studie 1.1 und 1.2). Nicht signifikant waren eine tendenziell negative Korrelation zwischen dem Wissen und den leistungsorientierten Unterrichtsinstruktionen und der Vergleich zwischen Lehrkräften, die kein vs. mindestens ein Training im Bereich Inklusion absolviert hatten. Die Analysen wurden ebenfalls mit den Fehlannahmen berechnet: Hier zeigten sich nur wenige signifikante Ergebnisse. Die Fehlannahmen korrelierten positiv mit dem beruflichen Kontakt zu Schüler:innen verschiedener Diversitätsmerkmale, und die Lehrkräfte hatten mehr Fehlannahmen als die Studierenden der Vergleichsstichproben. Subskalenspezifische Zusatzanalysen bestätigten mit wenigen Ausnahmen die Ergebnisse in Bezug auf das Gesamtwissen und die Gesamtfehlannahmen. Bedeutsame Unterschiede zwischen Lehrkräften verschiedener Schulformen konnten nicht gefunden werden. Zusammenfassend konnte diese Studie das nomologische Netz auch in Bezug auf eine Lehrkräftestichprobe bestätigen. Gleichzeitig werfen die Ergebnisse neue Erkenntnisse und Forschungsfragen nicht nur in Bezug auf das Wissen, sondern auch die Fehlannahmen auf.

Vor dem Hintergrund des Zusammenhangs zwischen dem Wissen und den

lernorientierten Unterrichtsinstruktionen sollte in Studie 3 (*Links between teachers' knowledge about special educational needs and students' social participation and academic achievement in mainstream classes*) geprüft werden, ob das Wissen der Lehrkräfte über die beiden SPF ESE und Lernen ebenfalls mit dem Unterrichtserfolg auf Ebene der Schüler:innen in Verbindung gebracht werden kann (Forschungsfrage 3). In dem Wissen, dass die beiden Förderschwerpunkte nur einen Teil schulischer Inklusion darstellen (vgl. Kapitel 1.1 zur Definition schulischer Inklusion), erschien eine Reduktion der Testinhalte auf Basis der hohen Prävalenz dieser Schüler:innen (KMK, 2022) und der Wichtigkeit, der diesen Förderschwerpunkten damit in der Praxis zukommt, nötig und sinnvoll für eine ökonomische Umsetzung der Studie. Untersucht wurden Schüler:innen mit und ohne SPF ESE (Stichprobe 1) bzw. Lernen (Stichprobe 2) in gemeinsamen Klassen und ihre Lehrkräfte. Die Analysen ergaben geringere Leistungen von Schüler:innen mit SPF im Vergleich zu Schüler:innen ohne SPF (unabhängig ob ESE oder Lernen). Dies galt auch für die soziale Integration von Schüler:innen mit SPF ESE im Vergleich zu ihren Peers. Je höher das Wissen der Lehrkräfte über den SPF ESE war, desto größer waren das Leseverständnis und die basalen Mathematik-Kompetenzen (= akademische Leistungen) sowie die empfundene soziale Integration und das Klassenklima (= soziale Partizipation) aller Schüler:innen in Klassen mit mindestens einem Kind mit SPF ESE. Diese Zusammenhänge zeigten sich nicht für das Wissen über den SPF Lernen. Auch zeigten sich keine signifikanten Interaktionseffekte zwischen dem Förderbedarf und dem Wissen. Die Unterschiede zwischen Schüler:innen mit und ohne SPF variierten dementsprechend nicht signifikant in Abhängigkeit des Lehrkräftewissens.

6.2 Interpretation der Ergebnisse zum Wissen und zu den Fehlannahmen über Diversität

Die einzelnen Studien werden in den jeweiligen Abschnitten in Kapitel 5 diskutiert.

Aus den Ergebnissen der drei Studien lassen sich in Ergänzung dazu Aspekte zur Einordnung in die professionelle Kompetenz und zur Erfassung und Struktur ableiten und zusammenfassen, die im Folgenden dargestellt werden.

6.2.1 Einordnung in die professionelle Kompetenz

In Anlehnung an die Darstellung empirischer Befunde in Kapitel 2.2 werden im Folgenden (1) formale Lerngelegenheiten, (2) ausgewählte Aspekte der professionellen Kompetenz und (3) die Performanz und der Unterrichtserfolg in den Blick genommen.

Bedeutung formaler Lerngelegenheiten. Nach dem COACTIV-Kompetenzmodell (Kunter, Kleickmann et al., 2013) haben formale Lerngelegenheiten einen Einfluss auf das Professionswissen. Dementsprechend konnten Zusammenhänge zwischen verschiedenen Lerngelegenheiten und Wissensfacetten bereits in vielen Studien nachgewiesen werden (Heyder, Vaskova et al., 2018; Hohenstein et al., 2017; König et al., 2017; Sanz-Cervera et al., 2017; Scitutto et al., 2000; Sousa et al., 2017; Wildbrett et al., 2022; Yin et al., 2019). Diese positiven Zusammenhänge konnten nun auch für das Diversitätswissen operationalisiert über fünf Diversitätsbereiche in drei unabhängigen Studien gezeigt werden.

Es zeigten sich wenige Ausnahmen, in denen kein Zusammenhang zwischen den Lerngelegenheiten und dem Diversitätswissen bzw. dem Wissen zu einzelnen Diversitätsbereichen gefunden wurde: Der Vergleich von Lehrkräften verschiedener Schulformen ergab, dass die wenigen Unterschiede zugunsten der Lehrkräfte an Grundschulen nicht auf die Schulform an sich sondern auf andere Hintergrundvariablen zurückzuführen sind (Studie 2). Darüber hinaus unterschieden sich Lehrkräfte, die kein vs. mindestens ein zusätzliches Training zu Inklusion besucht hatten, entgegen den Ergebnissen vorheriger Studien zu jeweils einem spezifischen Diversitätsbereich (Sanz-Cervera et al., 2017; Schmiedeler, 2013; Yin et al., 2019), nicht signifikant in ihrem Wissen voneinander. Dies könnte daran liegen, dass sich Trainings stark in Inhalt und Lernaktivitäten

unterscheiden (Kurniawati et al., 2014) und dass das zusätzliche Training nur global erfasst wurde. Dementsprechend bestand nicht notwendigerweise eine inhaltliche Nähe zum Diversitätswissen. Implikationen, die aus dieser Arbeit für die Aus- und Weiterbildung von (angehenden) Lehrkräften zum Beispiel im Rahmen von Trainings abgeleitet werden können, werden in Kapitel 6.4.2 diskutiert.

Lehrkräfte für sonderpädagogische Förderung hatten nur in Bezug auf die beiden SEN und Teilleistungsstörungen ein höheres Wissen als Lehrkräfte anderer Studiengänge (ebenfalls Studie 2). Dieses Muster erscheint plausibel, da die beiden anderen Dimensionen Geschlecht und Hochbegabung nicht direkt mit einem besonderen Unterstützungsbedarf verknüpft sind und die Lehrkräfte daher möglicherweise nicht besonders gut in diesen Bereichen ausgebildet sind. Die Lehrkräfte hatten zwar ein höheres Wissen als die Bachelor- und Master-Studierenden, in Bezug auf die Master-Studierenden galt dies allerdings nicht für die Skala Geschlecht. Das könnte dadurch erklärt werden, dass Lehrkräfte im beruflichen Alltag nicht explizit mit diesem Thema konfrontiert werden und das entsprechende Wissen vor allem im Studium vermittelt wird, da es sich hier im Vergleich zu den anderen Skalen nicht um eine Diversitätsdimension mit Diagnose handelt. Die Lehrkräfte müssten sich also darüber hinaus eigenständig aktiv mit dem Thema Geschlecht auseinandersetzen, um weitere Lerngelegenheiten zu erhalten.

Die tendenziell positive Korrelation der Lehrerfahrung in Jahren von Lehrkräften in Klassen mit Schüler:innen mit und ohne SPF ESE bzw. Lernen mit ihrem Wissen über den SPF ESE bzw. Lernen wurde in den kleinen Stichproben nicht signifikant (Studie 3). Befunde zu einem kurvilinearen Zusammenhang zwischen der Lehrerfahrung und dem GPK von Lehrkräften an Grundschulen und weiterführenden Schulen (höheres Wissen bei moderater Lehrerfahrung, niedrigeres Wissen bei besonders wenig und viel Lehrerfahrung; Lauermann & König, 2016) deuten darauf hin, dass auch der Zusammenhang zwischen der

Lehrerfahrung und dem Diversitätswissen komplexer sein könnte: Zur Studienzeit der Lehrkräfte, die zum Erhebungszeitpunkt viel Lehrerfahrung hatten, könnte Inklusion einen weniger hohen Stellenwert in der curricularen Ausbildung gehabt haben (vgl. Salamanca Erklärung, UNESCO, 1994; UN-Behindertenrechtskonvention, Beauftragter der Bundesregierung für die Belange von Menschen mit Behinderungen, 2018) als zur Studienzeit der Lehrkräfte, die zum Erhebungszeitpunkt moderate Lehrerfahrung hatten. Lehrkräfte mit besonders wenig Lehrerfahrung hingegen hatten im Vergleich zu den beiden anderen Gruppen die wenigsten Lerngelegenheiten und waren weniger Situationen ausgesetzt, in denen sie etwas über Diversität lernen konnten. Ob es tatsächlich einen kurvilinearen Zusammenhang zwischen der Lehrerfahrung und dem Diversitätswissen gibt und welche Mechanismen ihm zugrunde liegen, bleibt für zukünftige Forschung offen. Damit einhergehend bleibt offen, welche praktischen Lerngelegenheiten im Beruf mit dem Diversitätswissen zusammenhängen.

Zusammenfassend deutet die Gesamtheit der Studien auf positive Zusammenhänge zwischen formalen Lerngelegenheiten und dem Diversitätswissen hin. Da auf Basis des querschnittlichen Studiendesigns keine kausalen Schlüsse gezogen werden, kann über verschiedene Wirkrichtungen nur spekuliert werden: So ist es zwar plausibel, dass Lerngelegenheiten das Wissen positiv beeinflussen, jedoch könnte es auch sein, dass Personen mit einem hohen Wissen motiviert sind, weitere Lerngelegenheiten aufzusuchen und aktiv für den eigenen Wissenszuwachs zu nutzen. Außerdem ist nicht klar, wie die praktischen und theoretischen Lerngelegenheiten miteinander interagieren. So ist es zum Beispiel denkbar, dass Personen, die viel privaten Kontakt zu gewissen Diversitätsmerkmalen haben, eher eine Lernumgebung suchen, in der sie ihr Wissen darüber erweitern können, wohingegen Personen mit wenig Kontakt keine Notwendigkeit dafür sehen. Andersherum könnte es sein, dass Personen, die bereits viel über gewisse Diversitätsmerkmale wissen, auch

eher Situationen aufsuchen, in denen sie mit Menschen mit bestimmten Diversitätsmerkmalen in Kontakt kommen, zum Beispiel weil sie weniger Berührungsängste haben. Gleichzeitig ist es denkbar, dass das Wissen beeinflusst, wie Personen den Kontakt zu Menschen mit gewissen Diversitätsmerkmalen wahrnehmen. Befassen sich Personen mit dem Thema Hochbegabung, erkennen bzw. erfragen sie vielleicht eher, ob jemand in ihrem Umfeld hochbegabt ist. Im beruflichen Umfeld erkennen Personen mit einem hohen Wissen unter Umständen eher, dass eine Diagnostik vorgenommen werden könnte bzw. sollte.

Im Gegensatz zum Wissen, mit dem die Lerngelegenheiten durchweg positiv zusammenhängen, bleibt die Einordnung der Fehlannahmen deutlich diffuser. Um die Bedeutung von Lerngelegenheiten für die Entwicklung von Fehlannahmen einschätzen zu können, bedarf es weiterer Studien zu Hintergründen von Fehlannahmen über Diversität. In Studie 2 hatten Lehrkräfte mehr Fehlannahmen als Bachelor- und Master-Studierende, und der berufliche Kontakt zu Schüler:innen mit den erfragten Diversitätshintergründen korrelierte positiv mit den Fehlannahmen. Damit zeigten sich meistens keine und vereinzelt positive Zusammenhänge zwischen den Lerngelegenheiten und den Fehlannahmen, zu denen nach aktueller Recherche nur eine Studie mit uneindeutigen Befunden vorliegt (Sanz-Cervera et al., 2017). Methodische und inhaltliche Überlegungen, wie diese Zusammenhänge zustande kamen, werden in Kapitel 5.3 zu Studie 2 ausführlich diskutiert. Aus diesen Ergebnissen können zusammenfassend keine eindeutigen Schlüsse gezogen werden. Um detailliertere Aussagen treffen zu können, scheinen Studien zu bereichsspezifischen Variablen sinnvoll. Für diesen Zweck könnten die fünf Diversitätsbereiche des hier vorgestellten Tests genutzt werden.

Ausgewählte Aspekte professioneller Kompetenz. Im COACTIV-Kompetenzmodell (Baumert & Kunter, 2013a; Kunter, Kleickmann et al., 2013) werden verschiedene Aspekte professioneller Kompetenz genannt. Als gemeinsamen Nenner haben diese individuelle

Eingangsvoraussetzungen und formale Lerngelegenheiten, jedoch wird nicht definiert, wie die verschiedenen Aspekte miteinander konkret zusammenhängen. In den Studien 1 und 2 konnten positive Zusammenhänge zwischen dem Diversitätswissen der (angehenden) Lehrkräfte mit dem selbsteingeschätzten Wissen, dem pädagogischen Unterrichtswissen, den Einstellungen zu Inklusion und den Selbstwirksamkeitserwartungen in der Gestaltung inklusiven Unterrichts bzw. für das Unterrichten an einer inklusiven Schule aufgezeigt werden. Damit knüpft diese Arbeit an vergleichbare Befunde vorheriger Studien an (Ghanizadeh et al., 2006; König et al., 2017; Lu et al., 2020; Sciutto et al., 2000) und erweitert diese um das objektiv erfasste Diversitätswissen mit mehreren Diversitätsbereichen. Dennoch zeigten sich in Studie 3 keine signifikanten Zusammenhänge zwischen den Einstellungen und den Selbstwirksamkeitserwartungen mit dem Wissen über den SPF ESE bzw. Lernen. Vor dem Hintergrund bisheriger Forschung zu Einstellungen, Selbstwirksamkeitserwartungen und Wissen ist dies nicht ungewöhnlich (vgl. Depaepe & König, 2018; Heyder, Bergold & Steinmayr, 2018) und könnte zudem erklärt werden durch die vergleichsweise kleine Stichprobe und dadurch, dass die Einstellungen nur in Bezug auf Schüler:innen mit SPF im Allgemeinen und nicht die SPF ESE bzw. Lernen im Speziellen erfasst wurden. Die Selbstwirksamkeitserwartungen hingegen lagen für beide spezifischen SPF vor.

Zusammenfassend deuten die Studien dieser Arbeit darauf hin, dass das Diversitätswissen, erfasst mit dem hier vorgestellten Wissenstest, einen Platz in der professionellen Kompetenz von (angehenden) Lehrkräften einnimmt (vgl. Facette Heterogenität zu Unterschieden in den Voraussetzungen der Lernenden, Voss et al., 2015). Die Relevanz für die Umsetzung inklusiven Unterrichts wird durch Zusammenhänge zu Variablen, die sich bereits als unterrichtsrelevant herausgestellt haben (de Boer et al., 2011; Fernandes et al., 2021; Lautenbach & Heyder, 2019; Wray et al., 2022) unterstrichen. Auch

hier sind verschiedene Wirkrichtungen denkbar: Plausibel ist, dass sowohl das Wissen als auch die Einstellungen und Selbstwirksamkeitserwartungen unabhängig voneinander durch Variablen, die in der Vergangenheit liegen, beeinflusst werden. In Anlehnung an das Kompetenzmodell sind die individuellen Eingangsvoraussetzungen und formale Lerngelegenheiten (Kunter, Kleickmann et al., 2013, vgl. vorheriger Abschnitt). Auch könnte es sein, dass in Aus- und Weiterbildungsmaßnahmen alle drei Variablen gleichzeitig explizit oder implizit adressiert werden. Es könnte aber auch sein, dass Personen, die bereits positive Einstellungen zu Inklusion haben, sich eher für ein Studium für sonderpädagogische Förderung mit hohem Inklusionsbezug entscheiden und auch unabhängig davon eher Lehrveranstaltungen und Weiterbildungen aufsuchen, die sich mit Inklusion befassen, und damit ihr Wissen stärken. Das höhere Wissen könnte wiederum dazu führen, dass die betreffenden Personen sich besser auf das Unterrichten in inklusiven bzw. diversen Klassen vorbereitet fühlen und dementsprechend höhere Selbstwirksamkeitserwartungen haben. Andersherum könnte es sein, dass Veranstaltungen, die evidenzbasiertes Wissen fördern, gleichzeitig Fehlannahmen abbauen und damit einen positiven Einfluss auf die Einstellungen nehmen. Diese positiven Einstellungen könnten einen Optimismus begünstigen, der wiederum die Selbstwirksamkeitserwartungen fördert. Neben diesen Annahmen sind viele weitere Wirkmechanismen denkbar, die Fragestellungen für zukünftige Forschung bieten (vgl. Kapitel 6.4.1 zu Implikationen für die Forschung). Deutlich schwieriger ist auch hier die Einordnung der Fehlannahmen. In Studie 2 zeigten sich keine Zusammenhänge der Fehlannahmen mit den Einstellungen und Selbstwirksamkeitserwartungen. Es könnte aber sein, dass sich Zusammenhänge für bereichsspezifische Variablen zeigen würden (vgl. Heyder, Bergold & Steinmayr, 2018 zum negativen Zusammenhang zwischen den Fehlannahmen über Hochbegabung und den Einstellungen zur Förderung Hochbegabter).

Zusammenfassend deuten auch die Ergebnisse zu Zusammenhängen zwischen dem

Wissen und anderen Aspekten professioneller Kompetenz auf die Bedeutung dieser Wissensinhalte hin, wohingegen die Ergebnisse zu Fehlannahmen keine eindeutigen Schlüsse zulassen. Für die praktische Bedeutung besonders relevant sind darüber hinaus die im Folgenden beschriebene Performanz und der Unterrichtserfolg.

Performanz und Unterrichtserfolg. Im Rahmen dieser Arbeit konnten positive Zusammenhänge (1) zwischen dem Wissen über Diversität und den selbstberichteten lernorientierten Unterrichtsinstruktionen und (2) zwischen dem Wissen über den SPF ESE und der sozialen Partizipation und den akademischen Leistungen aller Schüler:innen in Klassen mit und ohne Schüler:innen mit SPF ESE in zwei Lehrkräftestichproben nachgewiesen werden. Lernorientierte Unterrichtsinstruktionen stehen für die Förderung von Lernprozessen und die Anpassung der Aufgaben an die individuellen Bedürfnisse der Schüler:innen (Retelsdorf et al., 2010), hängen positiv mit dem Interesse der Schüler:innen und Lehrkräfte zusammen (Retelsdorf et al., 2010; Schiefele, 2017) und sind somit ein Indikator für die Unterrichtsqualität – vor allem im Kontext schulischer Inklusion. Damit leistet diese Arbeit einen wichtigen Beitrag zu der Erforschung verschiedener Wissensfacetten im Rahmen der professionellen Kompetenz, da ähnliche Zusammenhänge bislang vor allem in Bezug auf das eher allgemeine PPK bzw. GPK aufgezeigt werden konnten (Depaepe & König, 2018; Gess-Newsome et al., 2017; König & Pflanzl, 2016; Pflanzl et al., 2013; Voss et al., 2011, 2022). Auch der Zusammenhang des Diversitätswissens mit Variablen auf Ebene der Schüler:innen schließt sich an Forschung mit dem fachdidaktischen Wissen (Baumert & Kunter, 2013b, Keller et al., 2017; Kunter, Klusmann et al., 2013) bzw. der adaptiven Lehrkompetenz (Beck et al., 2008; Brühwiler, 2014) an und erweitert diese dementsprechend substantiell um das inklusionsbezogene Wissen über Diversität. Beide Ergebnisse sprechen für eine hohe Relevanz dieses Wissens.

Auch hier kann darüber hinaus über verschiedene Wirkzusammenhänge spekuliert

werden: Lerngelegenheiten in der Aus- und Weiterbildung der Lehrkräfte könnten den Zusammenhang zwischen dem Gesamtwissen über Diversität und den lernorientierten Unterrichtsinstruktionen erklären. Da Diversität und ein adaptiver Umgang damit eng miteinander verknüpft sind, erscheint es plausibel, dass Lehrveranstaltungen beide Aspekte abdecken. Gleichwohl könnte ein hohes Wissen über Diversität die Nutzung lernorientierter Instruktionen erleichtern, andersherum könnten diese die Wahrnehmung für die Unterschiedlichkeit der Schüler:innen schärfen und die Aneignung neuen Wissens ermöglichen. Wichtig bei der Interpretation des Zusammenhangs zwischen dem Wissen und den Variablen auf Schüler:innen-Ebene ist, dass in dieser Arbeit das deklarative Faktenwissen und nicht das konkrete Unterrichtshandeln erfasst wurde. Da Zusammenhänge zwischen dem Wissen und Outcomes der Schüler:innen häufig durch das Unterrichtshandeln mediiert werden (z. B. Keller et al., 2017; Kunter, Klusmann et al., 2013), liegt dieser Mechanismus auch für den hier gefundenen Zusammenhang nahe, müsste aber zukünftig erforscht werden. Zusammenfassend weisen vor allem diese letzten beiden Ergebnisse zum Zusammenhang zwischen dem Wissen und den Unterrichtsinstruktionen bzw. der sozialen Partizipation und den akademischen Leistungen der Schüler:innen auf die Relevanz dieses Wissens für Schüler:innen, also die primären Empfänger:innen schulischer Inklusion, hin. Das Wissen über Diversität scheint damit zwar nicht die einzige, aber eine wichtige Variable zu sein, die Inklusion in der Praxis voranbringen könnte (vgl. Göransson & Nilholm, 2014).

Es zeigten sich keine Zusammenhänge zwischen den Fehlannahmen und den Unterrichtsinstruktionen (ein Zusammenhang zu Variablen auf Schüler:innen-Ebene wurde nicht untersucht). Da Mathe-Geschlechter-Fehlannahmen in einer aktuellen Studie positiv mit Mathe-Gender-Stereotypen und mit der Annahme, dass man für Mathe angeborene Fähigkeiten benötigt, die nicht erlernbar sind, korrelierten (Dersch et al., 2022), sind bereichsspezifische negative Zusammenhänge zu Facetten der Unterrichtsqualität aber

denkbar und bleiben Fragestellungen für zukünftige Forschung.

6.2.2 Erfassung und Struktur

Aus den Analysen von Studie 1 gingen Erkenntnisse auf Itemebene hervor, die im Folgenden diskutiert werden (alle Itemkennwerte sind einzusehen in Steinmayr et al., 2022, online verfügbar unter <https://doi.org/10.23668/psycharchives.5136>; Interpretation anhand Kelava & Moosbrugger, 2020): Die Schwierigkeit der meisten Items lag in dem vorgeschlagenen Bereich von 5 bis 95 (hier: Mittelwert \cdot 100) und beschränkte sich damit nicht nur auf Werte um 50. Bei diesen mittleren Werten kann besonders gut zwischen Personen verschiedener Merkmalsausprägungen unterschieden werden, da eine hohe Itemvarianz ermöglicht wird. Allerdings sind auch niedrige und hohe Werte nötig, wenn in Personengruppen mit besonders niedrigen bzw. hohen Merkmalsausprägungen differenziert werden soll. Die Itemschwierigkeiten lagen in den verschiedenen Stichproben tendenziell mehr unter 50 und weniger über 50. Das bedeutet, dass die Items eher schwierig waren und der Test somit besonders gut geeignet ist für Personen mit einem mittleren bis hohen Wissen. Das ist sinnvoll, da er sich an (angehende) Lehrkräfte und nicht an Laien richtet und auch in Stichproben mit einer besonders hohen Expertise (z. B. ausschließlich Lehrkräfte für sonderpädagogische Förderung) funktionieren soll. Deskriptiv wird der Test somit der Anforderung gerecht, dass die Items weder so leicht sind, dass sie von allen richtig beantwortet werden, noch so schwer, dass sie von niemandem richtig beantwortet werden können. Außerdem haben die Items eine hohe Varianz, da sich viele von ihnen im Bereich von 50 bewegen, und gleichzeitig ermöglicht er eine Differenzierung in Stichproben mit Personen mit einem insgesamt eher hohen Wissen. Die Trennschärfe der meistens Items war mit Werten zwischen .4 und .7 gut, was bedeutet, dass die Items adäquat mit den anderen Items der jeweiligen Skala korrelieren, die Skala somit gut repräsentieren und zwischen Personen mit einem niedrigen und hohen Wissen differenzieren können. Den definierten

Trennschärfen von $> .2$ der Pilotierung (Heyder, Vaskova et al., 2018) konnte die Evaluation der Langversion weiterhin gerecht werden. Allerdings war die Trennschärfe einzelner Items der Kurzversion niedriger. Die betreffenden Items könnten kritisch auf ihre inhaltliche Bedeutung für den Test überprüft und gegebenenfalls durch andere Items der Langversion ausgetauscht werden. Eine Analyse der Kennwerte der Fehlannahmen wurde nicht vorgenommen, sodass darüber keine Aussage getroffen werden kann.

Die Analysen ergaben einen guten Modellfit für das hierarchische 5-Faktoren-Modell in beiden Testversionen. Die angenommene Struktur konnte, vergleichbar mit anderen Leistungstests (Horn et al., 2004; Kunina-Habenicht et al., 2020), also durch die Daten gestützt werden. Zur Konstruktion wurden die inhaltlichen Bereiche SPF ESE, SPF Lernen, Teilleistungsstörungen, Geschlecht und Hochbegabung herangezogen. Damit schließt sich der Test an vorherige Tests an, in denen eine ähnliche Operationalisierung anhand verschiedener Diversitätsbereiche vorgenommen wurde (Geschlecht, sozialer Hintergrund, Kunina-Habenicht et al., 2020; Vorwissen, Intelligenz, Interesse, sozioökonomischer Status, Lern- und Entwicklungsstörungen, Geschlechterdifferenzen, ADHS, Begabung, Hohenstein et al., 2017; Testangst, ADHS, Dyslexie, Essstörungen, geistige Fähigkeiten, Begabungen, verschiedene ethnische Hintergründe, Voss et al., 2011). Diese Tests bilden die Diversitätsbereiche allerdings nur mit wenigen Items innerhalb einer Skala und zum Teil nur wenig reliabel ab. Mit den hier vorgestellten Studien zum DiWi(-K) konnten die bestehenden Tests dementsprechend substantiell erweitert werden. Dafür sprechen auch die guten Reliabilitätswerte für die Subskalen und den Gesamtttest. Es liegen darüber hinaus Messinstrumente vor, die für den Kontext schulischer Inklusion geeignet scheinen und in denen nicht anhand verschiedener Diversitätsbereiche unterschieden wird (Lernprozesse, Dispositionen/Unterschiede, Diagnose, Klassenführung, Strukturierung, Binnendifferenzierung/Individualisierung, König et al., 2017; Differenzierungsmaßnahmen,

Methodenvielfalt im Unterricht, König & Blömeke, 2009, 2010; individuelle Lernprozesse, Umgang mit Diversität, Diagnostik bei Lernstörungen, Wildbrett et al., 2022). Diese Unterscheidung könnte auch in Bezug auf den DiWi(-K) weitere Erkenntnisse ermöglichen, da auch hier Items zu verschiedenen Bereichen wie Prävalenzen, Diagnosen und Interventionen vorkommen. In Anlehnung an die Evaluation eines Tests zum GPK mittels IRT (Malva et al., 2020) könnte es zum Beispiel sein, dass Items zu Prävalenzen mit einem stärkeren Theoriebezug für Studierende besonders leicht sind, wohingegen Items zu Interventionen mit einem stärkeren Praxisbezug für Lehrkräfte besonders leicht sind.

Die Struktur der Fehlannahmen wurde in dieser Arbeit nicht überprüft. Die vereinzelt zusammenhängen im Rahmen dieser Arbeit und aus vorherigen bereichsspezifischen Studien (Dersch et al., 2022; Heyder, Bergold & Steinmayr, 2018; Sanz-Cervera et al., 2017) deuten darauf hin, dass sowohl die Erfassung als auch der Abbau von Fehlannahmen von praktischer Relevanz für Lehrkräfte und Schüler:innen sind, allerdings lassen sie sich auf dieser Basis nicht in ihrer Struktur bzw. in einem Modell darstellen und bleiben somit Bestandteil zukünftiger Forschung.

6.3 Stärken und Grenzen der Arbeit

Als Stärken dieser Arbeit sind zu nennen, (1) dass ein objektiver Wissenstest eingesetzt und das Wissen nicht nur im Selbstbericht erfasst wurde (vgl. Paju et al., 2016; Sulek et al., 2018), (2) dass verschiedene Diversitätsbereiche und nicht nur ein isolierter Bereich abgebildet wurden (z. B. Ginsburg et al., 2021; Haimour & Obaidat, 2013; Heyder, Bergold & Steinmayr, 2018; Scituito et al., 2000; Sousa et al., 2017), weshalb das Wissen über Diversität deutlich umfassender gemessen wurde, (3) dass diese Bereiche innerhalb eines Tests und nicht nur innerhalb einer Skala vorliegen, was dem heterogenen Konstrukt Diversität besser gerecht wird (Voss et al., 2011), (4) dass Stichproben aus Lehramtsstudierenden und praktisch tätigen Lehrkräften befragt wurden, (5) dass der

Evaluation in Bezug auf die Überprüfung eines nomologischen Netzes rund um das Wissen über Diversität ein aktuelles, fundiertes und vielfach genutztes Modell (COACTIV, Baumert & Kunter, 2013a; Kunter, Kleickmann et al., 2013) zugrunde liegt und (6) dass das Wissen der Lehrkräfte mit sozialen und akademischen Variablen auf Ebene der Schüler:innen in Verbindung gebracht wurde, wozu es im Kontext schulischer Inklusion bislang kaum Erkenntnisse gibt. Damit wird diese Arbeit wichtigen Ansprüchen gerecht, die an die Erforschung schulischer Inklusion gestellt werden: Es werden sowohl soziale als auch akademische Variablen untersucht (Göransson & Nilholm, 2014; Grosche, 2015), und die Arbeit leistet einen wichtigen Beitrag zur Erforschung von Gelingensfaktoren schulischer Inklusion (Göransson & Nilholm, 2014). Zur Einordnung der Ergebnisse müssen neben den Stärken auch die Grenzen der Arbeit beachtet werden. Nachdem die Limitationen der Einzelstudien in Kapitel 5 (Studie 1: S. 119 f., Studie 2: S. 148 f., Studie 3: S. 196 f.) bereits diskutiert wurden, werden im Folgenden die Grenzen thematisiert, die bei der studienübergreifenden Interpretation der Ergebnisse zu beachten sind.

Erstens wurde der Wissenstest zwar sowohl in der Lang- als auch der Kurzversion in Studie 1 anhand der KTT testtheoretisch überprüft, allerdings bleiben noch Fragen offen, die die Belastbarkeit der Ergebnisse aller drei Studien einschränken. Das hierarchische 5-Faktoren-Modell wurde mittels konfirmatorischer Faktorenanalysen durch die Daten gestützt, jedoch konnte auf Basis der zu kleinen Substichproben keine Messinvarianz über verschiedene Gruppen hinweg getestet werden. Dementsprechend kann es sein, dass die Werte verschiedener Gruppen unterschiedlich interpretiert werden müssen (Putnick & Bornstein, 2016). Dies kann nicht ausgeschlossen werden für den Vergleich zwischen Bachelor- und Master-Studierenden (Studie 1), zwischen Studierenden/Lehrkräften für sonderpädagogische Förderung vs. anderer Lehramtsstudiengänge (Studie 1 & 2), zwischen Studierenden und Lehrkräften (Studie 2) und zwischen Lehrkräften mit mindestens einer vs.

keiner Fortbildung zu Inklusion (Studie 2). Außerdem wurde die Faktorstruktur nur in Studierenden- nicht aber in Lehrkräftestichproben überprüft, weshalb die Belastbarkeit der Ergebnisse der Studien 2 und 3 zum Gesamttest und den beiden Subskalen zu den Förderschwerpunkten ESE und Lernen einschränkt ist. So könnte es zum Beispiel sein, dass die Studierenden im Laufe des Praxissemesters des Masterstudiums aufgrund des hohen Anteils der Schüler:innen mit SPF gemessen an allen Schüler:innen im Alter der Vollzeitschulpflicht (Förderrelation im Jahr 2020 von 7.7%, KMK, 2022) eher mit diesen und weniger mit hochbegabten Schüler:innen konfrontiert werden und sich dazu mehr Wissen aneignen. In diesem Fall wäre der Gruppenunterschied zwischen Master- und Bachelor-Studierenden in Studie 1 nicht auf einen Unterschied im Gesamtwissen, sondern in den beiden Förderschwerpunkten ESE und Lernen zurückzuführen. In Bezug auf die Faktorstruktur ist es auch denkbar, dass das Wissen über die beiden Förderschwerpunkte ESE und Lernen bei den befragten Lehrkräften gar keine distinkten Konstrukte darstellt. So könnten alle Items der beiden Subskalen auf einem Faktor laden, weil sich die Lehrkräfte gleichermaßen mit den Prävalenzen, Diagnosen und Interventionen auseinandersetzen müssen, unabhängig davon, um welchen Förderschwerpunkt es sich handelt. In diesem Fall könnte in Studie 3 das Wissen über den SPF ESE auch Varianz an den Schüler:innen-Outcomes in Klassen mit Kindern mit und ohne SPF Lernen aufklären. Dies war bei den mathematischen Kompetenzen der Schüler:innen in Studie 3 auch tatsächlich der Fall. Gleichwohl könnte dieser Befund ebenfalls auf eine Überlappung der Stichproben zurückzuführen sein. Um diese potenziellen Fehlerquellen ausschließen zu können, sollte in zukünftigen Studien eine Überprüfung der Messinvarianz über Gruppen hinweg und des hierarchischen 5-Faktoren-Modells in einer Lehrkräftestichprobe vorgenommen werden. Damit einhergehend kann der Test auf Basis der Annahmen der IRT überprüft werden (vgl. Kapitel 6.4.1 zu Implikationen für die Praxis). So könnte es zum Beispiel sein, dass die

Skalen des hier geprüften Wissenstests sehr reliabel sind bei Personen mit einem moderaten, nicht aber mit einem eher niedrigen Wissen, was die Aussagekraft in bestimmten Stichproben einschränken würde (vgl. de Ayala, 2022).

Zweitens sind alle drei Studien querschnittlicher Natur, weshalb keins der Ergebnisse kausal interpretiert werden darf. Zwar deutet das COACTIV-Modell (Kunter, Kleickmann et al., 2013) auf eine gewisse Wirkrichtung hin, jedoch sind darüber hinaus viele weitere Wirkmechanismen denkbar (vgl. Kapitel 6.2.1 zur Einordnung in die professionelle Kompetenz). Die vorsichtige Interpretation der Ergebnisse erscheint vor allem in Bezug auf Studie 3 mit einer sehr hohen Praxisrelevanz bedeutsam, da das Wissen nicht fälschlicherweise als Ursache für die soziale Partizipation und die akademischen Leistungen der Schüler:innen interpretiert werden darf. Längsschnittstudien können in Zukunft weitere Erkenntnisse liefern, damit die Mechanismen, die diesen Zusammenhängen zugrunde liegen, besser verstanden werden können.

Drittens haben alle drei Stichproben gemeinsam, dass sie aus Deutschland kamen (Studie 1: Studierende der TU Dortmund, Studie 2: Lehrkräfte aus deutschen Schulen, Studie 3: Lehrkräfte aus Grundschulen in Nordrhein-Westfalen), was den Vorteil hat, dass die Inhalte an die Anforderungen an ein Lehramtsstudium und die praktische Umsetzung von Inklusion an deutschen Schulen angepasst sind (vgl. Kapitel 3.1 zur Bedeutung der Diversitätsbereiche). Dennoch unterscheiden sich die Curricula zu bildungswissenschaftlichen Standards für die Lehramtsausbildung an Hochschulen sogar innerhalb von Deutschland (Hohenstein et al., 2014) und dementsprechend wahrscheinlich noch stärker im Vergleich zu Hochschulen außerhalb von Deutschland. Auch unterscheiden sich die Maßnahmen und Praktiken für schulische Inklusion innerhalb Europas (Ramberg et al., 2018), was noch größere Unterschiede außerhalb Europas nahelegt. Die Ergebnisse können dementsprechend nicht automatisch auf alle (angehenden) Lehrkräfte generalisiert

werden.

Viertens wurden neben dem objektiv gemessenen Wissen auf Ebene der Lehrkräfte alle weiteren Variablen im Selbstbericht erfasst, der verschiedenen Verzerrungen unterliegen kann. So gibt es potenzielle Einflussgrößen, die zu einer Überschätzung oder Unterschätzung der gefundenen Zusammenhänge geführt haben könnten. Zum Beispiel konnte in einer aktuellen Untersuchung gezeigt werden, dass Lehramtsstudierende Probleme haben, ihr Wissen korrekt einzuschätzen und sich entweder unter- oder überschätzen. Eine hohe Selbstwirksamkeit konnte dabei eine Überschätzung vorhersagen (Ernst et al., 2023). Übertragen auf den hier vorgestellten Wissenstest bedeutet das, dass Personen mit einer hohen Selbstwirksamkeit ihre Fähigkeiten eventuell fälschlicherweise als sehr hoch einschätzten und dementsprechend auch ein hohes selbsteingeschätztes Wissen berichteten und bei Unsicherheiten im Wissenstest eher dazu neigten zu raten anstatt die Weiß-nicht-Option zu wählen (vgl. Sherman, 1976). Dies könnte zu einer Überschätzung der Zusammenhänge zwischen dem objektiven Wissen, dem selbsteingeschätzten Wissen und den Selbstwirksamkeitserwartungen geführt haben. Auf der anderen Seite könnte es sein, dass Lehrkräfte mit einem hohen Wissen sich weniger dazu verleiten ließen, sozial erwünscht zu antworten, zum Beispiel in Bezug auf ihre Einstellungen zu Inklusion, einem Konstrukt mit einer hohen Sensitivität für sozial erwünschte Antworten (Lüke & Grosche, 2018; Tempel, 2022), da sie sich bereits mehr mit dem Thema auseinandergesetzt und ihre eigene Position dazu gefunden haben. In diesem Fall würde der Zusammenhang zwischen den Einstellungen und dem Diversitätswissen durch den Einfluss der sozialen Erwünschtheit unterschätzt.

Fünftens wurde bei der Erhebung der verschiedenen Variablen rund um das Diversitätswissen nicht das tatsächliche Unterrichtshandeln der (angehenden) Lehrkräfte erfasst. In Studie 2 wurden die selbsteingeschätzten Unterrichtsinstruktionen der Lehrkräfte

erfasst und in Studie 3 wurden die soziale Partizipation und die akademischen Leistungen der Schüler:innen hinzugenommen. Diese Variablen sind wichtige Indikatoren für die Performanz und den Unterrichtserfolg. Dennoch kann auf Basis der drei Studien keine Aussage darüber getroffen werden, was im Klassenzimmer tatsächlich passiert. Um diesen wichtigen Indikator für die Umsetzung schulischer Inklusion zu erfassen, könnten zukünftig Einschätzungen durch die Schüler:innen (König & Pflanzl, 2016; Pflanzl et al., 2013; Voss et al., 2022) und standardisierte Unterrichtsbeobachtungen (Gess-Newsome et al., 2017; Kuyini & Desai, 2007) vorgenommen werden.

6.4 Implikationen

Aus den Grenzen der Arbeit ergeben sich wichtige Implikationen für die Forschung und eine mögliche Weiterentwicklung des Tests, die im Folgenden durch weitere generelle Implikationen ergänzt werden. Daran anschließend folgen Implikationen für den praktischen Einsatz des Tests selbst bzw. seiner Inhalte.

6.4.1 Forschung

Um zu klären, wie der positive Zusammenhang zwischen dem Wissen und den Fehlannahmen (Studie 2) zustande kommt, könnte ein alternatives Antwortformat Aufschluss über das Wissen und die Fehlannahmen ohne den potenziellen Einfluss von Ratetendenzen geben. Die Weiß-nicht-Option wurde vor allem bei den Studierenden sehr häufig genutzt (Studie 1: 34 – 68 %), bei den Lehrkräften war der Anteil etwas geringer (Studie 2: 37 %; in Studie 3 wurde das Nicht-Wissen nicht ausgewertet). Ob die Werte das tatsächliche Nicht-Wissen widerspiegeln oder ob sie durch ein unterschiedliches Rateverhalten herbeigeführt wurden, kann abschließend nicht bewertet werden. Im besten Fall wurde gar nicht oder personenunabhängig geraten, es könnte aber auch sein, dass die Ratefrequenz von bestimmten Persönlichkeitsmerkmalen oder der Fähigkeit zur Selbsteinschätzung beeinflusst

wurde, was zu systematischen Verzerrungen führen und den prädiktiven Wert der Testwerte einschränken kann (Moosbrugger, Gäde et al., 2020). Als Alternative kann zum Beispiel ein Richtig-falsch-Antwortformat mit Angabe der Antwortsicherheit auf einer Likertskala getestet werden (z. B. Dersch et al., 2022; Eitel et al., 2021; Hasan et al., 1999). Wenn Wissen, Fehlannahmen und Nicht-Wissen eindeutig voneinander getrennt werden können, bietet dies eine wichtige Voraussetzung für die Konstruktion von Widerlegungstexten, die für den praktischen Umgang mit Fehlannahmen und ihrer Auflösung (vgl. Kapitel 6.4.2 zur Implikationen für die Praxis) in ihrer Wirkung auf (angehende) Lehrkräfte erforscht werden können. Klassischerweise werden Widerlegungstexte zu naturwissenschaftlichen Themen wie genetisch veränderte Lebensmittel, Evolution, Energie und Impfstoffe eingesetzt, können aber auch auf andere Bereiche ausgeweitet werden (Zengilowski et al., 2021). Dem Nachteil, dass viele Studien nur einen Text zu einem isolierten Thema einsetzen, könnte in diesem Fall direkt begegnet werden, indem Fehlannahmen der verschiedenen inhaltlichen Subskalen herausgefiltert und mittels Widerlegungstext bearbeitet werden. Die Angabe einer Antwortsicherheit könnte hier außerdem den Vorteil haben, dass die Stärke einer Fehlannahme mit evaluiert wird. Diese wiederum kann in Kombination mit weiteren Faktoren einen Einfluss darauf haben, wie leicht bzw. schwierig sie zu korrigieren ist.

Die Evaluation des vorliegenden Wissenstests wurde anhand der KTT vorgenommen. In Ergänzung dazu ist eine Überprüfung mittels IRT möglich, da diese als wichtige Ergänzung zur KTT angesehen werden kann und besonders geeignet für ein dichotomes Antwortformat ist (Moosbrugger, Schermelleh-Engel et al., 2020). Im Fall dieses Wissenstests wird in der IRT die Wahrscheinlichkeit, mit der eine Person die einzelnen Items korrekt (für das Wissen mit 1 kodiert, Rest mit 0 kodiert) bzw. inkorrekt (für die Fehlannahmen mit 1 kodiert, Rest mit 0 kodiert) lösen wird, anhand einer Wahrscheinlichkeitsfunktion modelliert. Die Lösungswahrscheinlichkeit hängt dabei sowohl

von der Aufgabenschwierigkeit als auch von der latenten Merkmalsausprägung (Wissen bzw. Fehlannahmen) dieser Person ab. Da beide Testtheorien als gleichwertig und in Ergänzung zueinander zu verstehen sind (Moosbrugger, Schermelleh-Engel et al., 2020), lässt sich auch ein Mehrwert der Nutzung beider Theorien ableiten. In diesem konkreten Fall ist ein Vorteil der IRT, dass eine Evaluation auf Itemebene vorgenommen werden kann (Embretson & Reise, 2000). Im Vergleich zur KTT wird hier nicht angenommen, dass die Reliabilität eines Konstrukts über eine Skala hinweg gleich ist, sondern es werden unterschiedlich reliable Bereiche in diesem Spektrum zugelassen. Dementsprechend könnte mittels IRT geprüft werden, ob die guten psychometrischen Eigenschaften über die Skalen hinweg und unabhängig von den untersuchten Populationen stabil sind oder ob sie sich in Abhängigkeit davon unterscheiden. Vor diesem Hintergrund erscheint eine Modellierung der in Studie 1 geprüften 5-faktoriellen Struktur (Wissen über den SPF ESE, den SPF Lernen, Teilleistungsstörungen, Geschlecht und Hochbegabung) mit einem allgemeinen Faktor (Wissen über Diversität) anhand der IRT sinnvoll, um den Test gegebenenfalls optimieren und weiterentwickeln zu können.

Für eine zukünftige Weiterentwicklung ist außerdem zu beachten, dass der vorgestellte Wissenstest Diversität nicht in ihrer Gesamtheit widerspiegeln kann. Aus der Literatur gehen zum Beispiel Diversitätsfacetten wie sozialer Hintergrund, Sprache, Religion, Intelligenz, kultureller Hintergrund und Behinderungsmerkmale hervor (Faulstich-Wieland, 2011; Grosche, 2015; Zulliger & Tanner, 2013). Deshalb ist die Hinzunahme weiterer Subskalen langfristig denkbar. Die Lehrkräfte aus Studie 2 gaben im Zuge einer offenen Frage am Ende der Umfrage die Diversitätsbereiche Deutsch als Fremdsprache, Lernmotivation, sozio-ökonomischer Status und kultureller Hintergrund als aus ihrer Sicht praxisrelevant an.

In Studie 1 zeigte sich ein positiver Zusammenhang zwischen dem Wissen über

Diversität und dem pädagogischen Unterrichtswissen. Auch aus dem COACTIV-Modell (Baumert & Kunter, 2013a) lässt sich ableiten, dass verschiedene Wissensfacetten Teil der professionellen Kompetenz sind und Zusammenhänge zwischen eben diesen anzunehmen sind. In den letzten Jahren und damit parallel zu der Testentwicklung im Rahmen des Projekts DoProfiL und des Testeinsatzes im Rahmen dieser Arbeit ließ sich eine starke Neu- bzw. Weiterentwicklung von inhaltlich nahen Messinstrumenten beobachten (Dersch et al., 2022; Ginsburg et al., 2021; Heyder, Bergold & Steinmayr, 2018; Hohenstein et al., 2017; König et al., 2017; Kunina-Habenicht et al., 2020; Wildbrett et al., 2022). Es erscheint deshalb ratsam, theoretisch und empirisch klar zu definieren, welche Überschneidungen und Unterschiede es gibt, um eine passgenaue Nutzung für Forschung und Praxis zu ermöglichen. Hohe Zusammenhänge sind zum Beispiel zwischen den Fehlannahmen der Subskala Geschlecht im DiWi und den Mathe-Geschlechter-Fehlannahmen (Dersch et al., 2022), zwischen Items der Skala SPF ESE mit Bezug zu Angst bzw. Ängstlichkeit und dem TKAAs (Teacher Knowledge of Anxiety Assessment mit Items zu Angstsymptomen und evidenzbasierten Strategien zum Abbau von Ängsten, Ginsburg et al., 2021) denkbar. Auch liegt es nahe, dass es Überschneidungen mit dem Diversitätswissen und der Skala zu Lernen und Lernenden des PPK-V (PPK erfasst über Vignetten, Wildbrett et al., 2022) gibt, bei der individuelle Lernprozesse, der Umgang mit Heterogenität und lern- und motivationspsychologisches Wissen erfasst werden. Unterschiede zu weiteren Tests bestehen unter anderem in der Struktur. Der DiWi(-K) ist inhaltlich anhand verschiedener Diversitätsbereiche strukturiert und enthält Items, die sich Bereichen wie Prävalenz, Zusammenhänge zu anderen Merkmalen, Diagnostik und Unterrichtsgestaltung zuordnen lassen könnten. Andere Tests, wie der BilWiss-2.0 (Kunina-Habenicht et al., 2020) enthalten Subskalen zur Unterrichtsgestaltung, zu Lernen und Entwicklung allgemein und dem Lehrberuf als Profession. Gerade die letzte Skala dürfte deshalb weniger nah am im DiWi(-

K) abgefragten Wissen sein. Interessant wäre es deshalb zu erforschen, welchen prädiktiven Wert das Diversitätswissen neben anderen inklusionsbezogenen Wissensfacetten für die Performanz und den Unterrichtserfolg hat. Vor dem Hintergrund der Verfügbarkeit verschiedener Tests erscheint es außerdem für zukünftige Forschung absolut ratsam, diese objektiven Tests zu nutzen und das Wissen nur ergänzend, nicht aber ausschließlich im Selbstbericht zu erfassen. Zwar zeigte sich in Studie 1 eine moderat positive Korrelation, es ist aber davon auszugehen, dass das selbstberichtete Wissen deutlich stärker Verzerrungen ausgeliefert ist (Ernst et al., 2023).

Dem Zusammenhang zwischen dem Wissen der (angehenden) Lehrkräfte und den formalen Lerngelegenheiten, den beruflichen Überzeugungen, der Performanz und dem Unterrichtserfolg können verschiedene Wirkzusammenhänge zugrunde liegen. Dem COACTIV-Modell (Kunter, Kleickmann et al., 2013; Voss et al., 2015) folgend erscheint es plausibel, dass formale Lerngelegenheiten einen Einfluss auf das Wissen haben und das Wissen wiederum einen Einfluss auf die Performanz und den Unterrichtserfolg hat. Annahmen über den Zusammenhang zwischen dem Wissen und den beruflichen Überzeugungen können auf Basis des Modells nicht getroffen werden. Für die Erforschung der Richtungen der Zusammenhänge und mögliche Moderatoren bzw. Mediatoren (vgl. Kapitel 6.2.1 zur Einordnung in die professionelle Kompetenz) eignen sich Längsschnittstudien. Diese haben den Vorteil, dass individuelle Entwicklungen über die Zeit praxisnah im Feld erfasst und ausgewertet werden können (Brünken et al., 2019). Dazu bieten sich zum Beispiel folgende Fragestellungen an, die das Wissen fokussieren und zum Teil auf die Fehlannahmen angewendet werden können: (1) Wie entwickeln sich das Wissen bzw. die Fehlannahmen von Lehramtsstudierenden im Laufe ihres Studiums und im Übergang zur praktischen Tätigkeit im Referendariat? Hierzu können verschiedene Arten formaler Lerngelegenheiten (z. B. Hochschulveranstaltungen, Praxiserfahrungen) erfasst und

mit der Entwicklung des Wissens bzw. der Fehlannahmen in Verbindung gebracht werden. Eine Studie zu den Einstellungen, Selbstwirksamkeitserfahrungen und dem Wissen über schulische Inklusion bietet erste Anhaltspunkte über einen Abfall aller drei Variablen in diesem Übergang (Mintz et al., 2020). Allerdings wurde hier das Wissen nur über den Selbstbericht und nicht objektiv gemessen, und es liegen nur zwei Messzeitpunkte vor (vor und im Referendariat). Die Nutzung des hier vorgestellten Wissenstests und die Erfassung während des Bachelors, während des Masters und im Referendariat könnten weitere Erkenntnisse zur Wissensentwicklung und für die Wissensförderung in der Lehrkräftebildung bieten. Dadurch ließen sich theoretische und praktische Lerngelegenheiten, die in allen drei Studien dieser Arbeit positiv (in der kleinen Stichprobe in Studie 3 allerdings nicht signifikant) mit dem Wissen korrelierten, und der in Studie 2 vorgenommene Vergleich zwischen Lehrkräften und Studierenden detaillierter betrachten. Auch für die Erforschung der Fehlannahmen erscheint dies hilfreich, damit der Zusammenhang zwischen dem beruflichen Kontakt und den Fehlannahmen inhaltlich eingeordnet und um Erkenntnisse zu theoretischen Lerngelegenheiten, zu denen sich in Studie 2 keine signifikanten Zusammenhänge zeigten, ergänzt werden kann. (2) Wie beeinflusst das Wissen über Diversität die Einstellungen über Inklusion und die Selbstwirksamkeitserwartungen zum inklusiven Unterrichten bzw. wird von ihnen beeinflusst? Eine aktuelle Längsschnittstudie zu den Einstellungen und Selbstwirksamkeitserwartungen von Lehrkräften im Kontext Inklusion ergab einen positiven Einfluss der Selbstwirksamkeitserwartungen auf die Einstellungen, nicht aber umgekehrt (Savolainen et al., 2022). Auch zeigte sich in längsschnittlichen Analysen, dass das bildungswissenschaftliche Wissen von Referendar:innen die Abnahme ihrer Selbstwirksamkeitserwartungen im Laufe des Referendariats negativ vorhersagen konnte (Dicke et al., 2015). Mit Ausnahme der relativ kleinen Stichprobe in Studie 3 korrelierte das Wissen über Diversität in dieser Arbeit konstant positiv mit den Einstellungen und den

Selbstwirksamkeitserwartungen der (angehenden) Lehrkräfte. Ob ein höheres Wissen die Einstellungen und Selbstwirksamkeitserwartungen steigern kann, ob diese beiden dafür sorgen, dass Lehrkräfte sich mehr Wissen aneignen möchten, ob Persönlichkeits- und Ausbildungsmerkmale alle drei Variablen beeinflussen und welche anderen Wirkmechanismen den Zusammenhängen zugrunde liegen, kann längsschnittlich überprüft werden. (3) Wird der Zusammenhang zwischen dem Wissen und den Outcomes der Schüler:innen mediiert durch die Unterrichtsinstruktionen bzw. das Unterrichtshandeln? Der hier vorgestellte Test misst das deklarative Faktenwissen, das nicht mit dem konkreten Unterrichtshandeln gleichzusetzen ist. In Studie 3 zeigten sich positive Zusammenhänge zwischen dem Wissen der Lehrkraft über ESE und der sozialen Partizipation und den akademischen Leistungen der Schüler:innen in Klassen mit und ohne Schüler:innen mit SPF ESE. Studien zum PCK von Lehrkräften deuten darauf hin, dass der Zusammenhang durch die Unterrichtsinstruktionen erklärt werden könnte (z. B. Keller et al., 2017; Kunter, Klusmann et al., 2013). Ob die lernorientierten Unterrichtsinstruktionen, die in Studie 2 positiv mit dem Wissen über Diversität zusammenhängen, den in Studie 3 gefundenen Zusammenhang zwischen dem Lehrkräftewissen und den Schüler:innen-Outcomes erklären können, lässt sich zukünftig überprüfen. Im Zuge einer längsschnittlichen Untersuchung könnten auch mögliche Wirkzusammenhänge zwischen dem Wissen und den lernorientierten Unterrichtsinstruktionen untersucht werden. (4) Neben den Unterrichtsinstruktionen als Moderatoren können weitere potenzielle Moderatoren der verschiedenen Zusammenhänge erforscht und für die Praxis nutzbar gemacht werden (vgl. Kapitel 6.4.2 zu Implikationen für die Praxis) – im Idealfall längsschnittlich, querschnittliche Designs sind für eine erste Annäherung aber ebenso denkbar. Das gilt zum Beispiel für Faktoren, die die Transformation von formalen Lerngelegenheiten in Wissen beeinflussen. Im COACTIV-Modell wird hier zum Beispiel die Nutzung angeführt, die wiederum von individuellen

Eingangsvoraussetzungen, wie kognitiven Fähigkeiten, Motivation und Persönlichkeit abhängt (Voss et al., 2015). Außerdem findet das Unterrichtshandeln der Lehrkräfte immer in einem Kontext (z. B. Bildungssystem, individuelle Schule; Voss et al., 2015) statt. Deshalb könnte die Frage untersucht werden, in welchen Rahmenbedingungen Lehrkräfte das Wissen in besonders effektives Unterrichtshandeln umsetzen können.

6.4.2 Praxis

Der Test selbst bzw. seine Inhalte können ebenso wie Erkenntnisse aus den Studien zum Wissen über Diversität vielfältige Anregungen für die Aus- und Weiterbildung von (angehenden) Lehrkräften geben. Im Folgenden werden verschiedene Ansätze präsentiert, die sich beide Bereiche zunutze machen. Aus Mangel an Studien zum objektiven Wissen werden auch Befunde zum wahrgenommenen Wissen einbezogen.

In Studie 1 und 2, in denen das Wissen über Diversität mit allen fünf Subskalen erfasst wurde, zeigten sich positive Zusammenhänge zwischen den Einstellungen, den Selbstwirksamkeitserwartungen und dem Wissen. Obwohl dieser Befund in der kleineren Lehrkräftestichprobe mit dem beiden Subskalen zu den SPF ESE und Lernen nicht repliziert werden konnte, deutet die Gesamtheit bisheriger Studien auf bestehende positive Zusammenhänge (Ghanizadeh et al., 2006; Lu et al., 2020; Sciotto et al., 2000; Yada et al., 2022). Für die Aus- und Weiterbildung von Lehrkräften im Rahmen von Hochschulveranstaltungen und Trainings bedeutet dies, dass diese drei Variablen zusammen betrachtet und bearbeitet werden könnten. Dies erscheint aus mehreren Gründen besonders wichtig: (1) Sowohl die Einstellungen als auch die Selbstwirksamkeitserwartungen und das wahrgenommene Wissen über Inklusion scheinen im Übergang von der Universität in die Praxis an der Schule in einer Art Realitätsschock schlagartig abzunehmen (Mintz et al., 2020). (2) Dabei erhöht ein höheres wahrgenommenes Wissen nicht automatisch die Einstellungen zu Inklusion und nimmt den Lehrkräften die Bedenken und Sorgen in Bezug

auf die Inklusion von Schüler:innen mit Behinderungen bzw. SPF (Forlin & Chambers, 2011). (3) Der Erfahrungsaustausch mit Menschen mit Behinderungen erwies sich auf der einen Seite als positiv für die Entwicklung von Selbstwirksamkeitserwartungen (Wray et al., 2022). Auf der anderen Seite hatten Lehrkräfte, die sich im Rahmen eines Trainings mit einer Person mit Behinderung austauschten, tendenziell mehr Sorgen über Inklusion als diejenigen, die das Training ohne direkten Kontakt absolvierten (Forlin & Chambers, 2011). (4) Studien dokumentieren einen Rückgang der Selbstwirksamkeitserwartung im inklusiven Unterrichten von Lehramtsstudierenden mit der Länge ihres Studiums bzw. im Laufe eines Trainings (Braksiek, 2022; Schwab et al., 2017; Tan & Amrhein, 2019). Dieser könnte durch ein gesteigertes Wissen über und ein gesteigertes Bewusstsein für die Komplexität von Inklusion und die Befürchtung, für die Umsetzung nicht genug gelernt zu haben, erklärt werden. Diese Befunde deuten darauf hin, dass die Konfrontation mit möglichen Schwierigkeiten von Inklusion für viele Lehrkräfte eine Herausforderung darstellt, die es abzupuffern und zu begleiten gilt, damit diese Lehrkräfte Inklusion als Chance und Bereicherung wahrnehmen und dementsprechend umsetzen können (vgl. Wischer, 2007). Deshalb sollten die verschiedenen Aspekte ihrer professionellen Kompetenz nicht isoliert, sondern gemeinsam betrachtet werden. Aus den Standards der Lehrkräftebildung (KMK, 2014) lassen sich verschiedene diversitätsbezogene Kompetenzen entnehmen, die in diesem Rahmen angesprochen werden könnten. Lehrkräfte sollen demnach eine wertschätzende und anerkennende Haltung gegenüber Diversität einnehmen und um die Bedeutung von Diversität für das Gelingen von Lernprozessen wissen (z. B. Reflexion eigener Einstellungen). Sie sollen außerdem den Begriff Diversität und Merkmale davon kennen (z. B. Sammlung von evidenzbasiertem Wissen). Außerdem reflektieren sie die professionellen Anforderungen des Umgangs mit Diversität (z. B. Erkennen eigener Selbstwirksamkeitserwartungen). Für die pädagogisch-psychologische Hochschullehre und pädagogisch-psychologische Trainings

bieten diese Standards somit einen möglichen Rahmen, um alle drei Konstrukte innerhalb einer Veranstaltung zu thematisieren, miteinander in Verbindung zu bringen und schlussendlich zu stärken.

In der Hochschullehre könnten der Test bzw. seine Inhalte und das damit einhergehende Wissen bzw. die aufgedeckten Fehlannahmen außerdem für Reflexionen genutzt werden. Als Rahmenmodell kann zu diesem Zweck zum Beispiel das EDAMA-Modell (Aeppli & Lötscher, 2016) herangezogen werden. Reflexion wird hier als kognitiver Prozess zur Strukturierung bzw. Umstrukturierung gesehen. Evidenzbasierte Fakten, gängige Fehlannahmen und eigene Recherchen bei Unsicherheiten der Studierenden könnten zum Beispiel als Auslöser für Reflexionen genutzt werden, in die ebenfalls in Anlehnung an die oben beschriebene Verknüpfung zum Wissen weitere Variablen wie Einstellungen und Selbstwirksamkeitserwartungen mit einbezogen werden.

Damit einhergehend können Lehrveranstaltungen konzipiert und durchgeführt werden, in denen die praktischen Erfahrungen aus dem Schulalltag oder aus einem Erfahrungsaustausch mit Menschen verschiedener Diversitätsbereiche mit dem deklarativen Faktenwissen über verschiedene Diversitätsbereiche abgeglichen werden. Gerade weil praktische Lernerfahrungen nicht automatisch mit positiveren Einstellungen, höheren Selbstwirksamkeitserwartungen und einem höheren Wissen einhergehen (Mintz et al., 2020; Forlin & Chambers, 2011), könnte die Verbindung beider Elemente die Lehrkräfte dabei unterstützen, erlerntes evidenzbasiertes Wissen in der Praxis zu nutzen und praktische Erfahrungen vor dem Hintergrund wissenschaftlicher Studien einzuordnen und – im Fall von Fehlannahmen – zu korrigieren. Diese Verbindung könnte darüber hinaus hilfreich sein, um die bestehende Skepsis vieler Studierender gegenüber evidenzbasierten fachunabhängigen Inhalten zum Lehren und Lernen (Voss, 2022) zu mindern. Besonders bei praktisch tätigen Lehrkräften erscheinen Trainings zur Aktualisierung ihres evidenzbasierten Wissens enorm

wichtig, da diese angeben, ihr Wissen vorrangig aus den Medien zu beziehen (Heyder, Bergold & Steinmayr, 2018) und nicht ohne weiteres in der Lage sind, sich neues evidenzbasiertes Wissen anzueignen (Mintz, 2022).

Dass Studierende bzw. Lehrkräfte für sonderpädagogische Förderung ein höheres Wissen haben als Studierende bzw. Lehrkräfte anderer Lehrämter, erwies sich in Einklang mit vorherigen Studien zu verwandten Wissensfacetten (Heyder, Vaskova et al, 2018; Sanz-Cervera et al., 2017; Sousa et al., 2017) auch in dieser Arbeit als robuster Befund. Dies könnten sich Lehrveranstaltungen und Fortbildungen zunutze machen, um einen Wissensaustausch anzuregen und die Kooperation von Lehrkräften verschiedener Studienhintergründe zu fördern.

Die genannten Vorschläge zur Nutzung des Tests und der Erkenntnisse daraus beziehen sich sowohl auf das Wissen als auch auf die Fehlannahmen. Um Fehlannahmen ganz explizit zu bearbeiten und zu verändern, können darüber hinaus Widerlegungstexte zum Einsatz kommen (z. B. Zengilowski et al., 2021), die das Ziel haben, Missverständnisse und naive Überzeugungen von Lernenden aufzuzeigen und sofort Korrekturen anzubieten. Diesem Vorgehen liegt unter anderem die Annahme zugrunde, dass die inkorrekte und korrekte Information simultan präsentiert werden müssen, damit die inkorrekte Information als bereits erworbene, aber nicht mehr korrekte Information erkannt und durch die korrekte Information ersetzt werden kann (Kendeou & O'Brien, 2014). Am Beispiel des vorliegenden Wissenstests (hier die Subskala Hochbegabung) könnte dies wie folgt formuliert werden: *Entgegen der häufigen Annahme, dass kognitiv hochbegabte Kinder und Jugendliche im Vergleich zu durchschnittlich begabten Gleichaltrigen mehr Probleme haben, zeigen wissenschaftliche Studien eine vergleichbare bis bessere sozio-emotionale Anpassung und weniger Verhaltensprobleme.* Auch wenn in Bezug auf Widerlegungstexte noch einige Forschungslücken identifiziert wurden (vgl. Zengilowski et al., 2021), handelt es sich um

einen vielversprechenden und ökonomischen Ansatz, dessen Vorzüge auch für Fehlannahmen über Diversität von (angehenden) Lehrkräften im Kontext schulischer Inklusion genutzt werden könnten.

Der positive Zusammenhang zwischen dem Wissen der Lehrkräfte über den SPF ESE und der sozialen Partizipation und den akademischen Leistungen ihrer Schüler:innen (Studie 3) weist auf die praktische Bedeutung dieses Grundlagenwissens hin. Wenngleich aufgrund des querschnittlichen Designs keine kausalen Schlüsse zulässig sind, erscheint es vor dem Hintergrund des COACTIV-Modells (Kunter, Kleickmann et al., 2013; Voss et al., 2015) plausibel, dass das Wissen als Teil der professionellen Kompetenz auch im Kontext schulischer Inklusion einen Einfluss auf die Performanz und den Unterrichtserfolg von Lehrkräften hat. Dabei sind Wissen und tatsächliches Unterrichtshandeln nicht gleichzusetzen (Voss et al., 2015). Ein möglicher Ansatz, um Praktiker:innen dabei zu unterstützen, ihr Wissen in erfolgreiches Unterrichtshandeln umzusetzen, könnte neben den oben beschriebenen Hochschulveranstaltungen und Trainings das Coaching für Lehrkräfte im Allgemeinen, für Lehrkräfte für sonderpädagogische Förderung im Speziellen und für weitere Hilfskräfte sein (für einen ersten Überblick s. Raley et al., 2022).

6.5 Fazit

Ziel dieser Arbeit war die Evaluation eines Tests zur Erfassung des Wissens – und ergänzend auch der Fehlannahmen – über Diversität von (angehenden) Lehrkräften. Damit einhergehend wurden Zusammenhänge zu unterrichtsrelevanten Variablen im Kontext schulischer Inklusion erforscht. Aus den drei Studien mit unabhängigen Stichproben aus Studierenden und Lehrkräften lässt sich folgendes Fazit ziehen: Der DiWi(-K) erwies sich als ökonomisches und reliables Messinstrument, dessen Struktur mit den fünf Diversitätsbereichen SPF ESE, SPF Lernen, Teilleistungsstörungen, Geschlecht und Hochbegabung und einem Generalfaktor Diversitätswissen durch die Daten gestützt werden

konnte. Damit ergänzt er die bisherigen Messinstrumente, die Diversität nur im Rahmen einzelner Items (Kunina-Habenicht et al., 2020; Wildbrett et al., 2022), als zusammenfassende Skala (König & Blömeke, 2009; Voss et al., 2011) oder in Bezug auf einen isolierten Diversitätsbereich (Ginsburg et al., 2021; Sousa et al., 2017) erfassen, um einen wichtigen Bestandteil des Professionswissens von Lehrkräften (in Anlehnung an das COACTIV-Modell zur professionellen Kompetenz, Baumert & Kunter, 2013a). Als Grundlage eines nomologischen Netzes konnten positive Zusammenhänge zu verschiedenen Lerngelegenheiten wie dem Studienabschnitt und dem Studienfach nachgewiesen werden, die sich an Befunde aus vorheriger Forschung anschließen (Heyder, Vaskova et al., 2018; Malva et al., 2020; Wildbrett et al., 2022). Außerdem korrelierte das Wissen der (angehenden) Lehrkräfte positiv mit weiteren Aspekten der professionellen Kompetenz wie den Einstellungen und Selbstwirksamkeitserwartungen und fügte sich damit ebenfalls wie erwartet in Befunde vorheriger Forschung ein (Ghanizadeh et al., 2006; Lu et al., 2020; Scitutto et al., 2000). Dass dieses Wissen darüber hinaus für den Unterricht im Kontext schulischer Inklusion von Bedeutung ist und die bereichsspezifische Auswertung sinnvoll sein kann, zeigte sich an Zusammenhängen zwischen dem Gesamtwissen über Diversität und den selbstberichteten lernorientierten Unterrichtsinstruktionen und zwischen dem Wissen über den speziellen SPF ESE und der sozialen Partizipation und den akademischen Leistungen aller Schüler:innen in Klassen mit und ohne Schüler:innen mit SPF ESE. Vor allem die letztgenannten Befunde konnten in dieser Form anhand der hier vorgestellten Studien erstmalig nachgewiesen werden. Mit der Veröffentlichung der Testbögen (Steinmayr et al., 2022) stehen die Items nun anderen Wissenschaftler:innen und auch Praktiker:innen zur Verfügung, die den Test nutzen und einen Beitrag zu Inklusion leisten möchten. Für die Forschung könnten sich zukünftige Studien zum Beispiel den Zusammenhängen zwischen dem Wissen und dem tatsächlichen Unterrichtsgeschehen, der bereichsspezifischen

Untersuchung von Wissen und Fehlannahmen und der Erklärung von Varianz am Unterrichtserfolg neben anderen Wissensfacetten widmen. In der Praxis könnten die gewonnenen Erkenntnisse für die Konzeption von Lehrveranstaltungen und Weiterbildungsmaßnahmen genutzt werden. Dazu bieten sich Schwerpunkte wie die Verbindung von Wissen, Einstellungen und Selbstwirksamkeitserwartungen, die Umsetzung des deklarativen Faktenwissens in tatsächliches Unterrichtshandeln und der Abbau von Fehlannahmen an. Damit kann ein Beitrag geleistet werden, um Schule und Unterricht und damit auch die Teilhabe an der Gesellschaft gerechter und inklusiver zu gestalten.

Literaturverzeichnis II

- Aeppli, J., & Lötscher, H. (2016). EDAMA-Ein Rahmenmodell für Reflexion. *Beiträge zur Lehrerinnen- und Lehrerbildung*, 34(1), 78–97. <https://doi.org/10.36950/bzl.34.2016.9540>
- Baumert, J., & Kunter, M. (2013a). The COACTIV model of teachers' professional competence. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss & M. Neubrand (Eds.), *Cognitive activation in the mathematics classroom and professional competence of teachers: Results from the COACTIV project* (pp. 29–53). Springer. <https://doi.org/10.1007/978-1-4614-5149-5>
- Baumert, J., & Kunter, M. (2013b). The effect of content knowledge and pedagogical content knowledge on instructional quality and student achievement. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss & M. Neubrand (Eds.), *Cognitive activation in the mathematics classroom and professional competence of teachers: Results from the COACTIV project* (pp. 175–205). Springer. https://doi.org/10.1007/978-1-4614-5149-5_9
- Beauftragter der Bundesregierung für die Belange von Menschen mit Behinderungen. (2018). *UN-Behindertenrechtskonvention*. Bundesministerium für Arbeit und Soziales. https://www.behindertenbeauftragter.de/SharedDocs/Downloads/DE/AS/PublikationenErklaerungen/Broschuere_UNKonvention_KK.pdf?__blob=publicationFile&v=8
- Beck, E., Baer, M., Guldemann, T., Bischoff, S., Brühwiler, C., Müller, P., Niedermann, R., Rogalla, M., & Vogt, F. (Eds.) (2008). *Adaptive Lehrkompetenz. Analyse und Struktur, Veränderung und Wirkung handlungssteuernden Lehrerwissens*. Waxmann.
- Braksiek, M. (2022). Pre-service physical education teachers' attitude toward, and self-efficacy in, inclusive physical education: Measurement invariance and influence factors. *Teaching and Teacher Education*, 109, 103547. <https://doi.org/10.1016/j.tate.2021.103547>
- Brühwiler, C. (Eds.) (2014). *Adaptive Lehrkompetenz und schulisches Lernen. Effekte handlungssteuernder Kognitionen von Lehrpersonen auf Unterrichtsprozesse und Lernergebnisse der Schülerinnen und Schüler*. Waxmann. <https://doi.org/10.1007/978-3-658-23187-3>
- Brünken, R., Münzer, S., & Spinath, B. (Eds.) (2019). *Pädagogische Psychologie – Lernen und Lehren*. Hogrefe. <https://doi.org/10.1026/02214-000>
- de Ayala, R. J. (2022). *The theory and practice of item response theory* (2nd ed.). Guilford Press.
- de Boer, A., Pijl, S. J., & Minnaert, A. (2011). Regular primary schoolteachers' attitudes

- towards inclusive education: A review of the literature. *International Journal of Inclusive Education*, 15(3), 331–353. <http://doi.org/10.1080/13603110903030089>
- Depaepe, F., & König, J. (2018). General pedagogical knowledge, self-efficacy and instructional practice: Disentangling their relationship in pre-service teacher education. *Teaching and Teacher Education*, 69, 177–190. <https://doi.org/10.1016/j.tate.2017.10.003>
- Dersch, A. S., Heyder, A., & Eitel, A. (2022). Exploring the nature of teachers' math-gender stereotypes: The math-gender misconception questionnaire. *Frontiers in Psychology*, 13, 1728. <https://doi.org/10.3389/fpsyg.2022.820254>
- Dicke, T., Parker, P. D., Holzberger, D., Kunina-Habenicht, O., Kunter, M., & Leutner, D. (2015). Beginning teachers' efficacy and emotional exhaustion: Latent changes, reciprocity, and the influence of professional knowledge. *Contemporary Educational Psychology*, 41, 62–72. <https://doi.org/10.1016/j.cedpsych.2014.11.003>
- Eitel, A., Prinz, A., Kollmer, J., Niessen, L., Russow, J., Ludäscher, M., Renkl, A., & Lindner, M. A. (2021). The misconceptions about multimedia learning questionnaire: An empirical evaluation study with teachers and student teachers. *Psychology Learning & Teaching*, 20(3), 420–444. <https://doi.org/10.1177/14757257211028723>
- Embretson, S. E. & Reise, S. P. (2000). *Item response theory for psychologists*. Lawrence Erlbaum Associates.
- Ernst, H. M., Wittwer, J., & Voss, T. (2023). Do they know what they know? Accuracy in teacher candidates' self-assessments and its influencing factors. *British Educational Research Journal*, 00, 1–25. <https://doi.org/10.1002/berj.3860>
- Faulstich-Wieland, H. (2011). *Umgang mit Heterogenität und Differenz*. Schneider. <https://doi.org/10.14361/9783839429617-003>
- Fernandes, P. R. D. S., Jardim, J., & Lopes, M. C. D. S. (2021). The soft skills of special education teachers: Evidence from the literature. *Education Sciences*, 11(3), 125. <https://doi.org/10.3390/educsci11030125>
- Forlin, C., & Chambers, D. (2011). Teacher preparation for inclusive education: Increasing knowledge but raising concerns. *Asia-Pacific Journal of Teacher Education*, 39(1), 17–32. <https://doi.org/10.1080/1359866X.2010.540850>
- Gess-Newsome, J., Taylor, J. A., Carlson, J., Gardner, A. L., Wilson, C. D., & Stuhlsatz, M. A. (2017). Teacher pedagogical content knowledge, practice, and student achievement. *International Journal of Science Education*, 41(7), 944–963. <https://dx.doi.org/10.1080/09500693.2016.1265158>
- Ghanizadeh, A., Bahredar, M. J., & Moeini, S. R. (2006). Knowledge and attitudes towards attention deficit hyperactivity disorder among elementary school teachers. *Patient*

- Education and Counseling*, 63(1-2), 84–88. <https://doi.org/10.1016/j.pec.2005.09.002>
- Ginsburg, G. S., Pella, J. E., Ogle, R. R., DeVito, A., Raguin, K., & Chan, G. (2021). Teacher knowledge of anxiety and use of anxiety reduction strategies in the classroom. *Journal of Psychologists and Counsellors in Schools*, 1–11. <https://doi.org/10.1017/jgc.2021.26>
- Göransson, K., & Nilholm, C. (2014). Conceptual diversities and empirical shortcomings – a critical analysis of research on inclusive education. *European Journal of Special Needs Education*, 29, 265–280. <https://doi.org/10.1080/08856257.2014.933545>
- Grosche, M. (2015). Was ist Inklusion? Ein Diskussions- und Positionsartikel zur Definition von Inklusion aus Sicht der empirischen Bildungsforschung. In P. Kuhl, P. Stanat, B. Lütje-Klose, C. Gresch, H. A. Pant & M. Prenzel (Eds.), *Inklusion von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf in Schulleistungserhebungen* (pp. 17–39). Springer.
- Haimour, A. I., & Obaidat, Y. F. (2013). School teachers' knowledge about autism in Saudi Arabia. *World Journal of Education*, 3(5), 45–56. <https://doi.org/10.5430/wje.v3n5p45>
- Hartig, J., Frey, A., & Jude, N. (2020). Validität von Testwertinterpretationen. In H. Moosbrugger & A. Kelava (Eds.), *Testtheorie und Fragebogenkonstruktion* (pp. 529–545). Springer. https://doi.org/10.1007/978-3-662-61532-4_21
- Hasan, S., Bagayoko, D., & Kelley, E. L. (1999). Misconceptions and the certainty of response index (CRI). *Physics Education*, 34(5), 294–299. <https://doi.org/10.1088/0031-9120/34/5/304>
- Heyder, A., Bergold, S., & Steinmayr, R. (2018). Teachers' knowledge about intellectual giftedness: A first look at levels and correlates. *Psychology Learning & Teaching*, 17(1), 27–44. <https://doi.org/10.1177/1475725717725493>
- Heyder, A., Vaskova, A., Hußmann, A., & Steinmayr, R. (2018). Wissen von angehenden Lehrkräften zu Diversität im Kontext schulischer Inklusion – Die Entwicklung eines Wissenstests und erste Ergebnisse. In S. Hußmann & B. Welzel (Eds.), *DoProfiL – Das Dortmunder Profil für inklusionsorientierte Lehrerinnen- und Lehrerbildung* (pp. 267–278). Waxmann.
- Hohenstein, F., Kleickmann, T., Zimmermann, F., Köller, O., & Möller, J. (2017). Erfassung von pädagogischem und psychologischem Wissen in der Lehramtsausbildung: Entwicklung eines Messinstruments. *Zeitschrift für Pädagogik*, 63, 91–112.
- Hohenstein, F., Zimmermann, F., Kleickmann, T., Köller, O., & Möller, J. (2014). Sind die bildungswissenschaftlichen Standards für die Lehramtsausbildung in den Curricula der Hochschulen angekommen. *Zeitschrift für Erziehungswissenschaft*, 17(3), 497–507. <https://doi.org/10.1007/s11618-014-0563-9>
- Horn, W., Lukesch, H., Mayrhofer, S., & Kormann, A. (2004). *PSB-R 6-13*,

- Prüfbildungssystem für Schul- und Bildungsberatung für 6. bis 13. Klassen (rev. ed.)*. Hogrefe.
- Kelava, A., & Moosbrugger, H. (2020). Deskriptivstatistische Itemanalyse und Testwertbestimmung. In H. Moosbrugger & A. Kelava (Eds.), *Testtheorie und Fragebogenkonstruktion* (pp. 143–158). Springer. https://doi.org/10.1007/978-3-662-61532-4_7
- Keller, M. M., Neumann, K., & Fischer, H. E. (2017). The impact of physics teachers' pedagogical content knowledge and motivation on students' achievement and interest. *Journal of Research in Science Teaching*, 54(5), 586–614. <https://doi.org/10.1002/tea.21378>
- Kendeou, P., & O'Brien, E. J. (2014). The knowledge revision components (KReC) framework: Processes and mechanisms. In D. N. Rapp & J. L. G. Braasch (Eds.), *Processing inaccurate information: Theoretical and applied perspectives from cognitive science and the educational sciences* (pp. 353–377). MIT Press. <https://doi.org/10.7551/mitpress/9737.003.0022>
- König, J., & Blömeke, S. (2009). Pädagogisches Wissen von angehenden Lehrkräften. *Zeitschrift für Erziehungswissenschaft*, 12(3), 499–527. <https://doi.org/10.1007/s11618-009-0085-z>
- König, J., & Blömeke, S. (2010). *Pädagogisches Unterrichtswissen (PUW) – Dokumentation der Kurzfassung des TEDS-M Testinstruments zur Kompetenzmessung in der ersten Phase der Lehrerausbildung*. Humboldt-Universität Berlin.
- König, J., & Pflanzl, B. (2016). Is teacher knowledge associated with performance? On the relationship between teachers' general pedagogical knowledge and instructional quality. *European Journal of Teacher Education*, 39(4), 419–436. <https://doi.org/10.1080/02619768.2016.1214128>
- König, J., Gerhard, K., Melzer, C., Rühl, A. M., Zenner, J., & Kaspar, K. (2017). Erfassung von pädagogischem Wissen für inklusiven Unterricht bei angehenden Lehrkräften: Testkonstruktion und Validierung [Assessing general pedagogical knowledge for inclusive teaching (GPK-IT) among pre-service teachers: Test construction and validation. *Unterrichtswissenschaft*, 45(4), 223–242.
- Kunina-Habenicht, O., Maurer, C., Wolf, K., Holzberger, D., Schmidt, M., Dicke, T., Teuber, Z., Koc-Januchta, M., Lohse-Bossenz, H., Leutner, D., Seidel, T., & Kunter, M. (2020). Der BilWiss-2.0-Test: Ein revidierter Test zur Erfassung des bildungswissenschaftlichen Wissens von (angehenden) Lehrkräften. *Diagnostica*, 66(2), 80–92. <https://doi.org/10.1026/0012-1924/a000238>
- Kunter, M., Kleickmann, T., Klusmann, U., & Richter, D. (2013). In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss & M. Neubrand (Eds.), *The development of teachers'*

- professional competence* (pp. 63–77). Springer. <https://doi.org/10.1007/978-1-4614-5149-5>
- Kunter, M., Klusmann, U., Baumert, J., Richter, D., Voss, T., & Hachfeld, A. (2013). Professional competence of teachers: effects on instructional quality and student development. *Journal of Educational Psychology, 105*(3), 805–820. <https://doi.org/10.1037/a0032583>
- Kurniawati, F., de Boer, A. A., Minnaert, A. E. M. G., & Mangunsong, F. (2014). Characteristics of primary teacher training programmes on inclusion: A literature focus. *Educational Research, 56*(3), 310–326. <https://doi.org/10.1080/00131881.2014.934555>
- Kuyini, A. B., & Desai, I. (2007). Principals' and teachers' attitudes and knowledge of inclusive education as predictors of effective teaching practices in Ghana. *Journal of Research in Special Educational Needs, 7*(2), 104–113. <https://doi.org/10.1111/j.1471-3802.2007.00086.x>
- Lauermann, F., & König, J. (2016). Teachers' professional competence and wellbeing – Understanding the links between general pedagogical knowledge, self-efficacy and burnout. *Learning and Instruction, 45*, 9–19. <http://doi.org/10.1016/j.learninstruc.2016.06.006>
- Lautenbach, F., & Heyder, A. (2019). Changing attitudes to inclusion in preservice teacher education: A systematic review. *Educational Research, 61*(2), 231–253. <https://doi.org/10.1080/00131881.2019.1596035>
- Lu, M., Zou, Y., Chen, X., Chen, J., He, W., & Pang, F. (2020). Knowledge, attitude and professional self-efficacy of Chinese mainstream primary school teachers regarding children with autism spectrum disorder. *Research in Autism Spectrum Disorders, 72*, 101513. <https://doi.org/10.1016/j.rasd.2020.101513>
- Lüke, T., & Grosche, M. (2018). What do I think about inclusive education? It depends on who is asking. Experimental evidence for a social desirability bias in attitudes towards inclusion. *International Journal of Inclusive Education, 22*(1), 38–53. <https://doi.org/10.1080/13603116.2017.1348548>
- Malva, L., Leijen, Ä., & Baucal, A. (2020). Towards measuring teachers' general pedagogical knowledge – A mixed method investigation of a pilot test. *Studies in Educational Evaluation, 64*, 100815. <https://doi.org/10.1016/j.stueduc.2019.100815>
- Mintz, J. (2022). The role of universities and knowledge in teacher education for inclusion. *International Journal of Inclusive Education, 26*, 1–11. <https://doi.org/10.1080/13603116.2022.2081877>
- Mintz, J., Hick, P., Solomon, Y., Matziari, A., Ó'Murchú, F., Hall, K., Cahill, K., Curtin, C., Anders, J., & Margariti, D. (2020). The reality of reality shock for inclusion: How does teacher attitude, perceived knowledge and self-efficacy in relation to effective inclusion

- in the classroom change from the pre-service to novice teacher year?. *Teaching and Teacher Education*, 91, 103042. <https://doi.org/10.1016/j.tate.2020.103042>
- Moosbrugger, H., Gäde, J. C., Schermelleh-Engel, K., & Rauch, W. (2020). Klassische Testtheorie (KTT). In H. Moosbrugger & A. Kelava (Eds.), *Testtheorie und Fragebogenkonstruktion* (pp. 275–304). Springer. https://doi.org/10.1007/978-3-662-61532-4_13
- Moosbrugger, H., Schermelleh-Engel, K., Gäde, J. C., & Kelava, A. (2020). Testtheorien im Überblick. In H. Moosbrugger & A. Kelava (Eds.), *Testtheorie und Fragebogenkonstruktion* (pp. 251–273). Springer. https://doi.org/10.1007/978-3-662-61532-4_12
- Paju, B., Rätty, L., Pirttimaa, R., & Kontu, E. (2016). The school staff's perception of their ability to teach special educational needs pupils in inclusive settings in Finland. *International Journal of Inclusive Education*, 20(8), 801–815. <https://doi.org/10.1080/13603116.2015.1074731>
- Pflanzl, B., Thomas, A., & Matischek-Jauk, M. (2013). Pädagogisches Wissen und pädagogische Handlungskompetenz. *Erziehung und Unterricht*, 163(1/2), 1–8.
- Putnick, D. L., & Bornstein, M. H. (2016). Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. *Developmental Review*, 41, 71–90. <https://dx.doi.org/10.1016/j.dr.2016.06.004>
- Raley, S. K., Shogren, K. A., & Hagiwara, M. (2022). Review of Existing Research on Coaching in Inclusive, Secondary Classrooms. *Education and Training in Autism and Developmental Disabilities*, 57(2), 151–166.
- Ramberg, J., Lénárt, A., & Watkins, A. (Eds.) (2018). *European Agency Statistics on Inclusive Education: 2016 Dataset Cross-Country Report*. European Agency for Special Needs and Inclusive Education. <https://www.european-agency.org/resources/publications/european-agency-statistics-inclusive-education-2016-dataset-cross-country>
- Retelsdorf, J., Butler, R., Streblov, L., & Schiefele, U. (2010). Teachers' goal orientations for teaching: Associations with instructional practices, interest in teaching, and burnout. *Learning and Instruction*, 20, 30–46. <https://doi.org/10.1016/j.learninstruc.2009.01.001>
- Sanz-Cervera, P., Fernández-Andrés, M. I., Pastor-Cerezuela, G., & Tárraga-Mínguez, R. (2017). Pre-service teachers' knowledge, misconceptions and gaps about autism spectrum disorder. *Teacher Education and Special Education*, 40(3), 212–224. <https://doi.org/10.1177/0888406417700963>
- Savolainen, H., Malinen, O. P., & Schwab, S. (2022). Teacher efficacy predicts teachers' attitudes towards inclusion—a longitudinal cross-lagged analysis. *International Journal of Inclusive Education*, 26(9), 958–972. <https://doi.org/10.1080/13603116.2020.1752826>

- Schiefele, U. (2017). Classroom management and mastery-oriented instruction as mediators of the effects of teacher motivation on student motivation. *Teaching and Teacher Education, 64*, 115–126. <https://doi.org/10.1016/j.tate.2017.02.004>
- Schmiedeler, S. (2013). Wissen und Fehlannahmen von deutschen Lehrkräften über die Aufmerksamkeitsdefizit-/Hyperaktivitätsstörung (ADHS). *Psychologie in Erziehung und Unterricht, 60*, 143–153. <https://doi.org/10.2378/peu2013.art12d>
- Schwab, S., Hellmich, F., & Görel, G. (2017). Self-efficacy of prospective Austrian and German primary school teachers regarding the implementation of inclusive education. *Journal of Research in Special Educational Needs, 17*(3), 205–217. <https://doi.org/10.1111/1471-3802.12379>
- Sciutto, M. J., Terjesen, M. D., & Bender Frank, A. S. (2000). Teachers' knowledge and misperceptions of attention-deficit/hyperactivity disorder. *Psychology in the Schools, 37*(2), 115–122. [https://doi.org/10.1002/\(sici\)1520-6807\(200003\)37:2<115::aid-pits3>3.0.co;2-5](https://doi.org/10.1002/(sici)1520-6807(200003)37:2<115::aid-pits3>3.0.co;2-5)
- Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2014, June). *Standards für die Lehrerbildung: Bildungswissenschaften*. http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2004/2004_12_16-Standards-Lehrerbildung-Bildungswissenschaften.pdf
- Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2022, January). *Sonderpädagogische Förderung in Schulen 2011 bis 2020*. https://www.kmk.org/fileadmin/Dateien/pdf/Statistik/Dokumentationen/Dok231_SoPaeFoe_2020.pdf
- Sherman, S. W. (1976, April). *Multiple choice test bias uncovered by use of an "I don't know" alternative*. [Paper presentation]. Annual Meeting of the American Educational Research Association, San Francisco.
- Sousa, P., Dias, P. C., & Cadime, I. (2017). Predictors of primary school teachers' knowledge about developmental dyscalculia. *European Journal of Special Needs Education, 32*(2), 204–220. <https://doi.org/10.1080/08856257.2016.1216635>
- Steinmayr, R., Heyder, A., & Tometten, L. (2022). DiWi. Test zum Wissen über verschiedene Diversitätsbereiche von (angehenden) Lehrkräften. Verfahrensdokumentation, Fragebogen Lang- und Kurzversion, Auswertungssyntax. In Leibniz-Institut für Psychologie (ZPID) (Ed.), *Open Test Archive*. ZPID. <https://doi.org/10.23668/psycharchives.5136>
- Sulek, R., Trembath, D., Paynter, J., & Keen, D. (2019). Empirically supported treatments for students with autism: General education teacher knowledge, use, and social validity ratings. *Developmental Neurorehabilitation, 22*(6), 380–389.

<https://doi.org/10.1080/17518423.2018.1526224>

- Tan, R., & Amrhein, B. (2019). Impact of training on preservice teachers' sense of self-efficacy to implement inclusive teaching in the English language classroom. *Herausforderung Lehrer_innenbildung : Zeitschrift zur Konzeption, Gestaltung und Diskussion*, 2(3), 365–382. <https://doi.org/10.4119/hlz-2474>
- Tempel, T. (2022). Asking about inclusion: Question order and social desirability influence measures of attitudes towards inclusive education. *European Journal of Special Needs Education*, 1–7. <https://doi.org/10.1080/08856257.2023.2162666>
- United Nations Educational, Scientific and Cultural Organization. (1994). *The Salamanca statement and framework for action on special needs education*. <https://unesdoc.unesco.org/ark:/48223/pf0000098427>
- Voss, T. (2022). Not useful to inform teaching practice? Student teachers hold skeptical beliefs about evidence from education science. *Frontiers in Education*, 7, 976791. <https://doi.org/10.3389/feduc.2022.976791>
- Voss, T., Kunina-Habenicht, O., Hoehne, V., & Kunter, M. (2015). Stichwort Pädagogisches Wissen von Lehrkräften: Empirische Zugänge und Befunde. *Zeitschrift für Erziehungswissenschaft*, 18(2), 187–223. <https://doi.org/10.1007/s11618-015-0626-6>
- Voss, T., Kunter, M., & Baumert, J. (2011). Assessing teacher candidates' general pedagogical/psychological knowledge: Test construction and validation. *Journal of Educational Psychology*, 103(4), 952–969. <https://doi.org/10.1037/a0025125>
- Voss, T., Zachrich, L., Fauth, B., & Wittwer, J. (2022). The same yet different? Teaching quality differs across a teacher's classes, but teachers with higher knowledge make teaching quality more similar. *Learning and Instruction*, 80, 101614. <https://doi.org/10.1016/j.learninstruc.2022.101614>
- Wildbrett, J., Lohse-Bossenz, H., & Dörfler, T. (2022). Empirische Arbeit: Vignettenbasierte Messung pädagogisch-psychologischen Professionswissens bei Lehramtsstudierenden. *Psychologie in Erziehung und Unterricht*, 69(4), 278–291. <https://doi.org/10.2378/peu2022.art20d>
- Wischer, B. (2007). Wie sollen LehrerInnen mit Heterogenität umgehen? Über 'programmatische Fallen' im aktuellen Reformdiskurs. *Die deutsche Schule*, 99(4), 422–433.
- Wray, E., Sharma, U., & Subban, P. (2022). Factors influencing teacher self-efficacy for inclusive education: A systematic literature review. *Teaching and Teacher Education*, 117, 103800. <https://doi.org/10.1016/j.tate.2022.103800>
- Yada, A., Leskinen, M., Savolainen, H., & Schwab, S. (2022). Meta-analysis of the relationship between teachers' self-efficacy and attitudes toward inclusive education.

- Teaching and Teacher Education*, 109, 103521.
<https://doi.org/10.1016/j.tate.2021.103521>
- Yin, L., Joshi, R. M., & Yan, H. (2019). Knowledge about dyslexia among early literacy teachers in China. *Dyslexia*, 1–19. <https://doi.org/10.1002/dys.1635>
- Zengilowski, A., Schuetze, B. A., Nash, B. L., & Schallert, D. L. (2021). A critical review of the refutation text literature: Methodological confounds, theoretical problems, and possible solutions. *Educational Psychologist*, 56(3), 175–195.
<https://doi.org/10.1080/00461520.2020.1861948>
- Zulliger, S. & Tanner, S. (2013). Der Begriff Heterogenität in empirischen Studien. *Schweizerische Zeitschrift für Bildungswissenschaften*, 35, 37–52.
<https://doi.org/10.24452/sjer.35.1.4900>

7 Anhang

7.1 Liste der Einzelarbeiten

Studie 1: Tometten, L., Heyder, A., & Steinmayr, R. (2022). Empirische Arbeit: Die Erfassung von Wissen über verschiedene Diversitätsbereiche im Kontext einer inklusionsorientierten Lehrkräftebildung. *Psychologie in Erziehung und Unterricht*, 69(4), 292–304. <https://doi.org/10.2378/peu2022.art23d>

Studie 2: Tometten, L., Heyder, A., & Steinmayr, R. (2023, May 4). Teachers' knowledge about diversity: Correlations with attitudes, self-efficacy, and instructions. *OSF Preprints*. <https://doi.org/10.31219/osf.io/qtfn7>

Studie 3: Tometten, L., Heyder, A., & Steinmayr, R. (2021). Links between teachers' knowledge about special educational needs and students' social participation and academic achievement in mainstream classes. *Contemporary Educational Psychology*, 67, 102022. <https://doi.org/10.1016/j.cedpsych.2021.102022>