OPEN ACCESS
OPUS AUGSBURG

Universität Augsburg
Universitätsbibliothek

Herkunftsbezogene Einstellungen und Stereotype Lehramtsstudierender: eine Interventionsstudie

Anita Tobisch, Sabine K. Lehmann-Grube, Lara Shames, Markus Dresel

Angaben zur Veröffentlichung / Publication details:

Tobisch, Anita, Sabine K. Lehmann-Grube, Lara Shames, and Markus Dresel. 2022. "Herkunftsbezogene Einstellungen und Stereotype Lehramtsstudierender: eine Interventionsstudie." In Stereotype in der Schule II: Ursachen und Möglichkeiten der Intervention, edited by Sabine Glock, 301–40. Wiesbaden: Springer VS. https://doi.org/10.1007/978-3-658-37262-0_8.

Nutzungsbedingungen / Terms of use: **licgercopyright**

Dieses Dokument wird unter folgenden Bedingungen zur Verfügung gestellt: / This document is made available under these conditions:
Deutsches Urheberrecht
Weitere Informationen finden Sie unter: / For more information see:
https://www.uni-augsburg.de/de/organisation/bibliothek/publizieren-zitieren-archivieren/publiz/

Herkunftsbezogene Einstellungen und Stereotype Lehramtsstudierender: Eine Interventionsstudie

8

Anita Tobisch, Sabine K. Lehmann-Grube, Lara Shames und Markus Dresel

8.1 Herkunftsassoziierte Lehrkrafturteile

Der Bildungsweg und -erfolg von Schüler*innen hängt weiterhin stark mit ihrer ethnisch-kulturellen sowie mit ihrer sozialen Herkunft zusammen (z. B. Organisation für wirtschaftliche Zusammenarbeit und Entwicklung [OECD], 2019). Dabei verweisen theoretische Modelle sowie empirische Untersuchungen auf die Bedeutung von Lehrkrafturteilen in diesem Zusammenhang. Sowohl experimentelle Untersuchungen (z. B. Glock et al., 2013) als auch Feldstudien (z. B. Rubie-Davies et al., 2006) zeigen regelmäßig, dass Lehrkräfte Schüler*innen mit Migrationshintergrund oder niedrigem sozialem Status im Vergleich zu Schüler*innen ohne Migrationshintergrund oder hohem sozialem Status – auch bei vergleichbaren Leistungen – dennoch schlechtere Noten geben und seltener die Empfehlung für ein Gymnasium erteilen. Einige Studien, in

A. Tobisch (✉) · S. K. Lehmann-Grube · L. Shames · M. Dresel
Lehrstuhl für Psychologie, Universität Augsburg, Augsburg, Bayern, Deutschland
E-Mail: anita.tobisch@phil.uni-augsburg.de

S. K. Lehmann-Grube
E-Mail: sabine.lehmann-grube@phil.uni-augsburg.de

L. Shames
E-Mail: lara.shames@student.uni-augsburg.de

M. Dresel
E-Mail: markus.dresel@phil.uni-augsburg.de

denen Herkunftseffekte eines Migrationshintergrundes für Herkunftseffekte des sozio-ökonomischen Hintergrundes kontrolliert werden, deuten dabei an, dass der sozio-ökonomische Hintergrund den Ausschlag für die Verzerrungen von Lehrkrafturteilen gibt. Berichtet wurden sogar Effekte positiver Diskriminierung (z. B. Diefenbach, 2010; Gresch & Becker, 2010; Maaz et al., 2011; Ülger et al., 2016), das heißt positive Urteilsverzerrung für Schüler*innen mit (türkischem) Migrationshintergrund. Dennoch verbleibt in einigen Untersuchungen auch ein eigenständiger Effekt des Migrationshintergrunds (z. B. Tobisch & Dresel, 2017). Insgesamt ist somit eine heterogene Befundlage zu konstatieren, aus der sich allerdings ableiten lässt, dass die Herkunft von Schüler*innen einen zwar nicht immer einheitlichen, aber dennoch bedeutsamen Einfluss auf Lehrkrafturteile haben kann.

Die Beurteilung von Schüler*innen im Hinblick auf lernrelevante Merkmale und Leistungen ist eine Kernaufgabe von Lehrkräften und sollte unabhängig von Herkunft, Geschlecht und weiteren, im Prinzip leistungsirrelevanten Merkmalen ablaufen. Somit stellt die diagnostische Kompetenz eine Kernkompetenz von Lehrkräften dar, die allerdings bei Lehrkräften durchaus verschieden und auch je nach Urteilsbereich mehr oder weniger gut ausgeprägt ist (z. B. Artelt & Gräsel, 2009; Spinath, 2005; Südkamp et al., 2012). Im Modell der Akkuratheit von Lehrkrafturteilen postulieren Südkamp et al. (2012), dass neben Urteils- und Testmerkmalen (z. B. spezifische versus globale Urteile; konkrete Aufgaben versus allgemeinere Fähigkeiten) auch Merkmale der Schüler*innen (z. B. Herkunft) und Merkmale der Lehrkräfte (z. B. Einstellungen und Stereotype) einen Einfluss auf die Urteilsakkuratheit haben können. Dabei ist davon auszugehen, dass Schüler*innenmerkmale nicht per se, sondern nur durch die Wahrnehmung der Lehrkraft und ihre*seine mit den Merkmalen verbundenen mentalen Assoziationen und Bewertungen die Beurteilung der Schüler*innen beeinflussen. Aus der Forschungslage können aber keine einfachen unidirektionalen Zusammenhänge in der Art, je günstiger (positiver) das assoziierte Stereotyp und/oder die Einstellungen der Lehrkraft, umso günstiger (positiver) die Beurteilung, abgeleitet werden (vgl. Karst & Bonefeld, 2020). Ableiten lässt sich jedoch, dass Einstellungen und Stereotype von Lehrkräften in diesem Zusammenhang von besonderer Bedeutung zu sein scheinen und negative wie positive Einstellungen und Stereotype in Zusammenhang mit Urteilsverzerrungen stehen, wenn Schüler*innen als Mitglieder bestimmter Gruppen wahrgenommen und entsprechende mentale Assoziationen aktiviert werden (z. B. Glock et al., 2016; Karst & Bonefeld, 2020).

8.2 Herkunftsbezogene Einstellungen und Stereotype von Lehrkräften

Einstellungen sind nach Eagly und Chaiken (1993) mentale Tendenzen, die in einer mehr oder weniger positiven oder negativen kognitiven, affektiven und/oder behavioralen Bewertungsreaktion auf ein Einstellungsobjekt ihren Ausdruck finden. Im Gegensatz zu Einstellungen beschreiben Stereotype im Kern zunächst nur die generalisierte Zuschreibung bestimmter Merkmale zu spezifischen sozialen Gruppen, beinhalten aber nicht notwendigerweise eine bewertende Komponente (z. B. Kunda & Thagard, 1996; Tajfel, 1969). Bezogen auf den schulischen Kontext könnten Lehrkräfte Schüler*innen bestimmte positive, neutrale oder negative Merkmale in Abhängigkeit von der Zugehörigkeit zu bestimmten Gruppen zuschreiben. So könnten sie zum Beispiel das Stereotyp haben, dass Schüler*innen aus Familien mit einem hohen sozialen Status typischerweise fleißig wären. Hierbei würden Lehrkräfte dieser Schüler*innengruppe ein positives Merkmal zuschreiben. Sie könnten aber auch eine tendenziell positivere Einstellung dieser Schüler*innengruppe gegenüber aufweisen und es beispielsweise als angenehmer empfinden, diese Schüler*innen zu unterrichten. Wie diese Beispiele bereits zeigen, können Stereotype und Einstellungen auch in Zusammenhang miteinander stehen. So könnten Lehrkräfte ihre positive Einstellung zu bestimmten Schüler*innengruppen auch aus der Verknüpfung mit den zugeschriebenen stereotypen Merkmalen ableiten. Eine Lehrkraft, die Schüler*innen einer bestimmten Gruppe im Allgemeinen für fleißiger hält, wird diesen eventuell auch eine positivere Einstellung gegenüber haben. Die Bedeutung dieser Einstellungen und Stereotype wird insbesondere dann deutlich, wenn berücksichtigt wird, dass sie auch Lehrkrafterwartungen und damit zum Beispiel die Unterrichtsgestaltung (z. B. Schrader & Helmke, 2001) und die Interaktionen mit Schüler*innen beeinflussen können, was sich dann durch Erwartungseffekte im Lernverhalten und den Leistungen von Schüler*innen niederschlagen kann (z. B. Brophy & Good, 1974; Jussim, 1990; Jussim & Harber, 2005). Theorien zur internen Struktur von Einstellungen schreiben Überzeugungen, beispielsweise hinsichtlich der Attribute von Einstellungsobjekten, eine Funktion für die Bildung und/oder den Abruf von Einstellungen zu, insbesondere das Erwartungs-Wert-Modell (Michigan Gruppe, u. a. Peak, 1955), das in Anlehnung an Erwartungs-Wert-Modelle der Motivationspsychologie entwickelt wurde, und das Informations-Integrations-Modell (Anderson, 1971). In Studien auf der Grundlage dieser Theorien werden dann Einstellungen meist mittels Überzeugungsaussagen erhoben (Erörterung in Eagly & Chaiken, 1993, Kap. 3).

Einstellungen können sowohl implizit als auch explizit vorliegen (z. B. Greenwald & Banaji, 1995). Implizite Einstellungen werden nach Greenwald und Banaji (1995) definiert als „introspectively unidentified (or inaccurately identified) traces of past experience that mediate favorable or unfavorable feeling, thought, or action towards social objects“ (S. 8). In der Regel sind implizite Einstellungen Individuen weniger bewusst zugänglich, spontaner und schwieriger zu kontrollieren als explizite Einstellungen; zudem existieren unterschiedliche theoretische Annahmen darüber, durch welche Faktoren sich implizite und explizite Einstellungen unterscheiden, zum Beispiel durch den Grad der Zugänglichkeit oder die Kontrolle expliziter, aber nicht impliziter Einstellungen durch soziale Erwünschtheit (vgl. Fazio, 2007; Gawronski & Bodenhausen, 2006; Goedderz & Hahn, 2020). Zahlreiche empirische Untersuchungen verweisen dabei auf nur geringe oder sogar keine Zusammenhänge zwischen impliziten und expliziten Einstellungen (z. B. die Meta-Analyse von Hofmann et al., 2005). Allerdings berichteten Hahn et al. (2014) in einer Reihe experimenteller Studien, dass Versuchspersonen die Ergebnisse der Messung ihrer impliziten Einstellungen akkurat vorhersagen konnten, bevor diese erhoben worden waren, was gegen die Erklärung der Unzugänglichkeit impliziter Einstellungen spricht. Hierbei könnte aber auch diskutiert werden, dass kognitive Kapazitäten für die Reflexion der eigenen impliziten Einstellungen notwendig sind, die in bestimmten Situationen, wie beispielsweise im schulischen Alltag und bei der Beurteilung von Schüler*innen, oft nicht zur Verfügung stehen. So zeigten Friese et al. (2008), dass implizite Einstellungen insbesondere spontane Verhaltensweisen vorhersagen können. Zwar lassen sich einige Unterschiede zwischen impliziten und expliziten Einstellungen finden, allerdings auch Verweise auf die Bedeutsamkeit beider Einstellungsfacetten, insbesondere für Urteilsprozesse (z. B. Glock et al., 2016).

Empirische Untersuchungen der Einstellungen von Lehrkräften verweisen häufig auf positivere Einstellungen zu Schüler*innen ohne Migrationshintergrund oder mit hohem sozialem Status im Vergleich zu Schüler*innen mit Migrationshintergrund oder niedrigem sozialem Status.[1] Glock und Karbach (2015) nutzten

[1] In den folgend referierten Studien wurden meist Schüler*innen mit türkischem Migrationshintergrund fokussiert. Es ist davon auszugehen, dass Lehrkräfte gegenüber Schüler*innen aus anderen Herkunftsgruppen verschiedene Einstellungen und Stereotype bilden, dass sich also angesichts der zunehmenden Heterogenität der Herkunftsgruppen differenzierte Einstellungsmuster zeigen.

in einer experimentellen Studie mit Lehramtsstudierenden drei unterschiedliche implizite Messverfahren und konnten dabei zeigen, dass die Versuchspersonen positivere implizite Einstellungen zu Schüler*innen ohne Migrationshintergrund aufwiesen. Im Vergleich dazu zeigten sich für Schüler*innen mit Migrationshintergrund neutrale bis negative implizite Einstellungen. Auch Kleen et al. (2019) konnten bestätigen, dass die impliziten Einstellungen Lehramtsstudierender für Schüler*innen ohne Migrationshintergrund positiver ausfallen. Allerdings wurde hier zudem die eigene Herkunft der Studierenden berücksichtigt. Die Befunde verweisen auf eine Eigengruppenpräferenz. So zeigten sich bei Lehramtsstudierenden mit türkischem Migrationshintergrund positivere implizite Einstellungen gegenüber Personen der eigenen Herkunftsgruppe. Für die expliziten Einstellungen zeigte sich hingegen kein Effekt der eigenen Herkunft. Die expliziten Einstellungen wurden jedoch nur gegenüber Schüler*innen mit türkischem Migrationshintergrund erfasst, sodass kein Vergleich mit den Einstellungen gegenüber Schüler*innen ohne Migrationshintergrund möglich war. Für stereotype Zuschreibungen zeigten sich für Schüler*innen aus Familien mit einem hohen sozialen Status positivere Eigenschaftszuschreibungen als für Schüler*innen aus Familien mit niedrigem sozialem Status (z. B. Tobisch & Dresel, 2020). Bonefeld und Karst (2020) berichteten zudem, dass angehende Lehrkräfte zwar nicht im Allgemeinen negativere Stereotype mit Personen mit türkischem Migrationshintergrund verbinden, in Bezug auf leistungsrelevante Aspekte zeigten sich aber mehr negative Stereotype für Personen mit Migrationshintergrund und mehr positive Stereotype für Personen ohne Migrationshintergrund.

Wie bereits dargestellt wurde, werden Einstellungen und Stereotype als potenzielle Einflussfaktoren auf Urteile und Verhaltensweisen von Lehrkräften diskutiert. So wird zum Teil angenommen, dass die Bevorzugung oder die Benachteiligung bestimmter Schüler*innen (z. B. im Hinblick auf ihren Migrationshintergrund oder ihre soziale Herkunft) durch besonders positive oder negative Einstellungen der Lehrkräfte entstehen. Als theoretisches Rahmenmodell wird hierbei häufig das *Kontinuum-Modell der Eindrucksbildung* von Fiske und Neuberg (1990) herangezogen. Das Modell beschreibt dabei eine kategoriegeleitete (automatisiertere) und eine individuumsgeleitete (kontrolliertere) Informationsverarbeitung als die Endpunkte eines Kontinuums, die auch als verschiedene Modi der Informationsverarbeitung verstanden werden. Dabei ist zunächst von einer kategoriebasierten Urteilsbildung auszugehen, bei der schrittweise mehr und mehr individuelle Informationen berücksichtigt werden können, wenn zum Beispiel Motivation und Ressourcen zur Verarbeitung zur Verfügung stehen, aber auch in Abhängigkeit von der Konformität der Informationen mit bestehenden Stereotypen. Böhmer et al. (2017) verweisen beispielsweise

darauf, dass stereotypinkonsistente Informationen eine individuumsbasierte Informationsverarbeitung von Lehrkräften begünstigen. Einstellungen und Stereotype sind demnach insbesondere dann von Bedeutung für die Urteilsbildung von Lehrkräften, wenn die zur Verfügung stehenden Informationen mit bestehenden Stereotypen konform sind und keine ausreichenden Kapazitäten und Motivation zur vertiefteren Informationsverarbeitung vorliegen. Werden Stereotype über typische lernrelevante Verhaltensweisen von Schüler*innen mit bestimmten Herkunftsmerkmalen als verallgemeinerte Beurteilung dieser Schüler*innen verstanden, könnte eine Intervention, die zur individuums- statt kategorienbasierten Urteilsbildung anregt, unter Umständen auf diese verallgemeinerten Beurteilungen im Sinne von Stereotypen wirken und diese verändern.

Glock et al. (2016) konnten in einer experimentellen Untersuchung mit Lehramtsstudierenden zeigen, dass besonders ausgeprägt positive implizite und explizite Einstellungen gegenüber Schüler*innen aus Familien mit einem hohen sozialen Status in einigen Urteilsbereichen im Zusammenhang mit einer positiveren Beurteilung dieser Schüler*innengruppe und zusätzlich im Zusammenhang mit einer negativeren Beurteilung von Schüler*innen aus Familien mit niedrigem sozialem Status standen. Auch Bonefeld und Dickhäuser (2018) berücksichtigten in einer experimentellen Untersuchung mit Lehramtsstudierenden die impliziten Einstellungen gegenüber Schüler*innen mit und ohne Migrationshintergrund bei der Beurteilung von Leistungen im Fach Deutsch. Dabei zeigte sich unter der Bedingung guter Schüler*innenleistungen kein bedeutsamer Einfluss der impliziten Einstellungen. Allerdings wurden Schüler*innen mit Migrationshintergrund bei schlechten Leistungen von Lehramtsstudierenden mit positiven impliziten Einstellungen zu dieser Gruppe schlechter bewertet als von Lehramtsstudierenden mit negativen impliziten Einstellungen. Als Erklärung für diesen zunächst unerwarteten Befund diskutierten die Autor*innen enttäuschte Erwartungen der Lehramtsstudierenden als mögliche Ursache; die schlechten Leistungen der Schüler*innen mit Migrationshintergrund könnten den positiven Erwartungen (im Einklang mit den positiven Einstellungen) widersprechen. Diese Befunde aus experimentellen Studien deuten zwar darauf hin, dass Einstellungen einen Einfluss auf Lehrkrafturteile haben können, dieser Zusammenhang aber durch weitere Faktoren beeinflusst beziehungsweise moderiert wird.

Ajzen und Fishbein (1980) postulierten in ihrer *Theory of Reasoned Action,* dass Einstellungen generell nicht direkt in bestimmte Verhaltensweisen münden. Unstrittig scheint aber, dass Einstellungen handlungsmotivierend wirken können, und auch Stereotypen wird eine Bedeutung für Verhaltenstendenzen zugeschrieben (z. B. Cuddy et al., 2008; Eagly & Chaiken, 1993). Einige Befunde

verdeutlichen die Bedeutung herkunftsassoziierter Einstellungen und Stereotype von Lehrkräften und verweisen zudem darauf, dass Lehrkräfte diese Einstellungen und Stereotype reflektieren und bei Bedarf verändern sollten (z. B. Glock et al., 2016). Mit anderen Worten, die Befunde zur positiveren Beurteilung bei positiveren Einstellungen oder Stereotypen sowie negativeren Beurteilung bei negativeren Einstellungen oder Stereotypen gegenüber bestimmten Schüler*innengruppen legen nahe, Urteilsverzerrungen entweder durch Reflektion und Kontrolle der Einstellungen und/oder durch eine Veränderung der relevanten Einstellungen und Stereotype zu vermindern.

8.3 Einstellungsänderung und bisherige Interventionsansätze

Die Grundlagenforschung zur Bildung und Änderung von Einstellungen ist hochkomplex und bietet eine Vielzahl von Modellen (vgl. Bohner & Wänke, 2006; Eagly & Chaiken, 1993). Dabei gelten Einstellungen und Stereotype gegenüber sozialen Gruppen und ihren Mitgliedern als relativ stabil. Wenn soziale Gerechtigkeit beziehungsweise gesellschaftliche Diskriminierung von Minderheiten im Fokus steht, wird oft von Vorurteilen gesprochen, die in einigen theoretischen Zugängen auch mit Einstellungen gleichgesetzt werden, wobei sich die Einstellungen beziehungsweise Vorurteile dann auf soziale Kategorien als Einstellungsobjekte beziehen (vgl. Eagly & Chaiken, 1993).[2] In der Forschung zur Veränderung interpersoneller Einstellungen von (zukünftigen) Lehrkräften gegenüber Schüler*innen anderer ethnisch-kultureller Herkunft, aus Familien mit niedrigem sozialem Status oder mit besonderem Förderbedarf können, grob gesagt, zwei Interventionsstrategien identifiziert werden, die unter theoretischen Gesichtspunkten als komplementär zu betrachten sind und in der Lehrkräftebildung auch kombiniert werden.

Die Intergruppenkontakttheorie (Überblick in Pettigrew & Tropp, 2006) besagt, dass der persönliche Kontakt zwischen Personen verschiedener sozialer

[2] In einigen Studien werden Vorurteile mit negativen Einstellungen gleichgesetzt. Das halten wir für eine sachlich irreführende Vereinfachung, da Vorurteile, wie Stereotype, auch positive Bewertungen beinhalten können. Der englische Begriff „prejudice“ wird in den relevanten Forschungsfeldern offenbar nur für negativ bewertende Vorurteile verwendet. Eine grundlagentheoretische Unterscheidung der insgesamt drei Konstrukte scheint für unsere Fragestellungen aber nicht wirklich relevant.

Gruppen (z. B. hinsichtlich ethnisch-kultureller Herkunft, sozio-ökonomischem Status) negative Vorurteile reduzieren kann. Allport leitete 1954 aus den bis dahin publizierten Ergebnissen zu Effekten von Intergruppenkontakten folgende Faktoren als notwendige Bedingungen für positive Effekte ab: In der Kontaktsituation haben die Personen beziehungsweise Gruppen gleichen Status, verfolgen gemeinsame Ziele, müssen kooperieren und der Kontakt ist institutionell „autorisiert", durch Gesetze gestützt oder durch Gewohnheit etabliert. Obwohl Lernende und Lehrende zweifelsfrei in einer asymmetrischen Statusbeziehung interagieren und ihnen oft auch verschiedene Ziele unterstellt werden, wurde in der Lehrer*innenbildung – meist ohne expliziten Verweis – mit diesem Modell gearbeitet. Insbesondere Maßnahmen des sogenannten *Service Learning,* bei dem Studierende im Rahmen von Lehrveranstaltungen (regelmäßige) praktische Service-Aufgaben, zum Beispiel in Gemeindeeinrichtungen, bekommen, können darauf zurückgeführt werden. So zeigte beispielsweise eine Trainingsstudie mit Lehrkräften zur Veränderung negativer Einstellungen und Stereotype gegenüber geistig Beeinträchtigten positive Effekte, wenn die Lehrkräfte Informationen über geistig Beeinträchtigte erhielten und mit diesen in ihrer Lehrkraftfunktion in Workshops arbeiteten, nicht aber wenn die Lehrkräfte nur Informationen erhielten (Hassanein, 2015). Die Evaluation eines einführenden Service-Learning-Kurses, in dessen Verlauf Lehramtsstudierende wöchentlich eine Stunde mit kleinen Gruppen *benachteiligter* Schüler*innen arbeiteten, ergab zwar positivere Einstellungen der Studierenden zur Inklusion beeinträchtigter Schüler*innen in den regulären Unterricht, aber keine Unterschiede in ihren Einstellungen zu kulturellem Pluralismus (Appreciation, Value, Implementation, Comfort with) im Vergleich zu Lehramtsstudierenden, die diesen Kurs noch nicht absolviert hatten (Lucas, 2011). Eine in der Lehrer*innenbildungsforschung vielzitierte qualitative Studie von Haberman und Post (1992) ergab andererseits, dass Lehramtsstudierende am Ende eines mehrwöchigen Service-Learning-Einsatzes in Schulklassen mit Schüler*innen aus Familien mit niedrigem sozialem Status und ethnisch-kultureller Minderheiten im Vergleich zur Vorbefragung zwar zuversichtlicher waren, mit diesen Schüler*innen effektiv arbeiten zu können (Instruktion, Umgang mit Störungen) und positivere Einstellungen äußerten (Akzeptanz), gleichwohl mehr negative Merkmalszuschreibungen machten. Als Fazit warnen Haberman und Post (1992) davor, Berichte über intensiv positiv erlebte Praxiserfahrungen mit der angestrebten Entwicklung einer kritischen Sicht auf Stereotype und Vorurteile zu verwechseln; vielmehr würden persönliche Stereotype ohne eine kritische Auseinandersetzung mit den Erfahrungen und ihrer Interpretation eher bestätigt. Die eher emotional gefassten positiveren Einstellungen nach der Einheit stehen im Einklang mit der Interpretation der Ergebnisse einer Meta-Analyse von Pettigrew und Tropp (2006,

2008). Sie schreiben den von Allport (1954) und anderen als notwendig erachteten Bedingungen zwar eine förderliche, nicht aber eine notwendige Funktion zu. Positive Effekte in zahlreichen Studien, in denen die Bedingungen nicht erfüllt oder gar konterkariert wurden, sprächen vielmehr dafür, dass steigende Vertrautheit vorhandene Unsicherheit und Bedrohungsgefühle reduziere und dadurch zu mehr gegenseitigem Gefallen führe. Die Meta-Analyse zeigte dabei, dass die insgesamt inverse Beziehung zwischen Kontakt und negativen Einstellungen, Stereotypen und Vorurteilen im Mittel stärker ausfiel, wenn die Kontakte durch gezielte (institutionelle) Maßnahmen zur Vorurteilsreduktion strukturiert wurden. Darüber hinaus belegt eine Meta-Analyse von Miles und Crisp (2014), dass auch vorgestellter Kontakt positive Effekte auf interpersonelle Einstellungen haben kann, und zwar umso stärker, je mehr die Personen dazu angehalten wurden, die Kontaktsituation gedanklich zu elaborieren. Dieser Effekt vertiefter Auseinandersetzung mit dem Einstellungsobjekt spricht für die Komplementarität der Intergruppenkontaktstrategie zur Einstellungsänderung mit dem zweiten Ansatz, der Einstellungsänderung durch Persuasion mittels Informationen über Einstellungsobjekte.

Dual-Process-Theorien der Einstellungsbildung und -änderung durch Persuasion (Überblick in Bohner & Wänke, 2006) postulieren genau diesen Effekt vertiefter Verarbeitung. Sowohl das *Elaboration Likelihood Model* (ELM; Petty & Cacioppo, 1986) als auch das *Heuristic* versus *Systematic Information Processing Model* (HSM; Chaiken, 1980) erklären Unterschiede in der Stärke und Persistenz von Einstellungsänderungen nach Präsentation persuasiver Informationen über ein Einstellungsobjekt mit zwei verschiedenen Routen beziehungsweise Prozessen der Informationsverarbeitung. Ähnlich wie im Kontinuum-Model der Eindrucksbildung (Fiske & Neuberg, 1990) wird auch in diesen Modellen zwischen einer automatisierten und einer kontrollierten Informationsverarbeitung unterschieden, die allerdings nicht als Kontinuum, sondern als zwei verschiedene Routen verstanden werden. Nach dem ELM (Petty & Cacioppo, 1986) stehen diese Informationsverarbeitungsrouten in einer inversen Beziehung zueinander und werden nur bei mittlerer Ausprägung parallel wirksam, wenn es um die Verarbeitung einstellungsrelevanter Informationen geht. Nach dem HSM (Chaiken, 1980) findet automatisierte Informationsverarbeitung praktisch immer (auch) statt, kontrollierte Informationsverarbeitung dagegen nur unter bestimmten Bedingungen. Beide Modelle postulieren, dass auf der peripheren Route weniger die Inhalte einer einstellungsrelevanten Botschaft verarbeitet werden und zum Beispiel die Qualität eines Arguments gerade nicht eingeschätzt wird, sondern es werden eher (oberflächliche) Reize, wie beispielsweise die Attraktivität oder die Expertise der Quelle (des Senders bzw. der

Senderin) einer einstellungsrelevanten Information oder auch eigene spontane affektive Reaktionen auf die Botschaft oder die Situation wahrgenommen und automatisiert beziehungsweise heuristisch verarbeitet. In Experimenten führt diese heuristische Verarbeitung peripherer Merkmale neuer Informationen über ein Einstellungsobjekt in der Regel nur zu einer kurzfristigen und instabilen Änderung einer Einstellung, die kaum verhaltensrelevant wird.[3] Die periphere Route mit heuristischer Informationsverarbeitung erfordert wenig kognitiven Aufwand und gilt als „Normalfall". Den Aufwand für vertiefte Verarbeitung (zentrale Route) investieren Personen nur dann, wenn sie genügend Motivation und Fähigkeiten haben, die Inhalte einer Botschaft systematisch zu verarbeiten (Chaiken, 1980; Petty & Cacioppo, 1986), das heißt „elaboration likelihood … is a function of two factors: *motivation* and *ability* to process a given message" (hervorgehoben im Original, Bohner & Wänke, 2006, S. 137 f.); weder Motivation noch Fähigkeit sind einzeln betrachtet hinreichend. Die *Motivation* zur Elaboration wird nach Petty und Cacioppo (1986) maßgeblich durch persönliche Relevanz, das Kognitionsbedürfnis (need for cognition; Cohen et al., 1955) sowie in Anlehnung an den Ringelmann-Effekt (Ingham et al., 1974) durch persönliche Verantwortung für die korrekte Bewertung eines Einstellungsobjektes, zum Beispiel eines*r Schülers*in, beeinflusst. Die *Fähigkeit* zur vertieften Verarbeitung der Inhalte und Qualität einer Information beschreiben Petty und Cacioppo (1986) als komplexes Zusammenspiel der folgenden drei Faktoren: 1) individuelle Voraussetzungen, wie beispielsweise das individuelle Wissen über das Einstellungsobjekt; 2) Aspekte der Botschaft, insbesondere deren Verständlichkeit; 3) Kontextbedingungen, insbesondere keine Ablenkungen in der Situation. Ob Personen nach der zentralen Verarbeitung der (neuen) Informationen über das Einstellungsobjekt, zum Beispiel über bestimmte Schüler*innen, ihre Einstellung ändern und in welcher Richtung sie diese ändern, wird dann auch davon beeinflusst, welche Richtung und Stärke ihre ursprüngliche Einstellung hatte. Die wechselseitige Beeinflussung von Bewertungstendenzen, Erfahrungen und Prozessen – gleichermaßen verhaltensbezogen, affektiv und kognitiv – beschreiben Petty und Cacioppo (1986) als Kreisprozesse, in deren Rahmen Einstellungsänderungen stattfinden (können).

[3] Gleichwohl kann oberflächliche Verarbeitung bei genügend hoher Frequenz über einen langen Zeitraum in stabilen Einstellungen resultieren (z. B. durch evaluative conditioning, vgl. Bohner & Wänke, 2006), was die Stabilität interpersoneller Einstellungen (einschließlich Stereotypen und Vorurteilen) erklärt.

Maßnahmen in der Lehrkräftebildung, die als eine Strategie auch die elaborierte Verarbeitung einstellungsrelevanter Informationen beinhalten, sind in der Regel curricular angelegt, zum Beispiel als integrierte Pflicht-Curricula oder als besondere Programme mit zum Teil freiwilliger Teilnahme der Studierenden. Im Feld können Motivation und Passung individueller Voraussetzungen mit den Lernbedingungen für einstellungsrelevante Informationen (Fähigkeiten) sowie ihre multiplen Wechselwirkungen kaum kontrolliert werden. Daher sind aus Evaluationsstudien in der Lehrkräftebildung eher kleine Effekte der vertieften Auseinandersetzung im Rahmen multikultureller Curricula zu erwarten, die sich erst nach längeren Entwicklungszeiträumen zeigen und auch gegenläufig (unerwünscht) ausfallen können. Grottkau und Nickolai-May (1989) verglichen mittels der Bogardus Social Distance Scale (Bogardus, 1933) die Entwicklung der Einstellungen Lehramtsstudierender gegenüber Personen anderer ethnisch-kultureller Herkunft im Rahmen eines breit angelegten integrierten Curriculums multikultureller Pädagogik, das im Studienverlauf auch praktische Unterrichtserfahrungen in multikulturellen Schulklassen vorsah, mit der Entwicklung der Einstellungen Studierender der Betriebswirtschaft und der Verwaltungswissenschaften über den Verlauf eines Semesters in je einem Kurs pro Vergleichsgruppe. Sie fanden keine Einstellungsveränderungen der Lehramtsstudierenden im ersten Semester im Vergleich zu Erstsemester-Studierenden anderer Studiengänge (Pre-post-Gruppenvergleich), aber geringere soziale Distanz (positivere ethnisch-kulturelle Einstellungen) der Lehramtsstudierenden im letzten Studiensemester im Vergleich zu den Lehramtsstudierenden im ersten Semester (Prätest-Kohortenvergleich). Den Kohorteneffekt führten sie auf fehlende oder unsystematische interethnische Kontakte der Erstsemester-Studierenden gegenüber nachhaltigen und kontrollierten Kontakten der Lehramtsstudierenden im Abschlusssemester zurück. Aus einer groß angelegten Studie mit vier Erhebungswellen berichteten Kumar und Hamer (2013) sowohl Kohorten- als auch Längsschnitteffekte vom ersten bis zum letzten Semester eines Lehrkräftebildungsprogramms, das Elemente multikultureller Erziehung beinhaltete. Im Kohortenvergleich (Beginn, Mitte, Ende) verringerte sich das Ausmaß negativer Stereotype gegenüber Schüler*innen aus ethnisch-kulturellen Minderheiten und gegenüber Schüler*innen aus Familien mit niedrigem sozialem Status sowie das Unbehagen, mit Schüler*innen und Kolleg*innen anderer Herkunft zu interagieren. Die Längsschnittanalysen der Angaben derjenigen Studierenden, die sowohl in der Mitte als auch zum Ende ihres Studiums befragt werden konnten, zeigten lediglich eine signifikante positive Veränderung ihrer Stereotype gegenüber Schüler*innen aus Familien mit niedrigem sozialem Status. Martin und Koppelman (1991) konnten dagegen aus der Evaluation eines ein-

semestrigen „Human Relations“ Kurses vom Prätest zum Posttest eine deutliche Sensibilisierung der teilnehmenden Lehramtsstudierenden für gesellschaftliche Diskriminierung (Sexism, Classism, Racism) berichten, die die Autor*innen als Einstellungsveränderungen gegenüber benachteiligten Gruppen interpretierten. Ihr Instrument, das die Ziele und Inhalte des Kurses spezifisch adressierte, hatten sie zuvor mit Lehramtsstudierenden anderer Universitäten validiert, deren Kurse Ethnizität und kulturellen Pluralismus statt sozialer Gerechtigkeit und Diskriminierung fokussierten. Diese zeigten keine Veränderung ihrer Sensibilität für gesellschaftliche Diskriminierung bestimmter Personengruppen über ein Semester hinweg und können als Kontrollgruppen gelten. Die Effekte führten Martin und Koppelman (1991) darauf zurück, dass die Studierenden sich im Kurs nicht nur kognitiv, sondern auch affektiv mit Ungerechtigkeit, Diskriminierung und ganz besonders mit den eigenen Einstellungen auseinandersetzten. Systematisch kombiniert wurden theoretische Lerneinheiten mit Informationen, Praxiselemente in Schulklassen, Reflexion der eigenen Einstellungen sowie Austausch mit Personen der diskriminierten Gruppen, die zum Beispiel für Vorträge und als Dozierende eingeladen worden waren (Expert*innen). Ergebnisse einer Feldstudie von Hammond (2010) auf der Basis von Interviews und Unterrichtsbeobachtungen im Verlauf eines Kurses zu multikultureller Bildung deuten darauf hin, dass die Richtung von Einstellungsänderungen nicht durch einzelne emotionale Erfahrungen oder kognitive Mechanismen, sondern durch deren Verbindung mit Informationssuche, Integration von Informationen, Reflexion eigener Emotionen und Fähigkeit zur Aufrechterhaltung von Empathie für andere kulturelle Orientierungen beeinflusst werde. Eine Meta-Analyse publizierter und unpublizierter Evaluationen in 40 Jahren Diversity Trainings (Bezrukova et al., 2016) ergab eine (nicht signifikant) geringfügig höhere Effektstärke von Trainings in organisationalen im Vergleich zu schulischen Kontexten, hohe Effekte für Teilnehmer*innenreaktionen, die geringsten Effekte für die Veränderung von Einstellungen (inkl. Stereotypen, Vorurteilen, Überzeugungen) und mittlere Effekte für Verhaltensänderungen und Wissenszuwachs.

Zwei experimentelle Studien belegten differenzielle positive Effekte kurzer Interventionen auf die Sensibilisierung beziehungsweise Stereotype verschiedener Studierender gegenüber Personen anderer ethnisch-kultureller Herkunft. Soble et al. (2011) ließen Studierende eine 20-minütige investigativ-journalistische Dokumentation des Lebens zweier Männer „(one Black the other one White)“ (S. 153) als Videoclip anschauen. Die Gegenüberstellung der gegensätzlichen Alltagserfahrungen machten gesellschaftliche Diskriminierung und Rassismus unmittelbar sichtbar. Die Kontrollgruppe sah ein Video ohne Bezüge

zu Rassismus (Karriereexploration College-Studierender). Die Interventionsgruppe zeigte vom Prä- zum Posttest im Vergleich zur Kontrollgruppe signifikante Steigerungen in Empathie, Schuldgefühlen gegenüber Minoritäten (White Guilt) und geringere Leugnung und Verharmlosung von Rassismus, aber keine Veränderung ihrer Furcht vor Begegnungen und ihrer Vorurteile gegenüber Personen anderer Herkunft. Die Ergebnisse interpretierten die Autor*innen als Effekte der kognitiven Auseinandersetzung mit Informationen ohne die affektiven Wirkungen interpersonellen Kontakts. Einen Effekt auf Stereotype, emotionale Vorurteile und Intentionen solidarischen Handelns berichteten Kotzur et al. (2019) von einer Intervention mit der „Fast-Friendship-Procedure", bei der sich Versuchsleiterverbündete mit den echten Versuchspersonen im Rahmen einer Studie gegenseitig nach einem festen Skript interviewten. In der Interventionsgruppe wurde der andere Studienteilnehmer (evtl. die andere Studienteilnehmerin)[4] als Flüchtling aus Syrien vorgestellt, in der Kontrollgruppe ein*e Student*in ohne besonderen Hintergrund. In der anschließenden Befragung, die als zweite Studie deklariert war, bewerteten die Studierenden der Experimentalgruppe den ‚anderen' Studienteilnehmer (die andere Studienteilnehmerin) (Outgroup) im Vergleich zur Kontrollgruppe (Ingroup-Teilnehmer*in) stärker als sympathisch, wohlgesonnen und warm(herzig) und äußerten gegenüber Asylbewerber*innen im Allgemeinen weniger Ablehnung.

Zusammenfassend zeigen die Ergebnisse der Studien Effekte einer ganzen Reihe von Einflussfaktoren, die von den theoretischen Modellen auch vorhergesagt werden: Vertiefte Auseinandersetzung mit einstellungsrelevanten Informationen (ELM, Petty & Cacioppo, 1986), (strukturierter) Intergruppenkontakt (Allport, 1954) und dadurch Steigerung der Vertrautheit (Pettigrew & Tropp, 2006). Kürzere Interventionen, einzelne oder indirekte Strategien zeigten überwiegend geringere beziehungsweise auf einzelne Facetten beschränkte Effekte (z. B. Haberman & Post, 1992; Lucas, 2011; Pettigrew & Tropp, 2008; Soble et al., 2011). Längere Interventionen, kombinierte Strategien und direkte (bewusste) Auseinandersetzung mit einstellungsrelevanten Informationen – oder Personen – zeigten stärkere beziehungsweise umfassendere Effekte (z. B. Grottkau & Nickolai-May, 1989; Hassanein, 2015; Kotzur et al., 2019; Kumar & Hamer, 2013; Pettigrew & Tropp, 2008).

[4] Im Original bleibt unklar, ob nur männliche Studierende untersucht wurden und/oder ob nur männliche Verbündete beteiligt waren.

Die Vielzahl der Einflussfaktoren und ihre Wechselwirkungen, die in Laborexperimenten ermittelt und in den Dual-Process-Theorien zusammengefasst wurden, können in der Lehrkräftebildung zwar nicht kontrolliert werden, wurden aber möglicherweise in einigen curricular angelegten Maßnahmen mit multiplen Elementen näherungsweise – mehr oder weniger systematisch – adressiert. Mit Petty und Cacioppo (1986), die in Anlehnung an Festinger (1954) davon ausgehen, dass Personen grundsätzlich motiviert sind, korrekte Einstellungen zu haben und diese gegebenenfalls anzupassen, kann für (angehende) Lehrkräfte eine überwiegend hohe Bereitschaft angenommen werden, sich mit Informationen über die Heterogenität von Schüler*innen und insbesondere mit Informationen über Bildungsungerechtigkeit für Schüler*innen aus Familien mit Migrationshintergrund oder mit niedrigem sozialem Status auseinanderzusetzen und die eigenen Einflussmöglichkeiten zu elaborieren. Wenn sie den Anspruch haben, neutral, gerecht und objektiv zu urteilen, bemühen sie sich vermutlich „zu positive“ Einstellungen „nach unten“ zu korrigieren und „zu negative“ Einstellungen „noch oben“ zu korrigieren. Erfahren sie zum Beispiel, dass sie bestimmten Schüler*innengruppen bestimmte Eigenschaften zuschreiben, die nicht mit empirischen Fakten übereinstimmen, oder dass sie bestimmte Schüler*innengruppen unbewusst präferieren, sollte sich die Wahrscheinlichkeit erhöhen, dass sie diese neuen Informationen vertieft verarbeiten und daraufhin ihre Einstellungen gegenüber diesen Schüler*innengruppen anpassen.

8.4 Eine experimentelle Interventionsstudie

Ausgehend von der Annahme, dass herkunftsassoziierte Einstellungen und Stereotype von Lehrkräften zu Verzerrungen in der Beurteilung bestimmter Schüler*innengruppen führen könnten und damit zu einer Benachteiligung beziehungsweise Bevorzugung von Schüler*innen spezifischer Herkunft beitragen, war das Ziel der hier vorgestellten Studie, Einstellungen und Stereotype Lehramtsstudierender gegenüber stigmatisierten Schüler*innengruppen durch eine Kurzintervention zu verändern. Unter der Annahme hoher Stabilität interpersoneller Einstellungen sehen bisherige Ansätze zur Änderung herkunftsbezogener Einstellungen und Stereotype in Bildungskontexten zeitbeziehungsweise lernintensive Interventionen in Form entsprechender Curricula vor. Experimentalstudien zur Prüfung theoretischer Modelle der Einstellungsänderung zeigten aber auch Effekte kurzzeitiger Interventionen auf bereits gebildete Einstellungen. Unter theoretischen wie bildungspraktischen Gesichtspunkten stellt sich angesichts dieser Forschungslage die Frage, ob die Aus-

prägungen herkunftsassoziierter expliziter Einstellungen und/oder Stereotype Lehramtsstudierender durch eine kurzzeitige Intervention verändert werden können.

8.4.1 Fragestellung

Die berichtete Forschungslage spricht für eine integrierte Sichtweise auf die theoretischen Modelle der Einstellungsänderung und ihre Implementation in Interventionsstudien. Dabei scheinen die Auseinandersetzungen mit bereits gebildeten persönlichen Einstellungen sowie mit eigenen Emotionen im Kontakt mit Personen, die in den theoretischen Modellen eher indirekt abgebildet werden, eigenständige Einflussgrößen darzustellen (vgl. Pettigrew & Tropp, 2006). Für den Ansatz, in der Lehrkräftebildung die herkunftsbezogenen Einstellungen und Stereotype von Studierenden zu verbessern, können folgende Fragen gestellt werden:

1. Zeigt eine systematische Kombination der Einflussfaktoren auch in einer Kurzintervention im Rahmen einer experimentellen Studie mit randomisierter Zuordnung zu einer Experimental- (= Interventions-) und einer Kontrollgruppe (= Placebogruppe) Effekte?
2. Zeigen sich gegebenenfalls Effekte nur für explizite Einstellungen, nur für explizite Stereotype oder für beide Aspekte der Bewertung bestimmter Schüler*innengruppen?
3. Falls sich Effekte zeigen: Bleiben diese Effekte bei einer Kontrolle der impliziten Einstellungen in Richtung und Ausmaß ähnlich oder verändern sie sich maßgeblich, sodass von einer Abhängigkeit der Wirkung der Intervention von der vorgängigen Ausprägung der impliziten Einstellungen ausgegangen werden muss, zum Beispiel als Moderator oder Mediator?
4. Falls sich Effekte zeigen: Wie nachhaltig sind diese Effekte über die unmittelbare Intervention hinaus?

Im Fokus der Studie standen als abhängige Variablen die expliziten Einstellungen und Stereotype zu Schüler*innen ohne und mit türkischem Migrationshintergrund sowie zu Schüler*innen aus Familien mit niedrigem und hohem sozialem Status, da sich für diese Herkunftsmerkale starke Disparitäten im Bildungserfolg zeigen und – korrespondierend mit den Disparitäten – auch herkunftsabhängige Urteilsverzerrungen durch Lehrkräfte bekannt sind (z. B. OECD, 2019; Tobisch & Dresel, 2017).

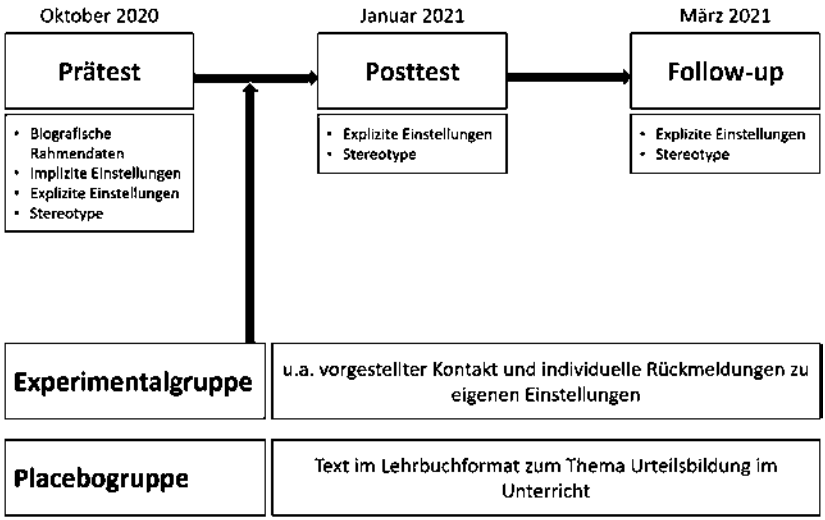

Abb. 8.1 Überblick über die Untersuchungsanlage

8.4.2 Design

Im Rahmen einer umfassenderen experimentellen Studie im Prä-Post-Follow-up-Design wurden Lehramtsstudierende randomisiert einer Interventions- und einer Placebogruppe zugewiesen (vgl. Abb. 8.1). Zu allen drei Messzeitpunkten wurden die expliziten Einstellungen und Stereotype gegenüber Schüler*innen mit türkischem und ohne Migrationshintergrund sowie gegenüber Schüler*innen aus Familien mit niedrigem und mit hohem sozialem Status in einem Online-Fragebogen erfasst. Zudem wurden auch implizite Einstellungen gegenüber diesen beiden sozialen Kategorien erfasst. Nachdem davon auszugehen ist, dass implizite Einstellungen relativ stabil sind (z. B. Baron, 2015) und entsprechend nicht von einer bedeutsamen Änderung auszugehen ist, wurden die impliziten Einstellungen ausschließlich zum ersten Messzeitpunkt erfasst, um diese als Kontrollvariable berücksichtigen zu können. Zudem wurden zum ersten Messzeitpunkt unter anderem biografische Rahmendaten der Studierenden erhoben.

8.4.3 Stichprobe

Insgesamt nahmen 221 Studierende des Grundschullehramts an der Studie teil.[5] Die Teilnahme wurde ihnen als Versuchspersonenstunde im Rahmen des Einführungsmoduls Psychologie anerkannt. Mit zwei Fällen nach dem ersten und vier weiteren Fällen nach dem zweiten Messzeitpunkt betrug der Dropout weniger als 3,0 %, sodass für die Analysen Daten von $N = 215$ Studierenden vorlagen, die überwiegend im 1. Semester studierten (98,6 %). Passend zum Studienzeitpunkt (überwiegend Erstsemester) waren die Studierenden im Durchschnitt erst $M = 19{,}9$ Jahre ($SD = 2{,}8$) alt und erwartungsgemäß lag mit 90,2 % für diese Stichprobe ein hoher Frauenanteil vor. Bei der Abfrage der Geburtsländer (eigenes, Vater, Mutter) gaben 20,8 % der Studierenden mindestens ein anderes Geburtsland als Deutschland in ihrer Kernfamilie an. Da der Fokus in dieser Studie auf den Einstellungen und Stereotypen gegenüber Schüler*innen mit türkischen Wurzeln lag, wurde ein besonderes Augenmerk auf die Verteilung dieser kulturellen Ingroup über Experimental- und Placebogruppe gelegt, deren Anteil mit 3,2 % an der Gesamtstichprobe zwar sehr gering ausfiel, allerdings ungleich verteilt auf die Gruppen, nämlich nur in der Interventionsgruppe (vgl. Tab. 8.1; $\text{Chi}^2(2) = 9{,}327$, $p < ,01$). Der überwiegende Anteil der

Tab. 8.1 Stichprobenbeschreibung für die Interventions- und Placebogruppe

	Interventionsgruppe	Placebogruppe
N	112	103
Alter	$M = 20{,}1$ ($SD = 2{,}7$)	$M = 19{,}9$ ($SD = 3{,}8$)
Fachsemester	$M = 1{,}0$ ($SD = 0{,}2$)	$M = 1{,}0$ ($SD = 0{,}2$)
Geschlecht (weiblich)	88,6 %	91,6 %
Sozialer Status	$M = 3{,}4$ ($SD = 0{,}7$)	$M = 3{,}4$ ($SD = 0{,}6$)
Migrationshintergrund	27,2 %	14,0 %
Türkischer Hintergrund	6,1 %	0,0 %
Drop-out	1,8 %	2,7 %

Anmerkung. $N = 215$

[5] Eine a-priori Poweranalyse in G*Power ergab eine notwendige Stichprobengröße von $N = 260$, die aus organisatorischen Gründen nicht erreicht werden konnte.

Studierenden ordnete sich selbst beziehungsweise die eigene Familie der Mittelschicht (53,4 %) oder der oberen Mittelschicht (40,7 %) zu. Lediglich 1,8 % der Studierenden gaben an, der Oberschicht anzugehören, 3,2 % fühlten sich der unteren Mittelschicht und 0,9 % der Unterschicht zugehörig. Die biografischen Rahmendaten der Interventions- und der Placebogruppe können Tab. 8.1 entnommen werden.

8.4.4 Interventionsbedingungen

Zu Beginn des zweiten Messzeitpunkts erfolgte die Intervention. In der *Interventionsgruppe* wurden die als wirksam geltenden Einflussfaktoren mit folgenden Elementen und in dieser Reihenfolge realisiert:

Informationen: Unter der Überschrift ‚Rückmeldung zu Ihren individuellen Einstellungen' erhielten die Studierenden als erstes eine kurze Erläuterung der Konstrukte ‚Einstellungen', ‚explizite Einstellungen' und ‚implizite Einstellungen'.

Auseinandersetzung mit eigenen Einstellungen: Es folgte die Rückmeldung über die individuellen Ausprägungen ihrer expliziten und impliziten Einstellungen gegenüber Schüler*innen ohne und mit türkischem Migrationshintergrund zum ersten Messzeitpunkt in Form von Balkendiagrammen (z-Werte). In den Anmerkungen wurde erläutert, was positive und negative Werte inhaltlich bedeuten. Unterhalb der Abbildungen wurde das individuelle Befundmuster interpretiert und mögliche Effekte dieser Einstellungsausprägungen auf die Beurteilung entsprechender Schüler*innen aufgezeigt. Hervorgehoben wurde dabei, ob sie mit ihren Einstellungen Gefahr laufen, bestimmte Schüler*innen ungewollt zu bevorzugen und/oder bestimmte Schüler*innen ungewollt zu benachteiligen.

Informationen: Unter der Überschrift ‚Der Einfluss von Einstellungen und Stereotypen auf Beurteilungen' wurde den Studierenden ein Text im Lehrbuchformat präsentiert (ca. eine Seite). Dargestellt wurden die Bedeutung von Beurteilungen für Bildungsverläufe, theoretische Grundlagen über automatisierte versus kontrollierte Urteilsbildung, empirische Ergebnisse zu herkunftsbezogenen Urteilsverzerrungen und dem Einfluss von Einstellungen und Stereotypen.

Selbstreflexion: Unter der Überschrift ‚Handlungsstrategien' erhielten die Studierenden auf der letzten Seite unter anderem Anregungen zur Selbstbeobachtung ihrer Reaktionen auf erwartungswidrige Schüler*innenleistungen, ob ihre Erwartungen mit der Schüler*innenherkunft in Zusammenhang stehen

und welche Stereotype sie in Bezug auf diese Gruppe (einschließlich der Ingroup) haben.

Vorgestellter Kontakt: Weiterhin wurden die Studierenden aufgefordert, sich verschiedene konkrete Personen der betreffenden Gruppe vorzustellen (s. o. die Ergebnisse der Meta-Analyse von Miles & Crisp, 2014), insbesondere solche, die ihren Erwartungen an diese Gruppe widersprechen.

Anregung und Auseinandersetzung mit eigenen Emotionen: Direkt im Anschluss wurden die Studierenden angeregt, sich ihre emotionalen Reaktionen auf erwartungswidriges Verhalten, wie zum Beispiel Leistungen, bewusst zu machen und als Beispiele drei potenzielle Emotionen genannt (Freude, Überraschung, Enttäuschung), die in Untersuchungen zu Effekten emotionaler Lehrkraftreaktionen auf Erfolg und Misserfolg von Schüler*innen eingesetzt wurden (vgl. Frenzel & Taxer, 2018) und von uns als genügend realitätsnah für schulischen Alltag eingeschätzt wurden.

Anregung kognitiver Verhaltenskontrolle: In Anlehnung an Strategien zur Unterbrechung automatisierter Reaktionen in Situationen unter Handlungsdruck (vgl. Wahl, 1991), mit denen elaborierte Informationsverarbeitung und (bewusste) Handlungsentscheidungen angeregt werden können, wurden die Studierenden dazu ermutigt, sich einen persönlichen Merksatz zu notieren, der sie in Situationen unter Handlungsdruck daran erinnert, individuelle Informationen zu beachten, anstatt stereotypengeleitet zu urteilen. Die individuellen Merksätze wurden als offene Antworten am Ende der Intervention erfasst.

Die *Placebogruppe* erhielt keine Rückmeldung ihrer Ergebnisse, sondern ausschließlich einen Text (ebenfalls im Lehrbuchformt), der sich auch mit dem Thema Schüler*innenbeurteilung befasste, allerdings auf Urteilsbildung im Unterricht fokussierte. Hier wurden die nützlichen Funktionen automatisierter Urteilsbildung unter Handlungsdruck hervorgehoben, wenn Lehrkräfte in der dynamischen Unterrichtsinteraktion schnell und adaptiv reagieren wollen und müssen. Damit sollte die Aufmerksamkeit dieser Studierenden vom Thema Einstellungen und Stereotype weggelenkt und außerdem der Modus einer eher automatisierten Informationsverarbeitung angeregt werden.

Im Anschluss an die Intervention wurden die Studierenden der Interventionsgruppe mit fünf Items (von 1 *stimmt gar nicht* bis 6 *stimmt voll und ganz*) nach ihrer Einschätzung der dargebotenen Informationen gefragt. Dabei wurden unter anderem motivationale Aspekte in Anlehnung an die *Erwartungs-Wert Theorie* (Wigfield & Eccles, 2000) erfasst, die den subjektiven Wert („Mir ist es wichtig, meine Einstellungen und Stereotype zu verändern“) sowie die Erfolgserwartung

(„Ich glaube, ich kann meine Einstellungen und Stereotype verändern“) zur Umsetzung der Einstellungsänderung erfassen sollten.

8.4.5 Messinstrumente

Zu allen drei Messzeitpunkten wurden explizite Einstellungen und Stereotype gegenüber Schüler*innen mit türkischem und ohne Migrationshintergrund sowie gegenüber Schüler*innen aus Familien mit niedrigem und mit hohem Status erfasst. Zudem wurden die impliziten Einstellungen zum ersten Messzeitpunkt erhoben.

Die expliziten Einstellungen wurden zu allen drei Messzeitpunkten parallelisiert für jede der vier Herkünfte mit einer sechsstufigen bipolaren Skala (adaptiert nach Lehmann-Grube et al., im Druck) mit je 10 Items erfasst (Itemstamm: „Für meinen Unterricht halte ich die Herkunft von Schüler*innen aus Familien ohne Migrationshintergrund/mit einem türkischen Migrationshintergrund/mit einem hohen sozioökonomischen Status/mit einem niedrigen sozioökonomischen Status für…“). Die Skala beinhaltet drei Subskalen (Einstellungsfacetten) der Einstellungen gegenüber den Schüler*innen der vier Gruppen, die Valenz, zum Beispiel *gut* versus *schlecht,* den Nutzen für Schüler*innen, beispielsweise *hinderlich* versus *förderlich* für das Lernen der Schüler*innen, und die Kosten der Lehrkraft, zum Beispiel *kaum zu bewältigen* versus *gut zu bewältigen.* Die Gesamtskalen wiesen über alle vier Herkunftsgruppen und über alle drei Messzeitpunkte gute interne Konsistenzen auf ($\alpha = ,80 - ,92$).

Die Stereotype wurden als Vergleichsmaß der Schüler*innen mit türkischem versus Schüler*innen ohne Migrationshintergrund beziehungsweise Schüler*innen aus Familien mit niedrigem versus hohem Status erfasst (Itemstamm: „Im Vergleich zu Schüler*innen ohne Migrationshintergrund/aus Familien mit hohem sozialem Status sind Schüler*innen mit türkischem Hintergrund/aus Familien mit niedrigem sozialem Status im Unterricht durchschnittlich…“). Hierzu wurde je ein sechsstufiges semantisches Differenzial mit 15 Eigenschaftspaaren typischer Schüler*innenmerkmale eingesetzt, die in einer Befragung Lehramtsstudierender zur Vorbereitung einer Studienabschlussarbeit als relevante Schüler*innenmerkmale angegeben worden waren. Die ausgewählten Eigenschaften bezogen sich vor allem auf lern- und leistungsrelevante Merkmale wie kognitive Fähigkeiten (z. B. *gar nicht intelligent* versus *sehr intelligent)*, motivationale Merkmale (z. B. *gar nicht motiviert* versus *sehr motiviert*), soziale Aspekte (z. B. *gar nicht teamfähig* versus *sehr teamfähig*) und

allgemeinere Merkmale des Arbeitsverhaltens (z. B. *gar nicht unzuverlässig* versus *sehr unzuverlässig*). Beide Skalen wiesen über alle drei Messzeitpunkte gute bis sehr gute interne Konsistenzen auf ($\alpha = ,89–,94$).

Die impliziten Einstellungen wurden zum ersten Messzeitpunkt mit zwei impliziten Assoziationstests (IAT; z. B. Greenwald et al., 1998, 2003) mit der Software Inquisit 5 Web (Millisecond Software, 2016) erfasst. Dabei wurden einerseits implizite Einstellungen gegenüber der ethnischen und andererseits zur sozialen Herkunft erfasst. Angenommen wird bei impliziten Assoziationstests, dass Attribute, die im Gedächtnis enger mit bestimmten Kategorien verknüpft sind als mit anderen Kategorien, auch schneller in einer reaktionszeitgestützten Erfassung einander zugeordnet werden können (vgl. Greenwald et al., 2003). Um implizite Einstellungen mit diesem Messverfahren abbilden zu können, wurden als Attributkategorien entsprechend positive (z. B. „ehrgeizig“) und negative (z. B. „vorlaut“) Eigenschaften genutzt. Als Zielkategorien zur Erfassung impliziter Einstellungen in Bezug auf den Migrationshintergrund wurden die Begriffe „deutsch“ und „türkisch“ genutzt. Dabei sollten jeweils acht typisch deutsche und typisch türkische Vornamen diesen beiden Kategorien zugeordnet werden (vgl. Tobisch, 2019). Zur Erhebung der impliziten Einstellungen gegenüber der sozialen Herkunft sollten je acht Berufe (ausgewählt nach der *International Standard Classification of Occupations;* z. B. Bundesamt für Statistik, 2017; Ganzeboom et al., 1992) den Kategorien „hoher Status“ beziehungsweise „niedriger Status“ zugeordnet werden (vgl. Tobisch & Dresel, 2020). Die Versuchspersonen erhielten vor beiden impliziten Assoziationstests eine Übersicht mit den jeweiligen Ziel- und Attributkategorien und den zugehörigen Begriffen. In beiden Testvarianten wurde die Differenz der Latenzzeit der kongruenten und inkongruenten Testblöcke berechnet und daraus der *D*-Score als Indikator für die Stärke der impliziten Assoziationen abgeleitet (vgl. Greenwald et al., 2003). Ein positiver Wert lässt sich in der vorliegenden Studie als Präferenz gegenüber Personen ohne Migrationshintergrund beziehungsweise Personen mit hohem Status im Vergleich zu Personen mit türkischem Migrationshintergrund beziehungsweise Personen mit niedrigem sozialem Status interpretieren.

8.4.6 Analysemethoden

Zur Hypothesenprüfung wurden Varianzanalysen mit Messwiederholung berechnet. Für die Analysen der herkunftsbezogenen Einstellungsentwicklung wurde ein 3(Messzeitpunkte)*2(Interventions- versus Placebogruppe)*4(Herkunft)-Modell berechnet. Um Veränderungen der Stereotype zu untersuchen

wurde ein 3(Messzeitpunkte)*2(Interventions- versus Placebogruppe)*2(Stereotype nach Herkunftskategorien)-Modell berechnet.

Beide Modelle wurden zur Kontrolle zusätzlich mit den impliziten Einstellungen als Kovariate geprüft. Beide Modelle wurden angesichts der ungleichen Stichprobenverteilung hinsichtlich des eigenen Migrationshintergrundes der Studierenden zusätzlich unter Ausschluss dieser Studierenden aus den Analysen geprüft, und für beide Modelle wurden univariate Varianzanalysen mit Messwiederholung je Gruppe gerechnet, um die Interaktionseffekte näher aufzuklären, das heißt, das Ausmaß der Veränderungen je Gruppe zu prüfen. Unterschiede der impliziten Einstellungen in Abhängigkeit vom eigenen Migrationshintergrund wurden mit einer einfachen ANOVA kontrolliert.

8.4.7 Ergebnisse

8.4.7.1 Deskriptive Befunde

Die deskriptiven Befunde zum Prätest verweisen auf insgesamt eher positive Einstellungen gegenüber allen Herkunftsgruppen (vgl. Tab. 8.2). Allerdings zeigten die Befunde auch deutliche Unterschiede der Einstellungen zu den unterschiedlichen Schüler*innengruppen, $F(3, 212) = 61{,}297$, $p < {,}001$, $\eta^2 = {,}47$. Besonders positiv wurden Schüler*innen ohne Migrationshintergrund bewertet, am wenigsten positiv, aber immer noch über dem Skalenmittel, wurden Schüler*innen aus Familien mit niedrigem sozialem Status bewertet. Auch für die Stereotype ergaben sich im Schnitt positive Assoziationen, das heißt im Vergleich mit den Schüler*innen des jeweils gegenübergestellten Herkunftsmerkmals wurden Schüler*innen mit türkischem Migrationshintergrund und Schüler*innen aus Familien mit niedrigem sozialem Status positiv eingeschätzt. Hier zeigten sich keine Unterschiede je nach Herkunftsvergleich (mit versus ohne Migrationshintergrund und niedriger versus hoher Status, $F(1, 206) = 1{,}358$, $p = {,}25$, $\eta^2 = {,}00$). Die impliziten Einstellungen der Studierenden zeigten deutliche Präferenzen für Personen ohne Migrationshintergrund im Vergleich zu Personen mit türkischem Migrationshintergrund sowie für Personen mit hohem sozialem Status im Vergleich zu Personen mit niedrigem sozialem Status. Dabei war die Präferenz für Personen mit hohem sozialem Status deutlich stärker ausgeprägt als die Präferenz für Personen ohne Migrationshintergrund, $F(1, 214) = 55{,}497$, $p < {,}001$, $\eta^2 = {,}21$. Die wenigen Studierenden mit eigenem türkischem Hintergrund zeigten nur in Bezug auf Personen mit türkischem Hintergrund (ihre Ingroup) andere Präferenzen als Studierende ohne oder mit anderem Migrations-

Tab. 8.2 Mittelwerte und Korrelationen der inhaltlichen Variablen im Prätest

	M (*SD*)	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Einstellungen								
(1) Mit türkischem Migrationshintergrund	4,39 (0,75)	–						
(2) Ohne Migrationshintergrund	4,83 (0,60)	,23**	–					
(3) Mit niedrigem sozialem Status	4,04 (0,85)	,54**	,28**	–				
(4) Mit hohem sozialem Status	4,44 (0,69)	,28**	,49**	,35**	–			
Stereotype								
(5) Türkischer Migrationshintergrund	3,98 (0,61)	,35**	,02	,17*	−,07	–		
(6) Niedriger sozialer Status	4,03 (0,75)	,22*	−,08	,33**	−,11	,61**	–	
Implizite Einstellungen								
(7) Kein Migrationshintergrund	0,48 (0,34)	−,24**	,08	−,14*	−,10	−,07	−,07	–
(8) Hoher sozialer Status	0,70 (0,38)	−,11	,04	−,05	−,03	−,04	−,03	,31**

$N = 215$. *** $p < ,001$. ** $p < ,01$. * $p < ,05$

hintergrund, $F(2, 212) = 7{,}846$, $p = ,001$, $\eta^2 = ,07$, nicht aber bezüglich Personen mit hohem sozialem Status, $F(2, 212) = 1{,}589$, $p = ,21$, $\eta^2 = ,02$.

Die Korrelationen zeigten mittlere positive Zusammenhänge zwischen den Einstellungen und Stereotypen sowohl gegenüber Schüler*innen mit türkischem Migrationshintergrund als auch gegenüber Schüler*innen mit niedrigem sozialem Status (vgl. Tab. 8.2). Studierende, die positivere Bewertungen dieser Schüler*innen für ihren zukünftigen Unterricht äußerten, assoziierten auch positivere Eigenschaften mit ihnen. Die expliziten Einstellungen gegenüber Schüler*innen mit türkischem Migrationshintergrund korrelierten auch mit den impliziten Einstellungen gegenüber Personen dieser Gruppe. Studierende,

die keine (besonders starke) implizite Präferenz für Personen ohne Migrationshintergrund zeigten, gaben auch positivere explizite Einstellungen gegenüber Schüler*innen mit türkischem Migrationshintergrund an. Allerdings ergaben sich keine entsprechenden Zusammenhänge für die expliziten und impliziten Einstellungen in Bezug auf die soziale Herkunft. Gleichwohl zeigten die impliziten Einstellungen gegenüber einem türkischen Migrationshintergrund und niedriger sozialer Herkunft einen mittleren positiven Zusammenhang. Studierende, die eine starke Präferenz für Personen ohne Migrationshintergrund aufwiesen, zeigten auch eine starke Präferenz für Personen mit hohem sozialem Status.

Die Vergleiche der beiden Versuchsgruppen ergaben im Prätest keine bedeutsamen Unterschiede für die abhängigen Variablen, Einstellungen: $F(4, 216) = 0{,}954$, $p = ,43$; Stereotype: $F(2, 210) = 1{,}025$, $p = ,36$. Die Gruppen unterschieden sich lediglich in den impliziten Einstellungen, ($F(2, 212) = 3{,}391$, $p < ,05$, $\eta^2 = ,03$. Die Interventionsgruppe zeigte im Mittel eine weniger ausgeprägt positive Präferenz für Personen ohne Migrationshintergrund, $F(1, 213) = 6{,}789$, $p < ,01$, $\eta^2 = ,03$, aber eine gleich hoch ausgeprägte Präferenz für Personen mit hohem sozialem Status, $F(1, 213) = 0{,}889$, $p = ,35$, $\eta^2 = ,00$, wie die Placebogruppe.

Die Intervention wurde von den Teilnehmenden insgesamt positiv bewertet (vgl. Tab. 8.3). So gaben die meisten Studierenden an, die Rückmeldung zu den eigenen Einstellungen als interessant und überraschend empfunden zu haben. Zudem zeigte sich, dass die Teilnehmer*innen der Interventionsgruppe den berichteten Ergebnissen zu ihren Einstellungen eine persönliche Bedeutsamkeit zuschrieben und es für wichtig und möglich hielten, ihre Einstellungen zu verändern. Der Großteil der Studierenden in der Interventionsgruppe (97,7 %) gab zudem auch einen inhaltlich passenden individuellen Merksatz an. Inhaltlich fokussierten diese überwiegend eine gerechte und herkunftsunabhängige

Tab. 8.3 Einschätzung der Intervention durch Studierende der Interventionsgruppe

Item	*M* (*SD*)
Die Informationen zu meinen Einstellungen und Stereotypen waren interessant	5,2 (0,8)
Die Ergebnisse zu meinen Einstellungen und Stereotypen haben mich überrascht	4,3 (1,2)
Die Ergebnisse haben für mich eine persönliche Bedeutung	4,7 (1,1)
Ich glaube, ich kann meine Einstellungen und Stereotype verändern	5,0 (0,9)
Mir ist es wichtig, meine Einstellungen und Stereotype zu verändern	5,2 (1,1)

Anmerkung. $N = 114$

Beurteilung der Schüler*innen und verwiesen oftmals auch auf eine bewusste Reflexion eigener Einstellungen und Stereotype.

8.4.7.2 Ergebnisse zu den Fragestellungen: Die Effekte der Intervention

Die Varianzanalysen mit Messwiederholung erbrachten für die *herkunftsbezogenen Einstellungen* einen bedeutsamen Haupteffekt über die drei Messzeitpunkte und einen Haupteffekt für die Schüler*innenherkunft (vgl. Tab. 8.4), das heißt, die Einstellungen veränderten sich über die Messzeitpunkte hinweg und waren gleichzeitig in Bezug auf die Herkunftsgruppen über alle Messzeitpunkte hinweg deutlich verschieden. Zudem ergaben sich bedeutsame Interaktionseffekte mit der Experimentalbedingung sowohl für die Messzeitpunkte als auch für

Tab. 8.4 Varianzanalysen mit Messwiederholung zu den Effekten der Messzeitpunkte, der Schüler*innenherkunft sowie der Interventionsbedingungen auf die Einstellungen

	F	η^2
Effekte ohne Kontrolle der impliziten Einstellungen[a]		
Haupteffekte		
Messzeitpunkte	10,516***	,05
Herkunft	84,136***	,28
Interaktionseffekte		
Messzeitpunkte*Bedingungen	6,919**	,03
Herkunft*Bedingungen	4,980**	,02
Messzeitpunkte*Herkunft*Bedingungen	4,785***	,02
Effekte mit Kontrolle der impliziten Einstellungen[b]		
Haupteffekte		
Messzeitpunkte	0,649	,00
Herkunft	9,434***	,04
Interaktionseffekte		
Messzeitpunkte*Bedingungen	6,378**	,03
Herkunft*Bedingungen	3,591*	,02
Messzeitpunkte*Herkunft*Bedingungen	5,538***	,03

Anmerkungen. $N = 215$. [a] Messzeitpunkte: $df_{Hypothese} = 2$ *und* $df_{Fehler} = 426$; Herkunft: $df_{Hypothese} = 3$ *und* $df_{Fehler} = 639$. [b]Messzeitpunkte: $df_{Hypothese} = 2$ *und* $df_{Fehler} = 422$; Herkunft: $df_{Hypothese} = 3$ *und* $df_{Fehler} = 633$
*** $p < ,001$. ** $p < ,01$. * $p < ,05$

die Herkunft und eine bedeutsame Dreifach-Interaktion. Das bedeutet, die Einstellungen der Studierenden in der Interventionsgruppe zeigten zum zweiten und dritten Messzeitpunkt andere Ausprägungen verglichen mit der Placebogruppe und die Ausprägungen der Einstellungen veränderten sich für die Herkunftsgruppen zum Teil verschieden voneinander. Auch unter Kontrolle der impliziten Einstellungen blieb dieses Befundmuster im Wesentlichen bestehen; die Effekte waren lediglich teilweise kleiner. Nur der Haupteffekt der Messzeitpunkte entfiel, wenn für die impliziten Einstellungen kontrolliert wurde, das heißt, die impliziten Präferenzen hatten Einfluss auf das Ausmaß der Einstellungsveränderungen, nicht aber auf die Richtung der Veränderungen. Sowohl die Kontrollanalysen unter Ausschluss aller Studierenden mit Migrationshintergrund als auch unter Ausschluss nur der Studierenden mit türkischem Migrationshintergrund ergaben die gleichen Muster.

Die differenzierten Analysen der Interaktionseffekte (u. a. mittels univariater Messwiederholung je Gruppe) ergaben, dass die Unterschiede der Einstellungsentwicklungen in den Gruppen vor allem in Bezug auf Schüler*innen mit Migrationshintergrund und in Bezug auf Schüler*innen mit niedrigem sozialem Status zu finden waren (vgl. Abb. 8.2). Die Interventionsgruppe zeigte gegenüber Schüler*innen mit türkischem Migrationshintergrund konstant positive Einstellungen und gegenüber Schüler*innen mit niedrigem sozialem Status eine signifikant positive Entwicklung (Steigung) direkt nach der Intervention (Differenz ,232, $p < ,05$ bonferroni-korrigiert, 95 % KI für die Differenz [0,06, 0,40]), die sich bis zum Follow-up aber relativierte, $F(1{,}788, 222) = 5{,}593$, $p < ,01$, $\eta^2 = ,048$; Greenhouse-Geisser korrigiert. Gegenüber Schüler*innen ohne Migrationshintergrund veränderten sich die Einstellungen der Interventionsgruppe in zunehmend negativer Richtung, $F(2, 222) = 4{,}849$, $p < ,01$, $\eta^2 = ,042$; Differenz Prätest zu Posttest: −0,185, $p < ,05$ bonferroni-korrigiert, 95 % KI für die Differenz [−,03, −0,35], gegenüber Schüler*innen mit hohem sozialem Status blieben sie konstant am positivsten. In der Placebogruppe veränderten sich die Einstellungen zu allen vier Gruppen vom Prätest bis zum Follow-up kontinuierlich in negativer Richtung, mit deutlich negativer Veränderung vom Posttest zum Follow up (alle Differenzen $p < ,05$ bonferroni-korrigiert).

Im Gesamtbild wiesen die Einstellungen der Studierenden gegenüber Schüler*innen mit türkischem Migrationshintergrund sowie gegenüber Schüler*innen mit niedrigem sozialem Status nach der Intervention bedeutsame Unterschiede zwischen der Interventions- und der Placebogruppe auf, die auch im Follow-up bestehen blieben.

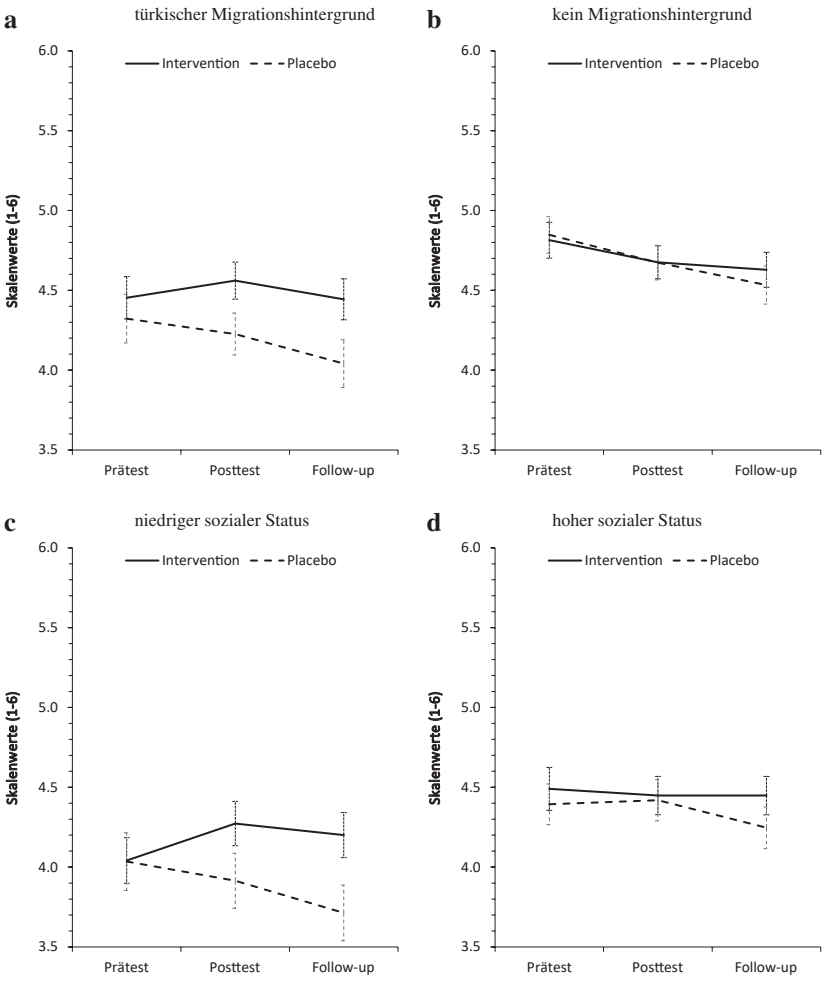

Abb. 8.2 Entwicklung der Einstellungen über alle drei Messzeitpunkte nach Bedingung (Interventions- versus Placebogruppe) und nach Schüler*innenherkunft (Fehlerbalken bilden das 95 % KI ab)

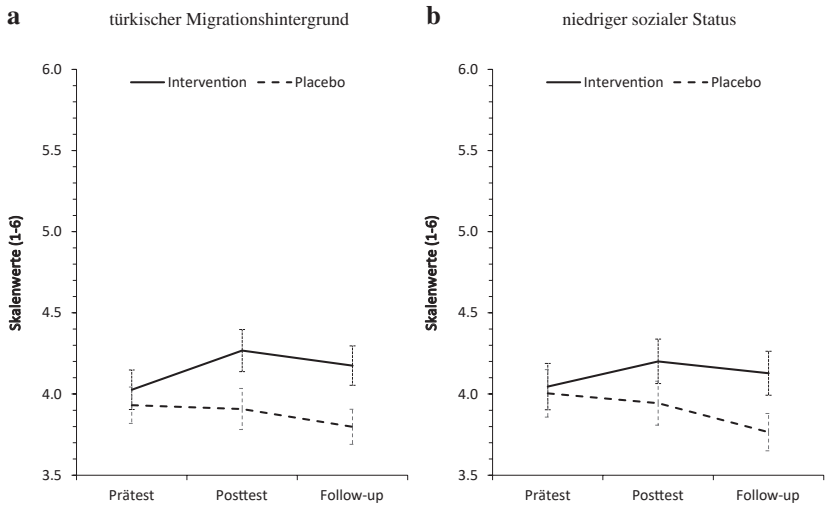

Abb. 8.3 Entwicklung der Stereotype über die Messzeitpunkte nach Bedingungen (Interventions- und Placebogruppe) und Schüler*innenherkunftsmerkmalen (Fehlerbalken bilden das 95 % KI ab)

Auch die Varianzanalyse zur Entwicklung der *herkunftsbezogenen Stereotype* ergab einen bedeutsamen Haupteffekt der Messzeitzeitpunkte und einen bedeutsamen Interaktionseffekt der Intervention mit den Messzeitpunkten, das heißt, die Stereotype veränderten sich über die drei Messzeitpunkte hinweg und sie veränderten sich in den beiden Versuchsgruppen in verschiedene Richtungen (vgl. Abb. 8.3). Es zeigten sich aber weder ein Haupteffekt noch Interaktionseffekte der Herkunftsmerkmale. Das bedeutet, die beiden Stereotype entwickelten sich je Versuchsgruppe nach dem gleichen Muster (vgl. Tab. 8.5; Abb. 8.3). Der Interaktionseffekt der Bedingungen mit den Messzeitpunkten blieb auch nach Kontrolle der impliziten Einstellungen in nahezu gleicher Höhe bestehen, das heißt, auch unter Berücksichtigung der impliziten Einstellungen der Studierenden gegenüber beiden Herkunftsmerkmalen zeigten sich unterschiedliche Entwicklungen der Stereotype im zeitlichen Verlauf in Abhängigkeit von der Interventionsbedingung.

Im Ergebnis resultierten sowohl für die Stereotype in Bezug auf den Migrationshintergrund als auch in Bezug auf die soziale Herkunft bedeutsame Unterschiede zwischen den Studierendengruppen nach der Intervention (vgl. Abb. 8.3). Dabei zeigte die Entwicklung der Stereotype ähnlich verschiedene

Tab. 8.5 Varianzanalysen mit Messwiederholung zu den Effekten der Messzeitpunkte, der Schüler*innenherkunft sowie der Interventionsbedingungen auf die Stereotype

	F	η^2
Effekte ohne Berücksichtigung der impliziten Einstellungen[a]		
Haupteffekte		
Messzeitpunkte	6,236**	0,06
Herkunft	0,022	0,00
Interaktionseffekte		
Messzeitpunkte*Bedingungen	9,490***	0,09
Herkunft*Bedingungen	0,568	0,00
Messzeitpunkte*Herkunft*Bedingungen	0,766	0,01
Effekte mit Berücksichtigung der impliziten Einstellungen[b]		
Haupteffekte		
Messzeitpunkte	2,373	0,02
Herkunft	0,005	0,00
Interaktionseffekte		
Messzeitpunkte*Bedingungen	8,410***	0,08
Herkunft*Bedingungen	0,427	0,00
Messzeitpunkte*Herkunft*Bedingungen	0,613	0,00

Anmerkungen. $N = 215$; [a] Messzeitpunkte: $df_{Hypothese} = 2$ *und* $df_{Fehler} = 204$; Herkunft: $df_{Hypothese} = 3$ *und* $df_{Fehler} = 205$. [b] Messzeitpunkte: $df_{Hypothese} = 2$ *und* $df_{Fehler} = 202$; Herkunft: $df_{Hypothese} = 3$ *und* $df_{Fehler} = 203$.
*** $p < ,001$. ** $p < ,01$. * $p < ,05$

Muster der beiden Gruppen wie ihre Entwicklung der Einstellungen. Die Studierenden in der Interventionsgruppe gaben im Posttest unmittelbar nach der Intervention im Vergleich zum Prätest vor der Intervention positivere Stereotype über Schüler*innen mit türkischem Migrationshintergrund an (Differenz 0,242, $p < ,05$ bonferroni-korrigiert, 95 % KI für die Differenz [0,096, 0,387], die sich im Follow-up deskriptiv zwar wieder ihren Stereotypen im Prätest annäherten, aber immer noch signifikant positiver waren als im Prätest (Differenz 0,149, $p < ,05$ bonferroni-korrigiert, 95 % KI für die Differenz [0,026, 0,272]), $F(2, 216) = 10{,}227$, $p < ,001$, $\eta^2 = ,087$. Gegenüber Schüler*innen mit niedrigem sozialem Status zeigte sich das gleiche Muster, allerdings nur als Tendenz, $F(1{,}868, 216) = 3{,}044$, $p = ,053$, $\eta^2 = ,027$; Greenhouse-Geisser

korrigiert. Die Studierenden der Placebogruppe zeigten wieder eine negative Entwicklung, das heißt, die Stereotype über Schüler*innen mit türkischem Migrationshintergrund und über Schüler*innen mit niedrigem sozialem Status wurden von Messzeitpunkt zu Messzeitpunkt weniger positiv, lagen aber auch im Follow-up immer noch über dem Skalenmittel, $F(2, 198) = 3{,}773$, $p < {,}05$, $\eta^2 = {,}037$ bzw. $F(2, 202) = 8{,}779$, $p < {,}001$, $\eta^2 = {,}080$. Diese negative Entwicklung war für ihre *Stereotype* über Schüler*innen mit niedrigem sozialem Status vom Posttest zum Follow-up am stärksten und etwa gleich stark negativ, wie die Entwicklung ihrer *Einstellungen* zu dieser Schüler*innengruppe.

8.5 Diskussion

Der vorliegende Beitrag verfolgte das Ziel, herkunftsbezogene Einstellungen und Stereotype von Lehramtsstudierenden durch eine kurze Intervention zu verändern. In einem experimentellen Prä-Post-Follow-up-Design erhielten Studierende des Grundschullehramts in der Interventionsgruppe unter anderem Rückmeldungen zu ihren individuellen Einstellungen sowie theoretische und empirische Informationen zur Bedeutung von herkunftsspezifischen Stereotypen und Einstellungen auf die Urteilsbildung. Zudem wurden weitere zentrale Aspekte zur Einstellungsänderung aus den theoretischen Modellen abgeleitet und in die Intervention integriert, unter anderem ein vorgestellter Kontakt und die Auseinandersetzung mit den eigenen Gefühlen. Die Placebogruppe erhielt hingegen Informationen zur Funktion automatisierter Prozesse der Urteilsbildung für die Dynamik der Unterrichtsinteraktion, die unter theoretischen Gesichtspunkten weder negiert noch pauschal negativ bewertet werden kann. Verweise auf Einstellungen und Stereotype wurden dabei absichtlich vermieden.

Die Befunde belegen, dass eine kurze Intervention bedeutsamen Einfluss auf die herkunftsbezogenen Einstellungen und Stereotype von Lehramtsstudierenden haben kann. So zeigten sich sowohl für die Einstellungen gegenüber Schüler*innen mit türkischem Migrationshintergrund als auch gegenüber Schüler*innen mit niedrigem sozialem Status Effekte der Intervention im Posttest, die auch im Follow-up noch gut zu erkennen waren, wenn auch weniger stark. Diese Effekte fanden sich ebenfalls für die herkunftsbezogenen Stereotype. Die Interaktionseffekte der Bedingungen mit den Messzeitpunkten sind zum Teil auf eine Veränderung der Einstellungen der Interventionsgruppe in eine positivere Richtung, teilweise auf eine Einstellungsänderung der Placebogruppe in eine negativere Richtung und zum Teil auf die Divergenz der Entwicklungen zurückzuführen. Positive Effekte auf Einstellungen und/oder Stereotype wurden auch aus

Maßnahmen in der Lehrkräftebildung berichtet, die curricular angelegt waren (z. B. Grottkau & Nickolai-May, 1989; Kumar & Hamer, 2013; Martin & Koppelman, 1991), das heißt als Pflicht- oder Wahlkurse mindestens ein Semester dauerten.

Auf den ersten Blick scheint die Entwicklung der Einstellungen und Stereotype in der Interventionsgruppe ähnlich wie die Entwicklung der Einstellungen in der Studie von Kumar und Hamer (2013) verlaufen zu sein. Die Effekte können angesichts der sehr unterschiedlichen Bedingungen aber nicht verglichen werden. In der Studie von Kumar und Hamer (2013) könnte die deutlich positive Entwicklung der Einstellungen bis zur Mitte des Curriculums als Effekt der ersten vertieften Auseinandersetzung mit einstellungsrelevanten Informationen interpretiert werden. Diese Erklärung liegt auch für die positiveren Einstellungen und Stereotype der Interventionsgruppe in unserer Studie nahe. In der Studie von Kumar und Hamer (2013) wurden die Studierenden aber von der Mitte bis zum Ende des Curriculums (eines ganzen Studienganges) mit weiteren einstellungsrelevanten Informationen konfrontiert und hatten eventuell auch praktische Lerngelegenheiten mit unmittelbaren Interaktionserfahrungen. Das Absinken der positiven Einstellungsausprägungen zum Ende des Curriculums könnte dann, den theoretischen Modellen folgend (HSM: Chaiken, 1980; ELM: Petty & Cacioppo, 1986), eher als Relativierung dieser Einstellungsänderungen im Zuge der weiteren Wissensvertiefung und Auseinandersetzung erklärt werden. Die Studierenden in unserer Interventionsgruppe hatten dagegen sehr wahrscheinlich außerhalb der Studie kaum oder wenig Gelegenheiten, die von uns angebotenen Informationen und Erfahrungen mit eigenen Einstellungen und Stereotypen systematisch zu vertiefen. Das Absinken der Effekte zwischen Posttest und Follow-up wäre dann eher auf geringere Nachhaltigkeit der Intervention, fehlende Lerngelegenheiten zur Vertiefung und unter Umständem gegenläufige Einflüsse im Studien- und Privatalltag der Studierenden zu interpretieren.

Ein Absinken von Trainingseffekten zwischen Posttest und Follow-up wird aus Interventionsstudien mit Trainings im engeren Sinne öfter berichtet (z. B. Tomonari, 2012). Für die Lehrkräftebildung (und auch für andere Kontexte, in denen herkunftsspezifische Einstellungen und Stereotype eine bedeutsame Rolle spielen) können hier mehrere Anschlussfragen gestellt werden: 1) Ist das Absinken der positiven Entwicklung ausschließlich als geringe oder fehlende Nachhaltigkeit der Intervention unabhängig von Alltagseinflüssen zu interpretieren oder wirken vielfältige Alltagseinflüsse zusätzlich in eine entgegengesetzte Richtung? 2) Welche Maßnahmen wirken besonders nachhaltig? 3) Welche Dauer und Intensität müssten diese Maßnahmen haben, um Einstellungen und Stereotype nachhaltig zu verändern und gegenüber möglichen gegenläufigen Alltagseinflüssen resistenter zu machen?

Im Unterschied zu den Ergebnissen von Haberman und Post (1992) änderten sich in unserer Studie nicht nur die Einstellungen der Studierenden gegenüber bestimmten Schüler*innengruppen, sondern auch ihre Eigenschaftszuschreibungen im Sinne von Stereotypen. Unter dem Vorbehalt einer notwendigen differenziellen Prüfung kann diskutiert werden, ob die Kombination von Einflussfaktoren gemäß der Intergruppenkontakttheorie (vgl. Pettigrew & Tropp, 2006, 2008) mit Einflussfaktoren gemäß der Dual-Process-Theorien verschiedene Ebenen adressiert und verschiedene Prozesse angeregt hat. Für verschiedene Ebenen und Prozesse spricht der Teileffekt im Experiment von Soble et al. (2011), in dem zwar Empathie und Schuldgefühle der Studierenden stiegen, Furcht vor Begegnungen aber nicht reduziert werden konnte. Die Schlussfolgerung von Pettigrew und Tropp (2006) aus den Ergebnissen ihrer Metanalyse, „reducing negative feelings such as anxiety and threat represents an important means by which intergroup contact diminishes prejudice“ (S. 767), verweist auf die affektive Ebene von Einstellungen und Stereotypen, die offenbar auch in Experimenten wirksam angesprochen werden kann (z. B. mit der fast-friendship-procedure von Kotzur et al., 2019). Ob in unserer Intervention die Aufforderung, sich die eigenen Emotionen bei erwartungswidrigem Verhalten vorgestellter Personen bewusst zu machen, diese Ebene angesprochen hat, können wir allerdings nicht differenziell aufklären. Die kognitive Ebene der herkunftsbezogenen Einstellungen wurde sicherlich mit den Informationsbausteinen über Stereotype, Urteilsverzerrungen und Bildungsungerechtigkeit angesprochen. Unsere Ergebnisse der Kombination beider Ebenen, so denn verschiedene Ebenen und Prozesse anzunehmen sind, stehen in Einklang mit den Ergebnissen der Studie von Hassanein (2015), der ebenfalls nur für die Kombination einstellungsrelevanter Informationen mit Kontakt positive Effekte auf Einstellungen fand, nicht aber von Kontakt ohne Elaboration von Informationen, wobei hier offenbleiben muss, inwieweit die Kontakte der Lehrkräfte in der Studie von Hassanein (2015) die theoretisch postulierten Bedingungen der Intergruppenkontakttheorie erfüllten.

Im Unterschied zu den Ergebnissen der Studie von Lucas (2011), in der sich Einstellungen zukünftiger Lehrkräfte zu Inklusion, nicht aber zu kulturellem Pluralismus änderten, zeigte die Intervention in unserer Studie sowohl Effekte auf die Einstellungen und Stereotype gegenüber Schüler*innen mit türkischem Migrationshintergrund als auch gegenüber Schüler*innen mit niedrigem sozialem Status. Gleichzeitig fielen die Korrelationen dieser beiden Einstellungen auf der einen Seite und die Korrelationen der ‚entgegengesetzten‘ Einstellungen gegenüber Schüler*innen ohne Migrationshintergrund und gegenüber Schüler*innen mit hohem sozialem Status auf der anderen Seite deskriptiv jeweils höher aus. Zusammen verweisen diese Ergebnisse auf das grundsätzliche Problem, dass

diese beiden Herkunftsmerkmale beziehungsweise sozialen Kategorien offenbar – und vermutlich nicht nur von (angehenden) Lehrkräften – als zusammenhängend wahrgenommen werden. In Bezug auf den tatsächlichen Bildungserfolg zeigen sich auch nach wie vor statistische Zusammenhänge (OECD, 2019).

Die Befunde dieser Studie deuten darauf hin, dass durch die Intervention ein negativer Trend im Verlauf des ersten Semesters, der sich in der Placebogruppe zeigte, in der Interventionsgruppe zum Teil verhindert und zum Teil reduziert werden konnte. Welche Faktoren dazu geführt haben könnten, dass die Studierenden der Placebogruppe negativere Einstellungen allen vier Schüler*innengruppen gegenüber entwickelten, kann durch diese Studie nicht beantwortet werden. Denkbar wäre zum Beispiel die Konfrontation mit Untersuchungen zu herkunftsbezogenen Disparitäten im Verlauf des Semesters, die, wenn sie nicht mit Handlungsmöglichkeiten zur Verringerung dieser Disparitäten verknüpft wird, als Belastung mit einer Verantwortung wahrgenommen werden könnte, die gar nicht erfüllbar erscheint und dann zu negativen Gefühlen gegenüber dem gesamten Thema führen könnte. Mediale Berichte, die auf die Gefahr einer Verschärfung dieser Disparitäten im Zuge des Wechselunterrichts oder Homeschoolings durch die Corona-Pandemie hinweisen, könnten die Wahrnehmung eines solchen Dilemmas noch verschärft haben.

Denkbar scheint auch, dass hier ein Designeffekt zu Buche schlägt, dem in der Interventionsgruppe in Bezug auf die im Fokus stehenden benachteiligten Schüler*innengruppen zwar entgegengewirkt werden konnte (Messzeitpunkt 2), aber in Bezug auf die Einstellungen zu Schüler*innen ohne Migrationshintergrund sowie Schüler*innen mit hohem sozio-ökonomischem Status in beiden Gruppen parallel verlief und eventuell anteilig zum leichten Absinken der positiven Effekte in der Interventionsgruppe bis zum Follow-up beigetragen hat: Die Studierenden haben die erste Erhebung im ersten Drittel ihres ersten Semesters an der Universität durchgeführt. Sie waren vermutlich aufgeregt und – trotz der Bedingungen – eher in positiver Stimmung. Die Interventionen (Experimental und Placebo) und die zweite Erhebung haben sie im letzten Drittel der Veranstaltungszeit durchgeführt und die Nacherhebung nach Ende der Veranstaltungszeit und nach verschiedenen Semesterprüfungen. Es ist unwahrscheinlich, dass die Mehrheit der Teilnehmer*innen zu diesem Zeitpunkt noch in einer positiven Gestimmtheit durch den Studienbeginn war und Stimmungen gelten in der Einstellungsforschung als einer von vielen Einflussfaktoren (vgl. z. B. Eagly & Chaiken, 1993; Petty & Cacioppo, 1986). In folgenden Untersuchungen sollte die Stimmung als Kontrollvariable gegebenenfalls berücksichtigt werden. Für Spekulationen, ob einzelne und gegebenenfalls welche spezifischen Merkmale der Kurzintervention dieser Entwicklung zum zweiten Messzeitpunkt entgegen-

gewirkt haben könnten, bieten die theoretischen Grundlagen und die empirischen Ergebnisse anderer Studien keine Anhaltspunkte.

Einschränkend ist als Limitation insbesondere anzumerken, dass sich im Prätest bedeutsame Unterschiede in den impliziten Einstellungen zwischen der Interventions- und der Placebogruppe gezeigt hatten. Die impliziten Einstellungen wurden allerdings in die Analysen einbezogen. Dabei blieben die wesentlichen Befundmuster bestehen, auch wenn sich die Effekte verringerten. Die positiveren impliziten Einstellungen in der Interventionsgruppe stehen vermutlich in Zusammenhang mit der ungleichen Verteilung der Studierenden mit Migrationshintergrund in den beiden Gruppen. Zwar ist der Anteil in dieser Untersuchung insgesamt nur sehr gering, was nur wenig Einfluss auf die Ergebnismuster vermuten lässt, wie auch die Kontrollanalysen ergaben. Dennoch wäre für zukünftige Studien ein ausbalanciertes Verhältnis der Herkunftsgruppen der Studierenden in den Versuchsgruppen anzustreben, und mindestens ebenso wichtig wäre zu prüfen, ob sich in Gruppen von Studierenden mit eigenem türkischen Migrationshintergrund und in Gruppen mit anderem Migrationshintergrund die gleichen Effekte zeigen oder keine oder ganz andere Veränderungen stattfinden. Für aussagekräftige Befunde hierzu wäre ein höherer Anteil dieser Studierendengruppe in der Stichprobe notwendig.

Insgesamt bleibt zu vermerken, dass trotz einer zeitlich und inhaltlich stark begrenzten Intervention Effekte auf die Einstellungen und Stereotype der Studierenden erzielt werden konnten. Offenbleiben muss hier, ob bestimmte Elemente und gegebenenfalls welche Elemente der Intervention für die Effekte maßgeblich waren. Dies müsste in weiteren experimentellen Untersuchungen systematisch geklärt werden, zum Beispiel, ob und inwiefern die individuellen Rückmeldungen zu eigenen Einstellungen einen Mehrwert gegenüber der reinen Darbietung von Informationen zu Einstellungen, Stereotypen und herkunftsabhängigen Urteilsverzerrungen haben oder vielleicht sogar ausreichen, um eine Einstellungsänderung anzuregen. Des Weiteren wurden unterschiedliche Interventionsteile aus den theoretischen Modellen abgeleitet, deren Wirksamkeit unter anderen Bedingungen (zeitlich, andere Personen, andere Themen) auch als eigenständige Einflussfaktoren empirisch belegt wurden (vgl. Bohner & Wänke, 2006; Eagly & Chaiken, 1993; Pettigrew & Tropp, 2006; Petty & Cacioppo, 1986). Mit zukünftigen Interventionen müsste aufgeklärt werden, welche der Elemente und gegebenenfalls welche Kombination notwendig und/oder hinreichend ist, um ähnliche Effekte mit einer kurzen Intervention zu erzielen. Weitere Studien sollten allerdings zunächst die Wirksamkeit der kurzen Intervention replizieren, beispielsweise an anderen Studienorten und in anderen Regionen. Zudem sollte die Intervention auch auf andere Populationen ausgeweitet werden (z. B. erfahrene

Lehrkräfte, Lehramtsstudierende anderer Schularten). Und in Folgestudien sollten mögliche Effekte der Studienzeitpunkte auf das Antwortverhalten ausgeschlossen sowie weiterführende Erkenntnisse zur Nachhaltigkeit der Intervention durch mehr Messzeitpunkte angestrebt werden.

Im Hinblick auf die Aus- und Weiterbildung von Lehrkräften deuten die Befunde darauf hin, dass bei der Thematisierung herkunftsassoziierter Einstellungen und Stereotype, zum Beispiel im Zusammenhang mit Schüler*innenbeurteilungen, wissenschaftlich fundierte Informationen zu Effekten von Einstellungen durch eine objektive Rückmeldung über die eigenen Einstellungen und eine Auseinandersetzung mit eigenen Emotionen sinnvoll ergänzt werden kann. Die Reflexion eigener Einstellungen anhand einer objektiven Rückmeldung stellt möglicherweise einen wichtigen Schritt zur Sensibilisierung für bislang unerkannte individuelle Faktoren im Umgang mit einer heterogenen Schüler*innenschaft dar. Dabei sollte allerdings ebenfalls deutlich werden, dass Einstellungen und Stereotype nicht nur in Bezug auf den Migrationshintergrund oder die soziale Herkunft von Schüler*innen von Bedeutung sind, sondern auch weitere Schüler*innenmerkmale (z. B. Geschlecht, sonderpädagogischer Förderbedarf, Religion), insbesondere in ihrer Kombination relevant sind. Derzeit ist das Thema Heterogenität der Schüler*innenschaft sowie Einstellungen und Stereotype in diesem Zusammenhang sicherlich an vielen Standorten der Lehrkraftausbildung präsent, allerdings meist durch frei wählbare Kurse, die somit in der Regel nicht garantieren, dass alle Studierenden durch Informationen und Reflexionsanregung in Kontakt mit dieser Thematik kommen. Die hier vorgestellte Untersuchung konnte zeigen, dass bereits kurze Anteile im Studium, wenn sie systematisch theoriegeleitet aufgebaut sind, nicht nur Reflexionsprozesse anregen können, sondern auch Effekte auf Einstellungen und Stereotype entfalten können.

Literatur

Ajzen, I., & Fishbein, M. (1980). *Understanding attitudes and predicting social behavior*. Prentice-Hall.

Allport, G. (1954). *The nature of prejudice*. Addison-Wesley.

Anderson, N. H. (1971). Integration theory and attitudes change. *Psychological Review, 78*(3), 171–206. https://doi.org/10.1037/h0030834.

Artelt, C., & Gräsel, C. (2009). Diagnostische Kompetenz von Lehrkräften. *Zeitschrift für Pädagogische Psychologie, 23*(34), 157–160. https://doi.org/10.1024/1010-0652.23.34.157.

Baron, A. S. (2015). Constraints on the development if implicit intergroup attitudes. *Child Development Perspectives, 9*(1), 50–54. https://doi.org/10.1111/cdep.12105.

Bezrukova, K., Spell, C. S., Perry, J. L., & Jehn, K. A. (2016). A meta-analytical integration of over 40 years of research on diversity training evaluation. *Psychological Bulletin, 142*(11), 1227–1274. https://doi.org/10.1037/bul0000067.

Böhmer, I., Gräsel, C., Krolak-Schwerdt, S., Hörstermann, T., & Glock, S. (2017). Teachers' school tracking decisions. In D. Leutner, J. Fleischer, J. Grünkorn, E. Klieme, & D. Leutner (Hrsg.), *Competence assessment in education. Research, models and instruments* (S. 131–147). Springer International Publishing.

Bogardus, E. S. (1933). A social distance scale. *Sociology & Social Research, 17,* 265–271.

Bohner, G. & Wänke, M. (2006). *Attitudes and attitude change* (repr.). Psychology Press.

Bonefeld, M., & Dickhäuser, O. (2018). (Biased) grading of students' performance: Students' names, performance level and implicit attitudes. *Frontiers in Psychology, 9,* 481. https://doi.org/10.3389/fpsyg.2018.00481.

Bonefeld, M., & Karst, K. (2020). Döner vs. Schweinebraten: Stereotype von (angehenden) Lehrkräften über Personen deutscher und türkischer Herkunft im Vergleich. In S. Glock & H. Kleen (Hrsg.), *Stereotype in der Schule* (S. 159–190). Springer.

Brophy, J. E., & Good, T. L. (1974). *Teacher-student relationships. Causes and consequences*. Holt, Rinehart and Winston.

Bundesamt für Statistik. (2017). *ISCO 08 (International Standard Classification of Occupations)*. Schweizerische Eidgenossenschaft. https://www.bfs.admin.ch/bfsstatic/dam/assets/4082534/master.

Chaiken, S. (1980). Heuristic versus systematic information processing and the use of source versus message cues in persuasion. *Journal of Personality and Social Psychology, 39*(5), 752–766. https://doi.org/10.1037/0022-3514.39.5.752.

Cohen, A. R., Stotland, E., & Wolfe, D. M. (1955). An experimental investigation of need for cognition. *The Journal of Abnormal and Social Psychology, 51*(2), 291–294. https://doi.org/10.1037/h0042761.

Cuddy, A. J. C., Fiske, S. T., & Glick, P. (2008). Warmth and competence as universal dimensions of social perception: The stereotype content model and the BIAS map. In M. P. Zanna (Hrsg.), *Advances in Experimental Social Psychology* (Bd. 40, S. 61–149). Academic. https://doi.org/10.1016/S0065-2601(07)00002-0.

Diefenbach, H. (2010). *Kinder- und Jugendliche aus Migrantenfamilien im deutschen Bildungssystem: Erklärungen und empirische Befunde*. Springer.

Eagly, A. H., & Chaiken, S. (1993). *The psychology of attitudes*. Harcourt Brace Jovanovich.

Fazio, R. H. (2007). Attitudes as object-evaluation associations of varying strength. *Social Cognition, 25*(5), 603–637. https://doi.org/10.1521/soco.2007.25.5.603.

Festinger, L. (1954). A theory of social comparison processes. *Human Relations, 7*(2), 117–140. https://doi.org/10.1177/001872675400700202.

Fiske, S. T., & Neuberg, S. L. (1990). A continuum of impression formation, from category-based to individuating processes. Influences of information and motivation on attention and interpretation. In M. P. Zanna (Hrsg.), *Advances in experimental social psychology* (Bd. 23, S. 1–74). Academic Press.

Frenzel, A. C., & Taxer, J. L. (2018). „Das kannst Du doch besser!": Effekte von Lehrerärger und -mitleid auf das Misserfolgs-Attributionsmuster und die Persistenz von

Lernenden. In G. Hagenauer & T. Hascher (Hrsg.), *Emotionen und Emotionsregulation in Schule und Hochschule* (S. 181–196). Waxmann.

Friese, M., Hofmann, W., & Schmitt, M. (2008). When and why do implicit measures predict behaviour? Empirical evidence for the moderating role of opportunity, motivation, and process reliance. *European Journal of Social Psychology, 19*(1), 285–338. https://doi.org/10.1080/10463280802556958.

Ganzeboom, H. B. G., De Graaf, P. M., & Treiman, D. J. (1992). A standard international socio-economic index of occupational status. *Social Science Research, 21*(1), 1–56. https://doi.org/10.1016/0049-089X(92)90017-B.

Gawronski, B., & Bodenhausen, G. V. (2006). Associative and propositional processes in evaluation: An integrative review of implicit and explicit attitude change. *Psychological Bulletin, 132*(5), 692–731. https://doi.org/10.1037/0033-2909.132.5.692.

Glock, S., & Karbach, J. (2015). Preservice teachers' implicit attitudes towards racial minority students: Evidence from three implicit measures. *Studies in Educational Evaluation, 45,* 55–61. https://doi.org/10.1016/j.stueduc.2015.03.006.

Glock, S., Krolak-Schwerdt, S., & Hörstermann, T. (2016). The higher the SES the better? Implicit and explicit attitudes influence preservice teachers' judgment of students. In C. Fields (Hrsg.), *Stereotypes and stereotyping. Misperceptions, percspectives and role of social media* (S. 1–20). Nova.

Glock, S., Krolak-Schwerdt, S., Klapproth, F., & Böhmer, M. (2013). Beyond judgment bias: How students' ethnicity and academic profile consistency influence teachers' tracking judgments. *Social Psychology of Education, 16*(4), 555–573. https://doi.org/10.1007/s11218-013-9227-5.

Goedderz, A., & Hahn, A. (2020). Haben wir alle versteckte Vorurteile? Implizite Einstellungen und ihre Bedeutung in der Gesellschaft. *In-Mind-Magazin, 1*. https://de.in-mind.org/article/haben-wir-alle-versteckte-vorurteile-implizite-einstellungen-und-ihre-bedeutung-in-der?page=2&utm_source=rssfeed&utm_medium=rssfeed&utm_campaign=general.

Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition. Attitudes, self-esteem, and stereotypes. *Psychological Bulletin, 102*(1), 4–27. https://doi.org/10.1037//0033-295x.102.1.4.

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology, 74*(6), 1464–1480. https://doi.org/10.1037/0022-3514.74.6.1464.

Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the implicit association test: An improved scoring algorithm. *Journal of Personality and Social Psychology, 85*(2), 197–216. https://doi.org/10.1037/0022-3514.85.2.197.

Gresch, C., & Becker, M. (2010). Sozial- und leistungsbedingte Disparitäten im Übergangsverhalten bei türkischstämmigen Kindern und Kindern aus (Spät-) Aussiedlerfamilien. In K. Maaz, J. Baumert, C. Gresch, & N. McElvany (Hrsg.), *Der Übergang von der Grundschule in die weiterführende Schule: Leistungsgerechtigkeit und regionale, soziale und ethnisch-kulturelle Disparitäten* (S. 181–200). BMBF.

Grottkau, B. J., & Nickolai-May, S. (1989). An empirical analysis of a multicultural education paradigm for preservice teachers. *Educational Research Quarterly, 13*(4), 27–33.

Haberman, M., & Post, L. (1992). Does direct experience change education students' perceptions of low-income minority children. *Midwestern Educational Researcher, 5*(2), 29–31.

Hahn, A., Judd, C., Hirsh, H. K., & Blair, I. V. (2014). Awareness of implicit attitudes. *Journal of Experimental Psychology: General, 143*(3), 1369–1392. https://doi.org/10.1037/a0035028.

Hammond, C. (2010). Multicultural education and the notion of privilege. An examination of White student attitude shift and identity development. *Dissertation Abstracts International Section A: Humanities and Social Sciences, 71*(3-A), 844.

Hassanein, E. E. A. (2015). Changing teachers' negative attitudes toward persons with intellectual disabilities. *Behavior Modification, 39*(3), 367–389. https://doi.org/10.1177/0145445514559929.

Hofmann, W., Gawronski, B., Gschwendner, T., Le, H., & Schmitt, M. (2005). A meta-analysis on the correlation between the implicit association test and explicit self-report measures. *Personality & Social Psychology Bulletin, 31*(10), 1369–1385. https://doi.org/10.1177/0146167205275613.

Ingham, A. G., Levinger, G., Graves, J., & Peckham, V. (1974). The Ringelmann effect. Studies of group size and group performance. *Journal of Experimental Social Psychology, 10*(4), 371–384. https://doi.org/10.1016/0022-1031(74)90033-X.

Jussim, L. (1990). Expectancies and social issues: Introduction. *Journal of Social Issues, 46*(2), 1–8. https://doi.org/10.1111/j.1540-4560.1990.tb01920.x.

Jussim, L., & Harber, K. D. (2005). Teacher expectations and self-fulfilling prophecies: Knowns and unknowns, resolved and unresolved controversies. *Personality and Social Psychology Review, 9*(2), 131–155. https://doi.org/10.1207/s15327957pspr0902_3.

Karst, K., & Bonefeld, M. (2020). Stereotype, Urteile und Urteilsakkuratheit von Lehrkräften: Eine Zusammenschau im Rahmen des Heterogenitätsdiskurses. In S. Glock & H. Kleen (Hrsg.), *Stereotype in der Schule* (S. 281–308). Springer.

Kleen, H., Bonefeld, M., Glock, S., & Dickhäuser, O. (2019). Implicit and explicit attitudes toward Turkish students in Germany as a function of teachers' ethnicity. *Social Psychology of Education, 22*(4), 883–899. https://doi.org/10.1007/s11218-019-09502-9.

Kotzur, P. F., Schäfer, S. J., & Wagner, U. (2019). Meeting a nice asylum seeker. Intergroup contact changes stereotype content perceptions and associated emotional prejudices, and encourages solidarity-based collective action intentions. *British Journal of Social Psychology, 58*(3), 668–690. https://doi.org/10.1111/bjso.12304.

Kumar, R., & Hamer, L. (2013). Preservice teachers' attitudes and beliefs toward student diversity and proposed instructional practices: A sequential design study. *Journal of Teacher Education, 64*(2), 162–177. https://doi.org/10.1177/0022487112466899.

Kunda, Z., & Thagard, P. (1996). Forming impressions from stereotypes, traits, and behaviors: A parallel-constraint-satisfaction theory. *Psychological Review, 103*(2), 284–308. https://doi.org/10.1037/0033-295x.103.2.284.

Lehmann-Grube, S. K., Hartinger, A., Grassinger, R., Brandl-Bredenbeck, H. P., Ohl, U., Riegger, M., & Dresel, M. (im Druck). Einstellungen Lehramtsstudierender zu Heterogenität: Modell, Messinstrument und Ergebnisse der Validierung. Erscheint in A. Hartinger, M. Dresel, E. Matthes, K. Peuschel, & U. E. Nett (Hrsg.), *Lehrkräfteprofessionalität im Umgang mit Heterogenität: Theoretische Konzepte, Förderansätze, empirische Befunde*. Waxmann.

Lucas, D. J. (2011). *The effects of a service-learning introductory diversity course on pre-service teachers' attitudes toward teaching diverse student populations*. Liberty University. https://digitalcommons.liberty.edu/cgi/viewcontent.cgi?article=1513&context=doctoral.

Maaz, K., Baeriswyl, F., & Trautwein, U. (2011). *Herkunft zensiert? Leistungsdiagnostik und soziale Ungleichheiten in der Schule. Eine Studie im Auftrag der Vodafone Stiftung Deutschland*. Vodafone Stiftung.

Martin, R. J., & Koppelman, K. (1991). The impact of a human relations/multicultural education course on the attitudes of prospective teachers. *The Journal of Intergroup Relations, 18*(1), 16–27.

Miles, E., & Crisp, R. J. (2014). A meta-analytic test of the imagined contact hypothesis. *Group Processes & Intergroup Relations, 17*(1), 3–26. https://doi.org/10.1177/1368430213510573.

Millisecond Software. (2016) Inquisit 5 Web (Windows) [Computer software]: Millisecond Software. http://www.millisecond.com/.

Organisation für wirtschaftliche Zusammenarbeit und Entwicklung [OECD]. (2019). *PISA 2018 results (Bd. 2): Where all students can succeed*. OECD Publishing.

Peak, H. (1955). Attitude and motivation. In M. R. Jones (Hrsg.), *Nebraska Symposium on Motivation* (Bd. 3, S. 149–188). University of Nebraska Press.

Pettigrew, T. F., & Tropp, L. R. (2006). A meta-analytic test of intergroup contact theory. *Journal of Personality and Social Psychology, 90*(5), 751–783. https://doi.org/10.1037/0022-3514.90.5.751.

Pettigrew, T. F., & Tropp, L. R. (2008). How does intergroup contact reduce prejudice? Meta-analytic tests of three mediators. *European Journal of Social Psychology, 38*(6), 922–934. https://doi.org/10.1002/ejsp.504.

Petty, R. E., & Cacioppo, J. T. (1986). The elaboration likelihood model of persuasion. In L. Berkowitz (Hrsg.), Advances in experimental social psychology (Bd. 19, S. 123–205). Academic Press. https://doi.org/10.1016/S0065-2601(08)60214-2

Rubie-Davies, C., Hattie, J., & Hamilton, R. (2006). Expecting the best for students: Teacher expectations and academic outcomes. *The British Journal of Educational Psychology, 76*(3), 429–444. https://doi.org/10.1348/000709905X53589.

Schrader, F.-W., & Helmke, A. (2001). Alltägliche Leistungsbeurteilung durch Lehrer. In F. E. Weinert (Hrsg.), *Leistungsmessung in Schulen* (S. 45–58). Beltz.

Soble, J. R., Spanierman, L. B., & Liao, H.-Y. (2011). Effects of a brief video intervention on White university students' racial attitudes. *Journal of Counseling Psychology, 58*(1), 151–157. https://doi.org/10.1037/a0021158.

Spinath, B. (2005). Akkuratheit der Einschätzung von Schülermerkmalen durch Lehrer und das Konstrukt der diagnostischen Kompetenz. *Zeitschrift für Pädagogische Psychologie, 19*(1/2), 85–95. https://doi.org/10.1024/1010-0652.19.12.85.

Südkamp, A., Kaiser, J., & Möller, J. (2012). Accuracy of teachers' judgments of students' academic achievement: A meta-analysis. *Journal of Educational Psychology, 104*(3), 743–762. https://doi.org/10.1037/a0027627.

Tajfel, H. (1969). Cognitive aspects of prejudice. *Journal of Social Issues, 25*(4), 79–97. https://doi.org/10.1111/j.1540-4560.1969.tb00620.x.

Tobisch, A. (2019). *Herkunftsassoziierte Lehrkrafturteile und -erwartungen: Soziale Kognitionen und Urteilsbildungsprozesse im Kontext einer ethnisch und sozial hetero-*

genen Schülerschaft. Universität Augsburg. https://opus.bibliothek.uni-augsburg.de/opus4/63835

Tobisch, A., & Dresel, M. (2017). Negatively or positively biased? Dependencies of teachers' judgments and expectations based on students' ethnic and social backgrounds. *Social Psychology of Education, 20*(4), 731–752. https://doi.org/10.1007/s11218-017-9392-z.

Tobisch, A., & Dresel, M. (2020). Fleißig oder faul? Welche Einstellungen haben angehende Lehrkräfte gegenüber Schüler*innen aus unterschiedlichen sozialen Schichten. In S. Glock & H. Kleen (Hrsg.), *Stereotype in der Schule* (S. 33–158). Springer.

Tomonari, D. A. M. (2012). *A professional development program evaluation: Teacher efficacy, learning, and transfer*. University of Southern California. http://www.proquest.com/en-US/products/dissertations/individuals.shtml

Ülger, Z., Dette-Hagenmeyer, D., & Reichle, B. (2016). Reduktion von Vorurteilen und Modifikation von interethnischen Einstellungen: Eine Möglichkeit zur Förderung des Bildungserfolgs von Migranten? *Psychologie in Erziehung und Unterricht, 63*(3), 169–179. https://doi.org/10.2378/peu2016.art15d.

Wahl, D. (1991). *Handeln unter Druck. Der weite Weg vom Wissen zum Handeln bei Lehrern, Hochschullehrern und Erwachsenenbildnern*. Beltz.

Wigfield, A., & Eccles, J. S. (2000). Expectancy-value theory of achievement motivation. *Contemporary Educational Psychology, 25*(1), 68–81. https://doi.org/10.1006/ceps.1999.1015.