

# Die Qualität des Lehrkraft- verhaltens in Lehrkraft-Kind-Dyaden

Unterschiede zwischen Kindern in Abhängigkeit ihrer Noten, ihres Geschlechts, ihrer sprachlichen Herkunft und eines sonderpädagogischen Förderbedarfs

Bettina Hannover , Karoline Koeppen und Madeleine Kreutzmann

Freie Universität Berlin, Deutschland

**Zusammenfassung:** Das Verhalten von Lehrkräften in Lehrkraft-Kind-Dyaden kann auf den beiden universalen orthogonalen Dimensionen interpersonalen Verhaltens beschrieben werden: Communion (z.B. Wärme, Bedürfnisbefriedigung) und Agency (z.B. Lenkung, Kontrolle). Bestimmte Kombinationen auf diesen Dimensionen haben sich als günstig erwiesen: Schülerinnen und Schüler lernen besonders motiviert und viel, wenn ihre Lehrkraft starke Communion zeigt. Weniger eindeutig ist, ob eine moderat starke oder starke Agency der Lehrkraft besonders vorteilhaft ist. Wir untersuchen unter Verwendung des neu konstruierten *Fragebogens zum Lehrkraftverhalten in dyadischen Lehrkraft-Lernen-Beziehungen*, ob Lehrkräfte ihre Agency komplementär zur Kompetenz des Kindes ausrichten und ob sich dies auch in stärkerer Agency gegenüber Gruppen von Kindern niederschlägt, deren mittlere Kompetenz geringer ist, nämlich (a) Kindern mit sonderpädagogischen Förderbedarf, (b) Kindern mit nichtdeutscher Erstsprache und (c) Jungen. Zweiundsiebzig Lehrkräfte beschrieben ihr Verhalten gegenüber fünf Kindern ihrer Klasse ( $N = 302$ ) auf jeweils 13 Items, die in einem Zirkumplex acht Facetten unterschiedlicher Kombinationen von Communion und Agency erfassen. Noten wurden als Proxy für Kompetenzen genutzt. Wie erwartet korrelierten Facetten mit sehr starker Agency positiv ( $r = .54$  und  $.65$ ) und Facetten mit sehr schwacher Agency negativ ( $r = -.46$  und  $-.59$ ) mit den Noten des Kindes. Nach Aggregation der Angaben über die Gruppen von Kindern zeigte sich, dass Lehrkräfte gegenüber Kindern mit sonderpädagogischem Förderbedarf oder mit nichtdeutscher Erstsprache auf Verhaltensfacetten mit starker Agency höhere Ausprägungen aufwiesen als gegenüber Kindern ohne das entsprechende Merkmal, unabhängig von der Stärke der Communion der Facetten. Diese Unterschiede verschwanden meist nach Kontrolle der Noten. Gegenüber Jungen (relativ zu Mädchen) gaben die Lehrkräfte höhere Ausprägungen auf Facetten mit (moderat) starker Agency bei gleichzeitig nur (moderat) schwacher Communion an, auch nach der Berücksichtigung der Noten. Die Ergebnisse werden bzgl. der Forschung (a) zum Zusammenhang zwischen Lehrkraftverhalten und Motivation der Lernenden und (b) zu den Ursachen des geringeren Bildungserfolgs von Jungen diskutiert.

**Schlüsselwörter:** Lehrkraftverhalten, Communion und Agency, interpersonale Komplementarität, Geschlechtsunterschiede, Sonderpädagogischer Förderbedarf, Erstsprache

**Agency and communion of teacher behavior in dyadic interactions with primary school children. Differences between children depending on their grades, gender, linguistic background, and special educational needs.**

**Abstract:** The behavior of teachers in teacher-child dyads can be described on the two universal orthogonal dimensions of interpersonal behavior: Communion (e.g., warmth, need satisfaction) and agency (e.g., steering, control). Certain combinations of manifestations on these dimensions are particularly favorable: students are motivated and learn more when their teacher demonstrates strong communion. It is less clear, however, whether a moderately strong or strong agency of the teacher is particularly beneficial. Via a newly developed questionnaire, we investigate whether teachers align their agency complementarily to the child's competence and whether this translates into stronger teacher agency towards groups of children whose mean competence is lower, namely (a) children with special educational needs, (b) children with a non-German first language, and (c) boys. Seventy-two teachers described their behavior toward five children in their class ( $N = 302$ ) on 13 items each which capture different combinations of communion and agency in a circumplex of eight facets. As expected, facets with very strong agency correlated positively ( $r = .54$  and  $.65$ ), and facets with very weak agency correlated negatively ( $r = -.46$  and  $-.59$ ) with the child's grades. After aggregating teachers' reports across groups of children, we found that teachers exhibited higher levels of behavior on facets with strong agency towards children with special educational needs or with non-German first language, irrespective of these facets' communion. Most of these differences disappeared once children's grades were controlled for. In comparison with their behavior towards girls, teachers showed significantly stronger behavioral manifestations towards boys on facets characterized by strong agency and, simultaneously, only (moderately) weak communion, even after controlling for children's grades. Results are discussed against the background of research (a) on the connection between teacher behavior and students' motivation, and (b) on the reasons for the lower success of boys at school.

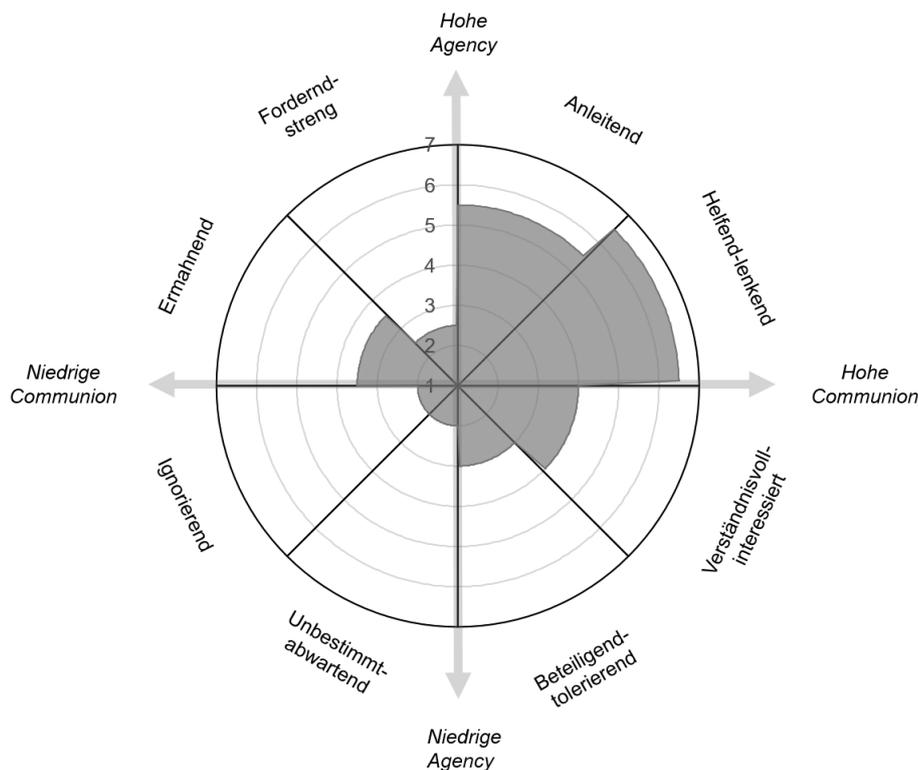
**Keywords:** teacher behavior, communion and agency, interpersonal complementarity, gender differences, special educational needs, first language

Insbesondere Grundschulklassen sind oft von großer Heterogenität der Schülerinnen und Schüler (SuS) gekennzeichnet (für Berlin siehe z.B. Thoren, Hannover & Brunner, 2019). Die inklusive Grundschule hat den Anspruch, für Kinder mit unterschiedlichen Lernvoraussetzungen eine gleichermaßen qualitätsvolle Lernsituation zu gewährleisten. Ein wesentlicher Aspekt ist dabei die Qualität der interpersonalen Beziehung zwischen Lehrkraft und dem einzelnen Kind. Aber ist tatsächlich jedes Kind in eine qualitätsvolle Beziehung zur Lehrperson eingebunden? Gelingt es Lehrkräften, in dyadischen Interaktionen mit jedem einzelnen Kind ein Verhalten zu zeigen, von dem angenommen werden kann, dass es für die soziale und akademische Entwicklung des Kindes günstig ist? Wir untersuchen, wie Lehrkräfte an deutschen Grundschulen ihr Verhalten gegenüber einzelnen Kindern ihrer Klasse beschreiben und wie die Qualität dieses Verhaltens zu bewerten ist.

## Dimensionen zur Beschreibung von Lehrkraftverhaltensweisen

Das Verhalten von Erziehungs- oder Lehrpersonen ist in unterschiedlichen Forschungstraditionen untersucht

worden: der Bindungsforschung, der Erziehungsstilforschung, der Führungstilforschung, der Forschung über Lehrstile, der Forschung zur Selbstbestimmungstheorie und zur Interpersonalen Theorie (Überblick: Knierim, Raufelder & Wettstein, 2017). Zur Beschreibung möglicher Verhaltensweisen werden typischerweise zwei orthogonale Achsen genutzt, von denen die eine auf Qualitäten bezogen ist, die für die Herstellung und Aufrechterhaltung sozialer Beziehungen erforderlich sind (z.B. Wärme, Herzlichkeit, Sensitivität, Affiliation, Liebe), und die andere auf Qualitäten, die für die Erreichung von Zielen erforderlich sind (z.B. Dominanz, Lenkung, Kontrolle, Macht). Für die beiden Dimensionen werden je nach Forschungstradition unterschiedliche Bezeichnungen verwendet (z.B. Responsivität und Anforderungen in den Erziehungsstilen nach Baumrind, 1968; Nähe und Einfluss in den Teaching Styles nach Wubbels, Créton, Levy & Hooyamers, 1993; Communion und Agency in Beziehungen in der Interpersonalen Theorie, z.B. Locke, 2011) und meist werden die Achsen bipolar konzipiert (z.B. Kooperation vs. Opposition als Pole der Dimension Nähe und Dominanz vs. Unterordnung als Pole der Dimension Einfluss bei Wubbels et al., 1993). In unserer eigenen Arbeit nutzen wir die Begrifflichkeiten *Communion* und *Agency* und konzipieren die beiden Achsen unipolar (hohe vs. geringe Communion; hohe vs. ge-



**Abbildung 1.** Zirkumplex zur Beschreibung des Verhaltens der Lehrkraft gegenüber einem Schüler oder einer Schülerin mit den beispielhaften Angaben einer Lehrkraft.

**Tabelle 1.** Beschreibung der acht Verhaltensfacetten im Fragebogen zum Lehrkraftverhalten in dyadischen Lehrkraft-Lernenden-Beziehungen mit Beispielimtem (anstelle von X ist der (pseudonymisierte) Name des Kindes einzusetzen)

Verhaltensfacette	Beschreibung	Beispielimtem
Anleitend	In einer wertschätzenden Art gibt die Lehrkraft dem Kind sehr präzise Anleitungen, überwacht und korrigiert ggf. jede seiner Verhaltensweisen.	Wenn ich Aufgaben stelle, überprüfe ich stets, ob X verstanden hat, was er / sie tun muss.
Helfend-lenkend	In einer herzlichen, zugewandten und fürsorglichen Art gibt die Lehrkraft klare Strukturen vor, an denen das Kind sich orientieren kann.	Wenn X eine Aufgabe noch nicht richtig lösen kann, helfe ich ihm / ihr dabei, etwaige Fehlkonzepte zu erkennen und so auf die Lösung zu kommen.
Verständnisvoll-interessiert	In einer herzlichen, zugewandten und fürsorglichen Art schafft die Lehrkraft Gestaltungsspielräume, in denen das Kind seine Interessen realisieren kann.	Ich höre speziell X immer geduldig und interessiert zu, wenn er / sie etwas im Unterricht beiträgt.
Beteiligend-tolerierend	In einer wertschätzenden Art signalisiert die Lehrkraft dem Kind, dass es seinen eigenen Interessen nachgehen kann und lässt es dann einfach gewähren.	Wenn ich Aufgaben gebe, darf X stets solche Aufgaben wählen, bei denen er / sie wirklich etwas Neues lernt.
Unbestimmt-abwartend	Die Lehrkraft befasst sich nicht mit den Interessen oder der Perspektive des Kindes im Besonderen, sondern lässt es einfach gewähren.	Bei meiner Unterrichtsvorbereitung plane ich nicht speziell für X – er / sie wird sich schon auch so in das Unterrichtsgeschehen einfinden.
Ignorierend	Die Lehrkraft hat eine distanzierte, abweisende Haltung gegenüber dem Kind und wendet sich ihm nur dann zu, wenn es sich nicht vermeiden lässt.	Wenn X bei einer herausfordernden Aufgabe stöhnt und jammert, ignoriere ich das erstmal, denn X muss lernen, auftretende Schwierigkeiten selbstständig zu überwinden.
Ermahnend	Die Lehrkraft hat eine distanzierte, abweisende Haltung gegenüber dem Kind und versucht, es durch häufiges Ermahnen in seinem Verhalten zu lenken und zu kontrollieren.	Ich mache X unmissverständlich deutlich, dass die vorgegebenen Aufgaben erledigt werden müssen – ob er / sie möchte oder nicht.
Fordernd-streng	Die Lehrkraft befasst sich nicht mit den Interessen oder der Perspektive des Kindes im Besonderen, sondern gibt sehr präzise Anleitungen, überwacht und korrigiert ggf. jede seiner Verhaltensweisen.	X sage ich oft etwas wie, dass er / sie sich beeilen soll, um eine angefangene Aktivität im Unterricht zu Ende zu bringen.

ringe Agency). Denn diese beiden Begrifflichkeiten stellen die breitest möglichen Bezeichnungen der universalen Inhaltsdimensionen dar, die, oft auch als „Big Two“ bezeichnet, nicht nur für die Beschreibung interpersonaler Beziehungen, sondern – typischerweise als unipolare Dimensionen – auch für die Beschreibung sozialer Wahrnehmungsprozesse, von Stereotypen, Werten, Motiven und Persönlichkeit genutzt werden (z.B. Abele et al., 2016; Fiske, 2018).

Interpersonales Verhalten in dyadischen Interaktionen ist nun stets durch eine *Kombination* von Ausprägungen auf den beiden Dimensionen Communion und Agency charakterisiert. In Zirkumplex-Modellen wird jeder der vier Quadranten, die durch Kreuzung der beiden orthogonalen Achsen entstehen, durch eine Diagonale in zwei Facetten zerteilt, sodass sich insgesamt acht Verhaltensfacetten ergeben (Abbildung 1). Die Dimension, an der eine Facette anliegt, ist in ihr stärker ausgeprägt als in der jeweils anderen Facette desselben Quadranten (z.B. Leary, 1957). Die Angemessenheit eines Zirkumplex-Modells kann durch korrelative Überprüfung der Strukturen empirisch getestet werden: benachbarte Facetten sollten dabei positiv korreliert sein, sich im Kreis gegenüberliegende Facetten negativ und Facetten, die im rechten Winkel zueinanderstehen, unkorreliert sein.

Um das interpersonale Verhalten von Lehrkräften als eine Kombination von Agency und Communion beschreiben zu können, haben sich Wissenschaftler und Wissenschaftlerinnen Zirkumplex-Modelle zu Nutzen gemacht. So wurden in der Forschung von Wubbels (Wubbels et al., 1993; Überblick: Wubbels & Brekelmans, 2005) oder von Aeltermann (z.B. Aeltermann, Vansteenkiste, Haerens, Soenens, Fontaine & Reeve, 2018) sog. Teaching Styles, d.h. Verhaltensstile von Lehrkräften gegenüber ihrer Schulklasse, mit Befragungsdaten der Lehrkraft oder der SuS in Zirkumplex-Modellen identifiziert. Abbildung 1 zeigt ein Zirkumplex-Modell, das wir aufbauend auf den Arbeiten der Forschungsgruppen um Wubbels und um Aeltermann entwickelt und in der vorliegenden Studie für die Beschreibung des Verhaltens der Lehrkraft *gegenüber einem konkreten Kind* genutzt haben. Jede der acht Verhaltensfacetten repräsentiert eine unterschiedliche Kombination von Ausprägungen auf den Dimensionen Agency und Communion. In Tabelle 1 findet sich eine Beschreibung der Verhaltensfacetten mit jeweils einem in unserem Befragungsinstrument verwendeten Item.

In dem in Abbildung 1 gezeigten Beispiel beschreibt die Lehrkraft ihr Verhalten gegenüber einem Kind auf einer Antwortskala von 1 (stimme gar nicht zu) bis 7 (stimme

völlig zu) mit einem mittleren Wert von 6,5 als besonders stark „helfend-lenkend“ (grau markierte Fläche), einer Facette mit hoher Communion und moderater Agency. Die zweitstärkste Ausprägung hat diese Lehrkraft auf der Facette „anleitend“ angegeben (mittlerer Wert von 5,5), also in Bezug auf Verhaltensweisen, die durch eine etwas schwächere Communion und eine etwas stärkere Agency als die der Facette „helfend-lenkend“ charakterisiert sind. Am wenigsten hat diese Lehrkraft den Items der Facetten „ignorierend“ und „unbestimmt-abwartend“ zugestimmt (mittlerer Wert jeweils bei 1). Dies bedeutet, Facetten mit (sehr) schwacher Communion und (sehr) schwacher Agency hält sie für wenig geeignet, um ihr Verhalten gegenüber dem Kind zu beschreiben.

## Für den Schüler oder die Schülerin günstige und ungünstige Verhaltensweisen der Lehrkraft

Empirisch ist gut belegt, dass eine hohe Communion des Lehrkraftverhaltens mit verschiedenen positiven Outcomes auf Seiten der SuS zusammenhängt. Eine Lehrkraft, die sich warm und herzlich verhält, sich dem Kind zuwendet und auf seine Bedürfnisse feinfühlig reagiert, erweist sich für die Entwicklung des Kindes als günstiger als eine Lehrkraft, die Gefühle von Ablehnung und Distanz ausdrückt und nicht auf die Interessen des Kindes eingeht. Dafür sprechen die Ergebnisse der Metaanalysen von Roorda, Koomen, Spilt und Oort (2011) über 99 Studien und von Roorda, Jak, Zee, Oort und Koomen (2017) über 189 Studien, die dyadische Beziehungen zwischen Lehrkraft und Schüler bzw. Schülerin untersucht haben. Sie fanden positive Zusammenhänge zwischen hoher Communion der Lehrkraft und psychologischen Indikatoren schulischen Engagements der SuS sowie – vermittelt über Engagement – ihren Lernergebnissen. Ähnlich fanden Aelterman et al. (2018) für drei Facetten von Lehrkraftverhaltensweisen besonders günstige Folgen für die Motivation der SuS, nämlich die Facetten mit hoher Communion und starker oder mäßig starker Agency (im in Abbildung 1 dargestellten Zirkumplex die Facetten „anleitend“, „helfend-lenkend“ und „verständnisvoll-interessiert“). In dem Maße, wie die SuS ihrer Lehrkraft hohe Ausprägungen auf diesen Facetten attestierten, gaben sie stärker autonome, selbstbestimmte Formen von Motivation und weniger Amotivation an – während sich keine Zusammenhänge zu kontrollierten, fremdbestimmten Motivationsformen zeigten. Als besonders ungünstig erwiesen sich die entgegengesetzten

Facetten, die in dem in Abbildung 1 dargestellten Zirkumplex „fordernd-streng“ und „ermahnend“ heißen und durch geringe Communion charakterisiert sind: in dem Maße, wie die SuS ihrer Lehrkraft hohe Ausprägungen auf diesen Facetten attestierten, gaben sie stärker Formen kontrollierter, fremdbestimmter Motivation und stärkere Amotivation an, während sich keine Zusammenhänge zu autonomen, selbstbestimmten Motivationsformen fanden.

Diese Befunde zeigen übereinstimmend, dass hohe Communion des Lehrkraftverhaltens mit verschiedenen positiven Outcomes auf Seiten der SuS zusammenhängt. Weniger eindeutig ist, welche Ausprägungen des Lehrkraftverhaltens auf der Agency-Dimension besonders günstig für das Kind sind: bei Aelterman et al. (2018) erwiesen sich sowohl sehr starke Ausprägungen („anleitend“), als auch mäßig starke Ausprägungen („verständnisvoll-interessiert“) als günstig. Wir interpretieren dieses Ergebnis als Hinweis darauf, dass in Abhängigkeit eines Merkmals des Kindes unterschiedliche Stärken von Agency besonders vorteilhaft sind: nämlich in Abhängigkeit seiner schulischen Kompetenz. Schulische Kompetenz kann sich dabei sowohl auf fachliche Kompetenzen beziehen, als auch auf die Kompetenz, selbstreguliert zu lernen. Beispielsweise kann schwache Agency des Lehrkraftverhaltens für ein Kind ungünstige Folgen haben, das eigentlich mehr Struktur und Lenkung bräuchte. Bei einem Kind, das in seinen Kompetenzen schon weit fortgeschritten ist, kann schwache Agency des Lehrkraftverhaltens hingegen Autonomieerleben und damit selbstbestimmte Formen von Motivation fördern. Analog kann starke Agency des Lehrkraftverhaltens für ein Kind mit noch relativ wenig entwickelten Kompetenzen lernunterstützend sein, während es bei einem schon sehr kompetenten Kind ungünstige Auswirkungen haben kann, z. B. wenn das Kind dann annimmt, die Lehrkraft traue ihm wenig zu oder es sich in seinen Handlungsmöglichkeiten beschnitten fühlt (vgl. Jang, Reeve & Deci, 2010).

Die Annahme, dass die Agency des Lehrkraftverhaltens in dem Maße günstig für das Kind ist, wie sie zu seiner Kompetenz komplementär ist, ist vereinbar mit dem sog. Komplementaritätsprinzip, das in verschiedenen Modellen interpersonalen Verhaltens postuliert wird, nämlich der Annahme, dass innerhalb von Dyaden das Verhalten der einen Person bei der jeweils anderen Person bestimmte Verhaltensweisen besonders wahrscheinlich macht, sog. komplementäre Verhaltensweisen. Dabei ist interpersonales Verhalten auf der Agency-Dimension komplementär, wenn es *entgegengesetzt* ist – also z. B. noch geringe fachliche Kompetenz oder eine noch geringe Kompetenz zum selbstgesteuerten Lernen eines Kindes mit starker Initiative und Lenkung (Agency)

der Lehrkraft einhergeht und umgekehrt (vgl. z. B. Sadler, Ethier & Woody, 2011; vgl. auch Metaanalyse von Roorda, Zee & Koomen, 2020).

## Ziele und Annahmen der vorliegenden Studie

Zusammenfassend zeigt die zitierte Forschung, dass sich Lehrkräfte in der Communion und Agency ihres Verhaltens gegenüber ihren Klassen voneinander unterscheiden und dass systematische Zusammenhänge zwischen dem Verhaltensstil der Lehrkraft und verschiedenen Outcomes auf Seiten der SuS bestehen. Mit unserer Studie verfolgen wir mehrere Ziele. Erstens wollen wir die *dyadischen Interaktionen* zwischen Lehrkraft und Kind aus der Sicht von Grundschullehrkräften unter Nutzung der Dimensionen Agency und Communion beschreiben lassen. Zweitens wollen wir das Komplementaritätsprinzip prüfen, nämlich die Annahme, dass Lehrkräfte die Agency ihres Verhaltens an der Kompetenz des jeweiligen Kindes ausrichten. Und drittens wollen wir die Frage beantworten, ob mögliche Unterschiede in der Agency des Lehrkraftverhaltens, die sich – als Implikation des Komplementaritätsprinzips – beim Vergleich von *Gruppen* von Kindern mit im Mittel unterschiedlicher Kompetenz zeigen sollten, nach Kontrolle des Einflusses der Kompetenz der Kinder schwächer werden.

In Bezug auf unsere erste Frage, wie Grundschullehrkräfte im Mittel ihr Verhalten auf den Dimensionen Agency und Communion beschreiben, nehmen wir an, dass Lehrkräfte in aller Regel bestrebt sind, ein durch hohe Communion charakterisiertes Verhalten zu zeigen, und zwar gleichermaßen gegenüber allen Kindern ihrer Klasse. Denn das Bedürfnis, mit Kindern zusammen zu sein und zu arbeiten, ist ein zentrales Berufswahlmotiv für das Grundschullehramt (z. B. Retelsdorf & Möller, 2012; Rothland, 2014). Als positiv erlebte Beziehungen zu den SuS verstärken das unterrichtsbezogene Engagement von Lehrkräften (Klassen, Perry & Frenzel, 2012) und reduzieren ihr Belastungserleben (Taxer, Becker-Kurz & Frenzel, 2018). Vor diesem Hintergrund vermuten wir, dass die Grundschullehrkräfte ihr Verhalten als besonders stark auf den beiden Facetten mit sehr hoher Communion („helfend-lenkend“, „verständnisvoll-interessiert“) und umgekehrt als besonders schwach ausgeprägt auf den beiden Facetten mit sehr geringer Communion („ermahnend“, „ignorierend“) beschreiben.

Zur zweiten Fragestellung betreffend das Komplementaritätsprinzip nehmen wir an, dass die Agency der Lehrkraft komplementär zur Kompetenz des Kindes ausgeprägt ist.

Da wir im Rahmen der geplanten Online-Befragung nicht auf Daten in Form standardisierter Kompetenzmessungen zurückgreifen können, haben wir die Lehrkräfte gebeten – als grobe Proxy-Variable für in der Schule relevante Kompetenzen des Kindes – seine aktuellen Noten in den Fächern Deutsch (bzw. Lesen) und Mathematik (bzw. Rechnen) und ggf. in bis zu zwei weiteren Fächern, in dem das Kind von der Auskunft gebenden Lehrkraft selbst unterrichtet wurde, mitzuteilen. Noten und Ergebnisse in standardisierten Leistungstests korrelieren typischerweise nur zu etwa  $r = .55$  (z. B. Bayer & Zinn, 2018; Ingenkamp, 1975; Stubbe, Bos & Schurig, 2017; Überblick Ingenkamp, 1995; Lintorf, 2012). Dies bedeutet, in Noten gehen nicht nur die fachlichen Kompetenzen des Kindes ein, sondern sie reflektieren auch die Beurteilung fachübergreifender Kompetenzen, wie sein Sozial-, Lern- und Arbeitsverhalten (z. B. Anders, McElvany & Baumert, 2010; Han, Elsässer, Lang & Ditton, 2017; Kuhl & Hannover, 2012; vgl. auch die Vorgaben der Kultusministerkonferenz (KMK, 2015), zur notenbasierten Empfehlung in Bezug auf die weiterführenden Schularten). Wenn es zutrifft, dass Lehrkräfte die Stärke der Agency an fachlichen und fachübergreifenden Kompetenzen des jeweiligen Kindes ausrichten, sollte die Agency der Lehrkraft mit den Noten des Kindes negativ korreliert sein. Genauer könnte sich dies darin zeigen, dass die Lehrkräfte ihr eigenes Verhalten auf Facetten mit sehr hoher Agency („fordernd-streng“, „anleitend“) als umso stärker und auf Facetten mit sehr geringer Agency („unbestimmt-abwartend“, „beteiligend-tolerierend“) als umso schwächer ausgeprägt beschreiben, je schlechter die mittlere Note eines Kindes ist.

Betreffend unsere dritte Fragestellung sollte sich, wenn Lehrkräfte ihre Agency an fachlichen und fachübergreifenden Kompetenzen des Kindes ausrichten, die mittlere Agency zwischen Gruppen von Kindern unterscheiden, die im Durchschnitt unterschiedlich starke Kompetenzen haben. Wir betrachten entsprechend die Agency des Lehrkraftverhaltens vergleichend zwischen Kindern mit oder ohne Sonderpädagogischen Förderbedarf (SPF), Kindern mit nichtdeutscher oder deutscher Erstsprache und zwischen Jungen und Mädchen.

Kinder mit SPF und Kinder mit nichtdeutscher Erstsprache bedürfen in der Grundschule im Mittel noch einer stärkeren Förderung ihrer fachlichen Kompetenzen (z. B. Spörer et al., 2015; Wendt, Schwippert & Stubbe, 2016; Wendt & Schwippert, 2017). Dies könnte bedeuten, dass Lehrkräfte ihr Verhalten gegenüber Kindern mit einem SPF und gegenüber Kindern mit nichtdeutscher Erstsprache auf den beiden Facetten mit sehr hoher Agency als stärker und auf den beiden Facetten mit sehr geringer Agency als schwächer ausgeprägt beschreiben als gegenüber Kindern ohne das entsprechende Merkmal. Mädchen zeigen im Mittel stärker entwickelte fachübergreifende

Kompetenzen als Jungen: sie setzen sich eher selbst Ziele und arbeiten zielstrebig an deren Erreichung als Jungen dies tun (z. B. Anders et al., 2010; Duckworth & Seligman, 2006; Kuhl & Hannover, 2012). Dies erklärt auch die besseren Noten, die Mädchen – nach Kontrolle von mit standardisierten Tests erfassten Kompetenzen – relativ zu Jungen erhalten (Review: Spinath, Eckert & Steinmayr, 2014; Metaanalyse: Voyer & Voyer, 2014). Bezogen auf unsere Untersuchungsfrage kann dies bedeuten, dass Mädchen weniger Anleitung und Kontrolle (Agency) durch die Lehrkraft benötigen, um die Lernziele zu erreichen, also unsere Lehrkräfte gegenüber Jungen – in Reaktion auf ihre noch geringere Selbstregulationskompetenz – auf Facetten mit sehr hoher Agency stärkere und auf Facetten mit sehr geringer Agency schwächere Verhaltensaussprägungen angeben als gegenüber Mädchen.

Schließlich betrachten wir, ob die Unterschiede in der Agency der Lehrkraft zwischen den verschiedenen Gruppen von Kindern schwächer werden, wenn für die mit Noten erfassten Kompetenzen der Kinder kontrolliert wird. In ihrer Notengebung reflektieren die Lehrkräfte die fachlichen und fachübergreifenden Kompetenzen eines Kindes (vgl. z. B. Ingenkamp, 1995; Kuhl & Hannover, 2012; Lintorf, 2012). Da die Unterschiede in der Agency der Lehrkraft zwischen Gruppen von Kindern durch deren im Mittel unterschiedlich stark ausgeprägten fachlichen und fachübergreifenden Kompetenzen bedingt sein sollten, erwarten wir, dass sich die Unterschiede abschwächen oder unbedeutsam werden, wenn für die Noten der Kinder kontrolliert wird.

## Methode

### Entwicklung des Fragebogens zum Lehrkraftverhalten in dyadischen Lehrkraft-Lernenden-Beziehungen

#### Vorstudie

Im Rahmen einer Online-Umfrage mit einer Gelegenheitsstichprobe von 16 Lehrkräften haben wir unser Befragungsinstrument erprobt. Die Lehrkräfte sollten ihr Verhalten gegenüber fünf verschiedenen Kindern ihrer Klasse mit Hilfe des Instrumentariums beschreiben ( $N = 80$ ), und zwar auf siebenstufigen Likert-Antwortskalen (1 = stimme gar nicht zu, 2 = stimme überwiegend nicht zu; 3 = stimme eher nicht zu; 4 = stimme teils zu/teils nicht zu; 5 = stimme eher zu; 6 = stimme überwiegend zu; 7 = stimme völlig zu).

Die Items für das Befragungsinstrument haben wir angelehnt an zwei Inventare entwickelt, mittels derer Inter-

aktionen zwischen Lehrkraft und SuS auf Grundlage eines Zirkumplex-Modells beschrieben werden können: dem *Questionnaire on Teacher Interaction* (QTI, Wubbels et al., 1993) und dem *Situations-in-School Questionnaire* (SiS, Aelterman et al., 2018). Der QTI ist in seiner Originalfassung von 1993 so konzipiert, dass SuS mit ihm das Interaktionsverhalten der Lehrkraft auf acht Facetten – ähnlich denen von uns in Tabelle 1 dargestellten – einschätzen können. Im SiS beschreibt die Lehrkraft ihr eigenes Verhalten gegenüber der Klasse auf acht Facetten – ähnlich denen in Tabelle 1 dargestellten.

Unser Messinstrument sollte von Lehrkräften für die Beschreibung ihres Verhaltens in dyadischen Interaktionen mit einzelnen Kindern ihrer Klasse genutzt werden. Da die Lehrkräfte insgesamt fünf Kinder zu beurteilen hatten, war das Ziel, die acht Facetten mit möglichst wenigen Items zu erfassen, damit die Befragung nicht zu viel Zeit in Anspruch nehmen und die Compliance der Lehrkräfte hoch bleiben würde. Wir haben deshalb angestrebt, jede der acht Facetten im finalen Instrument durch jeweils nur zwei Items zu repräsentieren. Die Formulierungen der Items aus dem QTI und dem SiS haben wir als Anregungen genutzt, um für jede der acht Facetten zwei deutschsprachige Items zu formulieren, mittels derer die Lehrkraft ihr Verhalten gegenüber einem bestimmten Kind beschreiben kann. Inhaltlich sollten die Items die Steuerung des Lernverhaltens (und nicht des Sozialverhaltens) der Kinder durch die Lehrkraft beschreiben. Jedes einzelne Item sollte jeweils eine bestimmte Kombination von starker, moderater oder geringer Agency bei gleichzeitig starker, moderater oder geringer Communion beinhalten. Mit den einzelnen Items wird also erfasst, wie stark lenkend das Verhalten der Lehrkraft dem Kind gegenüber ausgeprägt ist (Agency), wobei der Grad an Zugewandtheit, die die Lehrkraft dem Kind entgegenbringt (Communion), zwischen den Items gleichzeitig variiert. Die Bezeichnungen der acht Facetten haben wir gegenüber dem QTI und SiS ggf. ebenfalls leicht adaptiert, damit sie optimal zur Beschreibung dyadischer Beziehungen passen. Bei einigen Facetten hatten wir in unserer Vorstudie zunächst drei Items berücksichtigt, nämlich für die Facetten „anleitend“, „helfend-lenkend“, „verständnisvoll-interessiert“ und „ermahnend“, in der Hoffnung, diese Item-Zahlen auf empirischer Basis weiter reduzieren zu können. Für die verbleibenden Facetten „beteiligend-tolerierend“, „unbestimmt-abwartend“, „ignorierend“ und „fordernd-streng“ wurden bereits in der Vorstudie jeweils nur zwei Items berücksichtigt. Somit bestand das Instrument für die Vorstudie aus insgesamt 20 Items.

Die Ergebnisse zeigten erste Hinweise auf die Zirkumplex-Struktur der Daten in Form von Korrelationen zwischen den Facetten. Mit dem Ziel der Erhöhung der Reliabilität einzelner Subskalen wurden die Items teil-

weise sprachlich überarbeitet. Der entsprechend leicht modifizierte Satz von 20 Items wurde dann den an unserer Hauptuntersuchung teilnehmenden Lehrkräften mit denselben siebenstufigen Likert-Skalen wie in der Vorstudie vorgelegt.

### Stichprobenbeschreibung und Ablauf der Hauptuntersuchung

Über verschiedene Internetforen und soziale Netzwerke wurden 83 Lehrerinnen und Lehrer für die Teilnahme an einer Online-Befragung gewonnen. Da wir Lehrkraftverhalten gegenüber Kindern im Grundschulalter untersuchen wollten, mussten wir 11 Lehrkräfte, die sich in ihren Angaben auf Lernende über 12 Jahre bezogen haben, aus unseren weiteren Analysen ausschließen. Von den verbleibenden 72 Lehrkräften liegen uns Verhaltensbeschreibungen zu insgesamt 302 Kindern im Grundschulalter vor. Die Lehrkräfte erhielten 10 Euro als Aufwandsentschädigung, wenn sie den Fragebogen bis zum Ende ausgefüllt hatten und auch Angaben zu ihren demographischen Daten machten. Dies war bei 58 der 72 Befragten der Fall. In dieser Gruppe betrug das mittlere Alter  $M = 38.9$  Jahre ( $SD = 11.0$ ; Min/Max = 22–64 Jahre). Die 58 Lehrkräfte waren im Mittel bereits 11.2 Jahre ( $SD = 9.6$ ; Min/Max = 1–39 Jahre) als Lehrkraft tätig (siehe Tabelle 2 für eine ausführlichere Beschreibung).

Im Online-Fragebogen wurde die Lehrkraft zunächst gebeten, fünf Kinder ihrer Klasse auszuwählen, die sich in verschiedenen Merkmalen deutlich voneinander unterscheiden. Sie sollte dann ihr eigenes Verhalten gegenüber jedem dieser fünf Kinder nacheinander auf den jeweils 20 Items des Befragungsinstruments beschreiben. Abschließend wurde die Lehrkraft gebeten anzugeben, ob das Kind männlich oder weiblich ist (Jungen = 0, Mädchen = 1), ob Deutsch oder eine andere Sprache seine Erst-

**Tabelle 2.** Stichprobenbeschreibung der Lehrkräfte, die persönliche Angaben gemacht hatten

	<i>n</i>	%
<i>Geschlecht</i>		
männlich	7	12,1
weiblich	51	87,9
divers	1	1,7
<i>Erstsprache</i>		
deutsch	57	98,3
andere	1	1,7
<i>Bundesland</i>		
Bayern	7	12,1
Berlin	20	34,5
Brandenburg	4	6,9
Mecklenburg-Vorpommern	5	8,6
Niedersachsen	5	8,6
Nordrhein-Westfalen	12	20,7
Rheinland-Pfalz	1	1,7
Schleswig-Holstein	3	5,2
Thüringen	1	1,7
<i>Zusatzqualifikation zum Thema Sonderpädagogik/Inklusion</i>		
ja	7	12,1
nein	51	87,9

Anmerkungen:  $n = 58$ . Wir haben insgesamt 72 Lehrkräfte befragt. Von 14 Lehrkräften liegen uns keine sozio-demographischen Daten vor.

sprache ist (Deutsch = 0, andere Sprache = 1) und ob es einen SPF hat (kein Förderbedarf = 0, Förderbedarf = 1). Schließend sollte die Lehrkraft die zurückliegenden Zeug-

**Tabelle 3.** Verteilung der Untersuchungsmerkmale in der Gesamtstichprobe der Kinder, den Teilstichproben von Kindern mit und ohne Angaben zu Schulnoten und Ergebnisse eines die Verteilung der Merkmale in den Teilstichproben vergleichenden  $\chi^2$ -Tests

	Gesamtstichprobe ( $N = 302$ )		Teilstichprobe ohne Angaben zu Schulnoten ( $N = 132$ )		Teilstichprobe mit Angaben zu Schulnoten ( $N = 170$ )		$\chi^2$	<i>p</i>
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%		
<i>Geschlecht</i>								
Jungen	173	57	79	60,3	94	55,3	0,76	,383
Mädchen	128	42	52	39,7	76	44,7		
<i>Erstsprache</i>								
Deutsch	246	81	111	84,1	135	79,4	1,08	,299
Nichtdeutsch	56	19	21	15,9	35	20,6		
<i>SPF</i>								
Nein	272	90	119	90,2	153	90,0	0,00	,965
Ja	30	10	13	9,8	17	10,0		

nisnoten des Kindes in Deutsch und Mathe und ggf. Noten in bis zu zwei weiteren Fächern angeben, in dem die Lehrkraft selbst das Kind unterrichtet.

Zum Zeitpunkt der Befragung unterrichteten die Lehrkräfte die Kinder, denen gegenüber sie ihr Verhalten einschätzten, im Mittel seit zwei Jahren ( $SD = 0.97$ ). Von den 72 Lehrkräften hatten 54 ihr Verhalten entsprechend unserer Instruktion jeweils gegenüber fünf Kindern aus ihrer Klasse beschrieben, 18 Lehrkräfte gaben ihr Verhalten gegenüber weniger als fünf Kindern an. Unsere Analysen beziehen sich somit auf eine Gesamtstichprobe von 302 Kindern. Nur für 170 Kinder hatten die Lehrkräfte Angaben zu ihren Noten gemacht. Dabei haben die Lehrkräfte entweder zu allen Kindern oder zu keinem der Kinder, zu denen sie ihr Verhalten beschrieben hatten, Noten angegeben, so dass wir vermuten, dass ein Nichteintrag darauf zurückzuführen war, dass der Lehrkraft die Noten der Kinder zum Zeitpunkt der Online-Befragung nicht vorlagen. Tabelle 3 zeigt die Verteilung der Merkmale Geschlecht, Erstsprache und SPF in der Gesamtstichprobe von 302 Kinder und in den Teilstichproben von Kindern, für die Noten vorlagen ( $n = 170$ ) und nicht vorlagen ( $n = 132$ ). Um sicherzustellen, dass die fehlenden Daten nicht systematisch mit den Gruppenmerkmalen der Kinder zusammenhängen, wurden  $\chi^2$ -Tests berechnet. Die Ergebnisse zeigen, dass die fehlenden Angaben zufällig über die Kinder verteilt waren.

## Ergebnisse

### Skalenbildung und Überprüfung der Zirkumplex-Struktur

Ziel war es, ein für Lehrkräfte handhabbares und ökonomisches Messinstrument (d.h. mit möglichst wenig Items) zur Erfassung ihres Verhaltens in dyadischen Lehrkraft-Lernenden-Beziehungen zu entwickeln. Bei der Skalenbildung strebten wir für jede der acht Facetten des Lehrkraftverhaltens an, zwei Items zu identifizieren, durch die der gesamte durch die orthogonalen Dimensionen Communion und Agency aufgespannte zweidimensionale Raum repräsentiert wird. Bei der Itemauswahl waren also nicht nur Eigenschaften des einzelnen Items zu berücksichtigen, sondern auch, wie gut sie die jeweilige Facette in Abgrenzung von oder relativ zu den jeweils anderen Facetten des Zirkumplex repräsentieren.

In einem ersten Schritt wurde die Methode der multidimensionalen Skalierung (Bortz & Döring, 2016) dazu

genutzt, die Items in dem zweidimensionalen Raum graphisch anzuordnen (Scatterplot)<sup>1</sup>. In dem Fall, dass eine Facette durch mehr als zwei Items gemessen worden war („anleitend“, „helfend-lenkend“, „verständnisvoll-interessiert“, „ermahnend“), wurde von zwei Items, die sehr nah beieinander lagen, eines eliminiert. Ebenfalls wurden Items eliminiert, die zu stark im Zentrum des Zirkumplex lagen und somit nicht gut zwischen den verschiedenen Facetten differenzieren. Dies hatte zur Folge, dass bei der Facette „verständnisvoll-interessiert“ zwei Items entfernt werden mussten, die vermutlich aufgrund eines ähnlichen Wortlauts extrem hoch interkorreliert waren, jedoch gegenüber dem dritten (als einzigem beibehaltenem) Item den semantischen Inhalt der Facette weniger gut repräsentierten (siehe Tabelle 1).

In einem zweiten Schritt prüften wir die Reliabilität der verbleibenden 2-Item-Skalen mittels der Spearman-Brown Formel, wobei bei fünf Skalen akzeptable bis gute Werte erzielt wurden (siehe Tabelle 4). Nicht akzeptabel waren die Spearman-Brown-Koeffizienten für die Facetten „ignorierend“ bzw. „beteiligend-tolerierend“ ( $\rho = .29$  bzw.  $\rho = .54$ ), sodass aus diesen beiden Skalen in den weiteren Analysen nur jeweils ein Item berücksichtigt wurde. Die Itemauswahl erfolgte dabei erneut auf Grundlage der Anordnung des Items im Scatterplot und inhaltlicher Erwägungen. Der finale *Fragebogen zum Lehrkraftverhalten in dyadischen Lehrkraft-Lernenden-Beziehungen* besteht folglich aus insgesamt 13 Items, auf denen Lehrkräfte ihr Verhalten gegenüber einem einzelnen Kind beschreiben können. Diese Verhaltensbeschreibungen lassen sich insgesamt acht Facetten zuordnen, in denen das Ausmaß an Agency und Communion systematisch variiert.

Die in Tabelle 4 dargestellten Interkorrelationen zwischen den Verhaltensfacetten sind mit den Annahmen eines Zirkumplex-Modells konsistent, wenn auch nicht in jedem einzelnen Fall: Benachbarte Facetten waren positiv korreliert, gegenüberliegende Facetten negativ korreliert bzw. unkorreliert und Facetten im rechten Winkel waren unabhängig bzw. leicht positiv korreliert. In Tabelle 4 sind die Ladungen auf den Faktoren 1 und 2 angegeben, die sich aus einer explorativen 2-faktoriellen Maximum Likelihood-Faktoranalyse mit orthogonaler Rotation (Varimax) über die 8 Facetten ergab. Die Ergebnisse zeigen erwartungsgemäß, dass auf Faktor 1 Facetten mit hoher Agency deutlich positiv laden (z.B. „fordernd-streng“, Ladung: .93) bzw. Facetten mit geringer Agency deutlich negativ laden (z.B. „unbestimmt-abwartend“, Ladung: -.42). Facetten mit mittlerer Agency (z.B. „helfend-lenkend“: .25; „verständnisvoll-interessiert“: -.17) laden

<sup>1</sup> Items, die empirisch in einer anderen Facette verortet gewesen wären als theoretisch für sie angenommen worden war, wären als erste entfernt worden – dieser Fall kam jedoch (wie gewünscht) nicht vor.

**Tabelle 4.** Mittelwerte, Standardabweichungen, Reliabilitäten (Spearman-Brown-Koeffizient  $\rho$ ), Pearson Korrelationen zwischen Lehrkraftverhalten und Schulnoten sowie Ergebnisse (Faktorladungen) einer explorativen Faktorenanalyse mit orthogonaler Rotation über die acht Verhaltensfacetten

	M	SD	$\rho$	Faktor									
				1	2	3	4	5	6	7	8		
1 Anleitend	4.04	1.85	.86	–								.62	.25
2 Helfend-lenkend	5.41	1.22	.68	.39***	–							.25	.92
3 Verständnisvoll-interessiert <sup>a</sup>	5.38	1.44	–	.00	.48***	–						–.17	.60
4 Beteiligtend-tolerierend <sup>a</sup>	4.54	1.37	–	–.25***	.19***	.42***	–					–.27	.28
5 Unbestimmt-abwartend	4.22	1.63	.71	–.58***	–.08	.20***	.24***	–				–.42	.07
6 Ignorierend <sup>a</sup>	3.86	1.80	–	.06	.01	.02	.08	.13*	–			.27	–.03
7 Ermahnend	3.44	1.97	.83	.42***	.07	–.20**	–.17**	–.27***	.36***	–		.79	–.19
8 Fordernd-streng	3.84	1.97	.85	.57***	.19**	–.19**	–.27***	–.38***	.28***	.78***	–	.93	–.07
9 Schulnoten	2.28	0.95	.92	.65***	.12	–.21**	–.46***	–.59***	.03	.40***	.54***		

Anmerkungen: Lehrkraftverhalten von 1 (stimme gar nicht zu) bis 7 (stimme völlig zu); Schulnoten von 1 (sehr gut) bis 6 (ungenügend);  $N = 302$  ohne Schulnoten,  $n = 170$  mit Schulnoten.

<sup>a</sup> Einzelitem. Keine Berechnung des Spearman-Brown-Koeffizienten möglich.

†  $p < .10$ , \*  $p < .05$ , \*\*  $p < .01$ , \*\*\*  $p < .001$ .

schwach bzw. um Null herum auf Faktor 1. Gleichzeitig laden die Facetten in dem Maße auf Faktor 2 positiv oder negativ, wie stark die Communion dem Zirkumplex-Modell zufolge in ihnen ausgeprägt sein sollte (z. B. „helfend-lenkend“, Ladung: .92; „ermahnend“, Ladung: –.17) – wobei allerdings die Ladungen für Facetten mit geringer Communion insgesamt auf Faktor 2 nur schwach negativ ausgeprägt waren.

Die Struktur des Modells wurde zusätzlich mit der Methode der strukturellen Zusammenfassung für Zirkumplex-Daten untersucht (Structural Summary Method (SSM); Gurtman & Pincus, 2003; Wright, Pincus, Conroy, & Hilsenroth, 2009). Hierfür wurde das *circumplex*-Paket (Girard, Zimmermann & Wright, 2020) für das Statistikprogramm R (R Core Team, 2020) genutzt. Die Passung der theoretisch angenommenen Zirkumplex-Struktur zu den Daten zeigte mit  $R^2 = .863$  einen guten Modellfit. Ein Vergleich der Modellgüte bestätigte, dass die Einzelitems die drei Facetten „ignorierend“, „verständnisvoll-interessiert“ und „beteiligtend-tolerierend“ besser repräsentieren als die Zwei-Item-Lösungen, bei denen der globale Modellfit jeweils nur unzureichend war ( $R^2 = .471 - .634$ ).

## Ausprägungen der Lehrkraftverhaltensweisen auf den acht Zirkumplex-Facetten

Wir hatten angenommen, dass die Lehrkräfte ihr Verhalten gegenüber allen Kindern im Mittel als besonders hoch ausgeprägt auf den beiden Facetten mit sehr star-

ker Communion („helfend-lenkend“, „verständnisvoll-interessiert“) und als besonders niedrig ausgeprägt auf den beiden Facetten mit sehr schwacher Communion („ermahnend“, „ignorierend“) beschreiben würden. Zur Prüfung dieser Annahme haben wir t-Tests für gepaarte Stichproben berechnet, bei denen wir jeweils den Mittelwert einer der beiden Facetten mit sehr starker bzw. sehr niedriger Communion mit den Mittelwerten aller anderen Facetten verglichen haben. Wie die in Tabelle 4 dargestellten mittleren Werte zeigen, bestätigten sich unsere Erwartungen im Wesentlichen. Sowohl die Facette „helfend-lenkend“, als auch die Facette „verständnisvoll-interessiert“ waren im Mittel gegenüber allen anderen sechs Verhaltensfacetten signifikant stärker ausgeprägt (alle  $ps < .001$ ,  $d = .53 - .88$ ). Erwartungskonform zeigte sich die geringste Ausprägung auf der Facette „ermahnend“ – und zwar im Vergleich zu allen anderen Verhaltensfacetten im Zirkumplex (alle  $ps < .001$ ,  $d = -.88 - .20$ ). Die Zustimmung auf der Facette „ignorierend“ war hingegen vergleichbar niedrig wie auf Facetten mit mäßiger Communion („anleitend“  $p = .219$ ,  $d = -.07$ ; „fordernd-streng“  $p = .879$ ,  $d = .01$ ).

## Komplementarität des Lehrkraftverhaltens zur Kompetenz des Kindes

Unsere zweite Untersuchungsfrage war, ob mit zunehmend besserer mittlerer Note eines Kindes die Agency des Lehrkraftverhaltens abnehmen würde. Wie in der letzten Zeile von Tabelle 4 dargestellt ist, bestätigte sich diese

**Tabelle 5.** Ergebnisse der Regressionsanalysen zur Vorhersage des Lehrkraftverhaltens durch die Merkmale Geschlecht, Erstsprache und Sonderpädagogischer Förderbedarf (SPF)

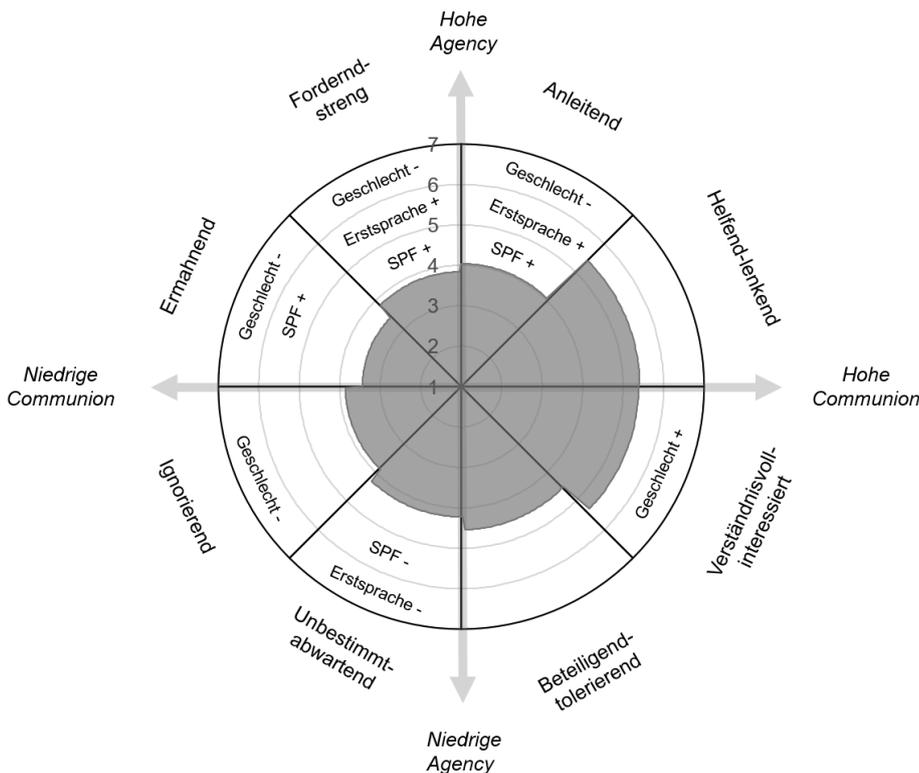
	Anleitend				Helfend-lenkend				Verständnisvoll-interessiert				Beteiligend-tolerierend			
	B	SE(B)	β	p	B	SE(B)	β	p	B	SE(B)	β	p	B	SE(B)	β	p
Konstante	3.91	0.14		<.001	5.19	0.10		<.001	5.19	0.12		<.001	4.49	0.11		<.001
Geschlecht	-0.61	0.20	-.16	.002	0.20	0.14	.08	.151	0.41	0.17	.14	.014	0.37	0.16	.14	.018
Erstsprache	1.19	0.25	.25	<.001	0.25	0.18	.08	.166	-0.01	0.21	.00	.955	-0.33	0.20	-.10	.095
SPF	1.70	0.33	.28	<.001	0.84	0.23	.21	<.001 <sup>a</sup>	0.34	0.27	.07	.218	-0.32	0.26	-.07	.216 <sup>a</sup>
R <sup>2</sup>	.18				.05				.02				.04			

	Unbestimmt-abwartend				Ignorierend				Ermahnend				Fordernd-streng			
	B	SE(B)	β	p	B	SE(B)	β	p	B	SE(B)	β	p	B	SE(B)	β	p
Konstante	4.36	0.13		<.001	4.07	0.15		<.001	3.78	0.15		<.001	4.26	0.15		<.001
Geschlecht	0.43	0.18	.13	.016 <sup>a</sup>	-.52	0.21	-.14	.014	-1.29	0.21	-.32	<.001	-1.56	0.20	-.39	<.001
Erstsprache	-0.85	0.22	-.20	<.001	-.02	0.26	.00	.947	0.38	0.27	.07	.161	0.67	0.26	.13	.011
SPF	-1.56	0.29	-.29	<.001	.28	0.35	.05	.427	1.30	0.35	.20	<.001	1.12	0.34	.17	.001
R <sup>2</sup>	.15				.02				.17				.22			

Anmerkungen: Kodierung Geschlecht: 0 (männlich) bzw. 1 (weiblich); Kodierung Erstsprache: 0 (deutsch) bzw. 1 (nichtdeutsch); Kodierung SPF: 0 (nein) bzw. 1 (ja); N = 302.

<sup>a</sup> Dieser Effekt stellte sich in der Teilstichprobe mit Kindern, für die Noten vorlagen, anders dar und wird deshalb nicht interpretiert.



**Abbildung 2.** Mittlere Ausprägung des Lehrkraftverhaltens gegenüber allen Kindern auf den acht Facetten des Zirkumplex unter Kontrolle aller Merkmale (ausgenommen Schulnoten).

Anmerkung: Eingezeichnet sind sämtliche signifikant unterschiedliche Ausprägungen für Gruppen von Kindern und die Richtung des Effekts. Kodierung: Geschlecht (0 = männlich, 1 = weiblich); Erstsprache (0 = deutsch, 1 = nichtdeutsch); SPF (0 = nein, 1 = ja).

Vermutung für die 170 Kinder, für die Noten vorlagen, in deutlichen Pearson-Korrelationen: Je besser die mittlere Note eines Kindes war, als umso schwächer ausgeprägt beschrieb die Lehrkraft ihr Verhalten auf den Facetten mit sehr hoher Agency, „fordernd-streng“ ( $r = .54$ ) und „anleitend“ ( $r = .65$ ), und als umso stärker ausgeprägt beschrieb sie ihr Verhalten auf den Facetten mit sehr geringer Agency, „beteiligtend-tolerierend“ ( $r = -.46$ ) und „unbestimmt-abwartend“ ( $r = -.59$ ).

## Unterschiede im Verhalten der Lehrkräfte gegenüber verschiedenen Gruppen von Kindern

Im nächsten Schritt haben wir geprüft, ob die Lehrkräfte ihr Verhalten gegenüber Kindern mit einem SPF, mit nicht-deutscher Erstsprache und gegenüber Jungen auf den beiden Facetten mit sehr hoher Agency („fordernd-streng“, „anleitend“) als stärker und auf den beiden Facetten mit sehr geringer Agency („unbestimmt-abwartend“, „beteiligtend-tolerierend“) als schwächer ausgeprägt beschreiben würden als gegenüber Kindern der jeweiligen Vergleichsgruppe ohne das entsprechende Merkmal. Dazu haben wir für die Gesamtstichprobe der 302 Kinder für jede der acht Verhaltensfacetten eine Regression berechnet, in der die Ausprägung des Lehrkraftverhaltens simultan aus einem möglichen SPF des Kindes, aus seiner Erstsprache und seinem Geschlecht vorhergesagt wurde (Tabelle 5). Um die Stabilität der Effekte innerhalb der Stichprobe zu prüfen und um – in einer weiteren Analyse – für den Einfluss der Noten kontrollieren zu können, haben wir dieselben Regressionsanalysen auch für die Teilstichprobe der 170 Kinder berechnet, für die Noten vorlagen (Tabelle E1 in Elektronischem Supplement [ESM] 1). Ein Vergleich der Ergebnisse (Tabelle 5 und Tabelle E1 in ESM 1) zeigte, dass von drei Ausnahmen abgesehen, sämtliche Effekte analog waren, d. h. die Beta-Gewichte das gleiche Vorzeichen und eine vergleichbare Stärke hatten (in manchen Fällen ergaben sich bei vergleichbaren Beta-Gewichten in der Teilstichprobe keine Signifikanzen mehr).<sup>2</sup> Im Folgenden werden nur die Effekte dargestellt, die sich stabil in der Gesamtstichprobe und der Teilstichprobe der Kinder mit Notenangaben gezeigt hatten (Abbildungen 2 und 3). Kindern mit nichtdeutscher Erstsprache gegenüber zeigten die Lehrkräfte erwartungsgemäß stärker „fordernd-stren-

ges“ ( $\beta = .13, p = .011$ ) und „anleitendes“ ( $\beta = .25, p < .001$ ) und weniger „unbestimmt-abwartendes“ ( $\beta = -.20, p < .001$ ) Verhalten als gegenüber Kindern mit deutscher Erstsprache. Lediglich auf der Facette „beteiligtend-tolerierend“ zeigte sich der erwartete Unterschied zwischen Kindern mit unterschiedlicher Erstsprache nicht ( $\beta = -.10, n.s.$ ). Beim Vergleich von Kindern mit oder ohne SPF zeigten sich Unterschiede in erwarteter Richtung für die Facetten „fordernd-streng“ ( $\beta = .17, p = .001$ ), „anleitend“ ( $\beta = .28, p < .001$ ) und „unbestimmt-abwartend“ ( $\beta = -.29, p < .001$ ), während sich für die Facette „beteiligtend-tolerierend“ der erwartete Unterschied nur in der Teilstichprobe der 170 Kinder ergab, für die Noten vorlagen ( $\beta = -.19, p = .015$ ; Gesamtstichprobe  $\beta = -.07, n.s.$ ). Darüber hinaus gaben die Lehrkräfte für Kinder mit SPF an, sich ihnen gegenüber signifikant stärker „ermahnend“ ( $\beta = .20, p < .001$ ) zu verhalten als gegenüber Kindern ohne SPF. Dies bedeutet, die Lehrkräfte beschrieben ihr Verhalten auf drei der vier Facetten mit relativ hoher Agency als signifikant stärker ausgeprägt als gegenüber Kindern ohne Förderbedarf. Zusammengefasst zeigten die Lehrkräfte nach ihren eigenen Angaben gegenüber Kindern mit nichtdeutscher Erstsprache und gegenüber Kindern mit SPF stärker agentisches Verhalten, und zwar sowohl in Kombination mit hoher als auch mit geringer Communion, als gegenüber Kindern ohne das entsprechende Merkmal.

Der Vergleich von Mädchen und Jungen ergab, dass die Lehrkräfte im Umgang mit Jungen stärkere Ausprägungen auf den Facetten mit sehr hoher Agency („anleitend“  $\beta = -.16, p = .002$ ; „fordernd-streng“  $\beta = -.39, p < .001$ ) und schwächere Ausprägungen auf der Facette mit sehr geringer Agency („beteiligtend-tolerierend“  $\beta = .14, p = .018$ ), zeigten als gegenüber Mädchen; für die Facette mit sehr geringer Agency „unbestimmt-abwartend“ war der erwartete Unterschied nur in der Gesamtstichprobe signifikant ( $\beta = .13, p = .016$ , Teilstichprobe  $\beta = .06, n.s.$ ). Darüber hinausgehend gaben sie aber auch auf den beiden Facetten mit der geringst möglichen Ausprägung von Communion (und moderat starker Agency), „ignorierend“ ( $\beta = -.14, p = .014$ ) und „ermahnend“ ( $\beta = -.32, p < .001$ ), gegenüber Jungen stärkere Verhaltensausprägungen als gegenüber Mädchen an und gleichzeitig auf der Facette mit der höchst möglichen Ausprägung von Communion (und moderat starker Agency), „verständnisvoll-interessiert“ ( $\beta = .14, p = .014$ ), eine stärkere Ausprägung gegenüber Mädchen als gegenüber Jungen an.

<sup>2</sup> Die drei bedeutsamen Unterschiede in den Betagewichten (und Signifikanzen) waren die Folgenden: (a) der in der Gesamtstichprobe auftretende Geschlechtseffekt, nach dem Jungen stärker unbestimmt-abwartend behandelt werden als Mädchen, trat in der Teilstichprobe der Kinder mit Notenangaben nicht auf; (b) der in der Gesamtstichprobe auftretende Effekt für den SPF, nach dem Kinder mit SPF stärker helfend-lenkend behandelt werden als Kinder ohne SPF, trat in der Teilstichprobe nicht auf, und (c) während sich in der Gesamtstichprobe kein Effekt für den SPF auf der Facette „beteiligtend-tolerierend“ zeigte, beschrieben die Lehrkräfte ihr Verhalten gegenüber Kindern mit SPF in der Teilstichprobe als stärker „beteiligtend-tolerierend“ als gegenüber Kindern ohne SPF. Diese drei nicht mehr interpretierten Effekte haben wir in Tabelle 5 mit einem Superscript markiert.

**Tabelle 6.** Ergebnisse der Regressionsanalysen zur Vorhersage des Lehrkraftverhaltens durch die Merkmale Geschlecht, Erstsprache, Sonderpädagogischer Förderbedarf (SPF) und Schulnoten

	Anleitend				Helfend-lenkend				Verständnisvoll-interessiert				Beteiligtend-tolerierend			
	B	SE(B)	$\beta$	p	B	SE(B)	$\beta$	p	B	SE(B)	$\beta$	p	B	SE(B)	$\beta$	p
Konstante	1.40	0.33		<.001	5.07	0.26		.000	6.14	0.31		<.001	6.07	.29		<.001
Geschlecht	-0.09	0.22	-.02	.682	0.13	0.17	.06	.463	0.17	0.21	.07	.402	0.11	0.20	.04	.566
Erstsprache	0.65	0.27	.14	.016	0.24	0.21	.09	.253	0.24	0.25	.07	.339	-0.09	0.24	-.03	.693
SPF	0.35	0.38	.06	.350	0.19	0.30	.05	.527	0.47	0.35	.11	.188	-0.22	0.34	-.05	.515
Schulnoten	1.13	0.12	.59	<.001	0.12	0.10	.10	.246	-0.33	0.12	-.24	.005	-0.62	0.11	-.43	<.001
$R^2$	.45				.03				.06				.22			

	Unbestimmt-abwartend				Ignorierend				Ermahnend				Fordernd-streng			
	B	SE(B)	$\beta$	p	B	SE(B)	$\beta$	p	B	SE(B)	$\beta$	p	B	SE(B)	$\beta$	p
Konstante	6.61	0.29		<.001	4.42	0.42		<.001	2.70	0.42	-.33	<.001	2.45	0.37		<.001
Geschlecht	-0.18	0.20	-.06	.378	-0.74	0.28	-.21	.009	-1.37	0.28	.01	<.001	-1.30	0.25	-.33	<.001
Erstsprache	-0.30	0.24	-.08	.213	0.16	0.34	.04	.634	0.05	0.34	.06	.885	0.10	0.30	.02	.731
SPF	-0.78	0.34	-.15	.024	-0.37	0.48	-.06	.440	0.40	0.49	.29	.409	0.69	0.42	.11	.103
Schulnoten	-0.87	0.11	-.53	<.001	-0.02	0.16	-.01	.906	0.62	0.16		<.001	0.84	0.14	.41	<.001
$R^2$	.38				.05				.27				.40			

Anmerkungen: Kodierung Geschlecht: 0 (männlich) bzw. 1 (weiblich); Kodierung Erstsprache 0 (deutsch) bzw. 1 (nichtdeutsch); Kodierung SPF: 0 (nein) bzw. 1 (ja); Kodierung Schulnoten von 1 (sehr gut) bis 6 (ungenügend); n = 170.

Insgesamt beschrieben die Lehrkräfte damit ihr Verhalten gegenüber Schülern auf sechs der acht Facetten als anders als ihr Verhalten gegenüber Schülerinnen (vgl. Tabelle 5, Abbildung 2).

### Unterschiede im Verhalten der Lehrkräfte gegenüber verschiedenen Gruppen von Kindern nach Kontrolle der Noten der Kinder

Unsere dritte Fragestellung war, ob Unterschiede in der Agency des Lehrkraftverhaltens zwischen Gruppen von Kindern mit oder ohne SPF, mit deutscher oder nichtdeutscher Erstsprache und zwischen Mädchen und Jungen durch Unterschiede in fachlichen oder fachübergreifenden Kompetenzen, die in den Noten der Kinder reflektiert sind, begründet sein würden, so dass sie sich abschwächen sollten, wenn für die Noten der Kinder kontrolliert wird.

Zur Prüfung dieser Annahme wurden dieselben wie in Tabelle 5 berichteten Regressionsanalysen unter Einbezug der mittleren Note der Kinder wiederholt berechnet, wobei aufgrund fehlender Angaben der Lehrkräfte hierbei nur 170 Kinder in den Analysen berücksichtigt werden konnten (Tabelle 6). Um zu gewährleisten, dass etwaige Unterschiede tatsächlich auf die Kontrolle der Noten und

nicht auf zufallsbedingte Unterschiede zwischen der Gesamtstichprobe und der Teilstichprobe der 170 Kinder zurückzuführen sind, werden wiederum nur die Effekte interpretiert, die sich in der Gesamtstichprobe und der Teilstichprobe stabil gezeigt hatten. Wie bereits in den bivariaten Korrelationen zu erkennen war, zeigten sich auf sechs von acht Facetten systematische Zusammenhänge zwischen den Noten der Kinder und der Ausprägung des Lehrkraftverhaltens in der Form, dass bessere Noten des Kindes Lehrkraftverhalten mit geringer Agency wahrscheinlich machten („verständnisvoll-interessiert“  $\beta = -.24$ ,  $p = .005$ ; „beteiligtend-tolerierend“  $\beta = -.43$ ,  $p < .001$ ; „unbestimmt-abwartend“  $\beta = -.53$ ,  $p < .001$ ) und schlechtere Noten des Kindes Lehrkraftverhalten mit relativ hoher Agency („anleitend“  $\beta = .59$ ,  $p < .001$ ; „ermahnend“  $\beta = .29$ ,  $p < .001$ ; „fordernd-streng“  $\beta = .41$ ,  $p < .001$ ; Tabelle 6) hervorriefen.

Wie Abbildung 3 zeigt, wurden von den auf drei Facetten gefundenen Unterschieden im Lehrkraftverhalten gegenüber Kindern mit deutscher oder nichtdeutscher Erstsprache zwei durch die Noten der Kinder erklärt („unbestimmt-abwartend“, „fordernd-streng“); nur der Effekt auf der Facette „anleitend“ blieb bestehen, demnach die Lehrkräfte auch bei gleichen Noten gegenüber Kindern mit nichtdeutscher Erstsprache eine stärkere

Ausprägung auf dieser Facette mit sehr starker Agency zeigten als gegenüber Kindern mit deutscher Erstsprache ( $\beta = .14, p = .016$ ). Für Kinder mit oder ohne SPF war auf vier der fünf Facetten, für die sich ein Unterschied im Verhalten der Lehrkräfte gezeigt hatte, nach Kontrolle der Noten kein Unterschied mehr zu erkennen. Die Ausnahme stellte die Facette „unbestimmt-abwartend“ dar, auf der die Lehrkräfte ihr Verhalten weiterhin als gegenüber Kindern mit Förderbedarf geringer ausgeprägt beschrieben ( $\beta = -.15, p = .024$ ). Beim Vergleich des Lehrkraftverhaltens gegenüber Mädchen und Jungen blieben auch nach Kontrolle der Noten der Kinder drei der sechs beobachteten Unterschiede bestehen: bei gleichen Noten erfuhren Jungen auf den beiden Facetten mit der geringst möglichen Communion („ignorierend“  $\beta = -.21, p = .009$ ; „ermahnend“  $\beta = -.33, p < .001$ ) und auf der Facette mit relativ geringer Communion und stärkster möglicher Agency („fordernd-streng“  $\beta = -.33, p < .001$ ) signifikant höhere Ausprägungen des Verhaltens ihrer Lehrkraft als Mädchen.

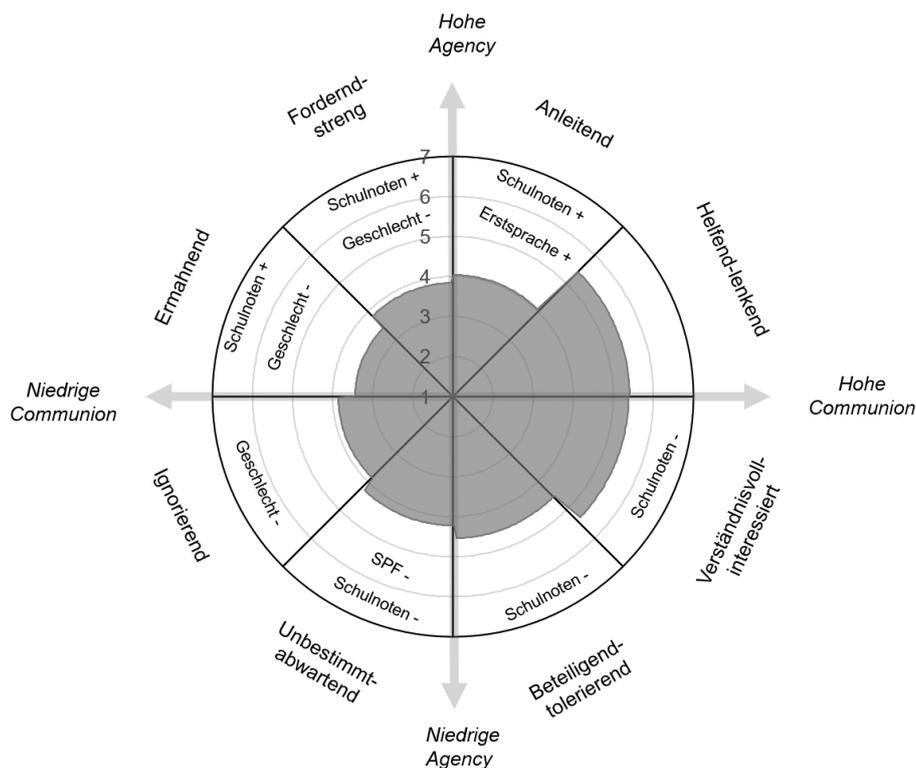
## Diskussion

Gelingt es Grundschullehrkräften, in einer heterogenen Klasse gegenüber jedem Schüler und jeder Schülerin lern-

förderliches Verhalten zu zeigen, welches durch hohe Communion und eine zur Kompetenz des Kindes komplementäre Agency geprägt ist? In einer Befragung von Grundschullehrkräften, die ihren Umgang mit jeweils fünf Kindern ihrer Klasse auf acht Facetten eines Zirkumplex-Modells beschrieben, zeigte sich erwartungsgemäß, dass die Lehrerinnen und Lehrer ihr Verhalten in besonders starkem Maße durch Facetten mit sehr hoher Communion und gleichzeitig in besonders schwachem Maße durch Facetten mit sehr geringer Communion charakterisierten. Frühere Arbeiten zum Zusammenhang zwischen der Communion des Lehrkraftverhaltens und Outcome-Variablen auf Seiten der SuS belegen, dass ein solches Verhaltensmuster besonders günstig für Motivation und Lernerfolg von Kindern und Jugendlichen ist (Metaanalysen Roorda et al., 2011, 2017).

## Komplementarität der Agency des Lehrkraftverhaltens

In Arbeiten, in denen Lehrstile von Lehrkräften gegenüber ihrer Klasse mit Hilfe von Zirkumplex-Modellen untersucht worden sind, ist gezeigt worden, dass moderat starke Agency für die Motivation der Lernenden besonders günstig ist (Aelterman et al., 2018; Wubbels & Brekelmans, 2005). Im Rahmen des Paradigmas der Selbstbestimmungstheorie (Deci & Ryan, 2000) wird dies damit er-



**Abbildung 3.** Mittlere Ausprägungen des Lehrkraftverhaltens gegenüber allen Kindern auf den acht Facetten des Zirkumplex unter Kontrolle aller Merkmale einschließlich Schulnoten.

Anmerkung: Eingezeichnet sind sämtliche signifikant unterschiedliche Ausprägungen für Gruppen von Kindern und die Richtung des Effekts. Kodierung: Geschlecht (0 = männlich, 1 = weiblich); Erstsprache (0 = deutsch, 1 = nichtdeutsch); SPF (0 = nein, 1 = ja); Schulnoten (von 1 = sehr gut bis 6 = ungenügend).

klärt, dass eine zu starke Agency die Lernenden in ihrem Erleben von Autonomie einschränkt und entsprechend kontrolliert-fremdbestimmte, statt autonom-selbstbestimmte Formen von Motivation begünstigt werden. Zu schwache Agency der Lehrkraft kann jedoch auch bedeuten, dass die Kinder keine klare Struktur des Unterrichts und Handlungsprogramms erkennen können (Jang et al., 2010). Vor diesem Hintergrund haben wir die Annahme formuliert, dass für verschiedene Kinder durchaus unterschiedliche Stärken von Agency besonders angemessen sein können. Genauer haben wir vermutet, dass Lehrkräfte eine zur in den Noten reflektierten fachlichen und fachübergreifenden Kompetenz des Kindes komplementäre Ausprägung der Agency zeigen.

Um diese Annahme zu prüfen, haben wir die mittleren Noten der Kinder in Deutsch, Mathematik und in bis zu zwei weiteren Fächern verwendet und ihren Einfluss auf das Lehrkraftverhalten geschätzt. Wie erwartet fanden wir moderate bis starke Effekte der mittleren Noten der Kinder auf das Lehrkraftverhalten, welche in Richtung und Stärke systematisch mit dem Ausmaß an Agency des Lehrkraftverhaltens variierten. Wir interpretieren dieses Befundmuster als eine Bestätigung des sog. Komplementaritätsprinzips in der Interpersonalen Theorie (z. B. Sadler et al., 2011): Lehrkräfte verhalten sich gegenüber einem Kind umso stärker agentisch, je schlechter seine Noten sind, d. h., je mehr dieser Schüler oder diese Schülerin noch fachlicher Unterstützung sowie Leitung und Kontrolle bedarf. Die moderaten bis starken Korrelationen, die wir zwischen der Agency der Lehrkraft und den mittleren Noten des Kindes gefunden haben, sind gleichzeitig ein Hinweis auf die Validität der Messung des Lehrkraftverhaltens durch den neu konstruierten *Fragebogen zum Lehrkraftverhalten in dyadischen Lehrkraft-Lernenden-Beziehungen*. Unseres Wissens ist unsere Studie die erste, die empirische Evidenz für das Komplementaritätsprinzip – unter Nutzung von Noten als Kompetenzmaß für das Kind und einer Beschreibung der Lehrkraft ihres Verhaltens gegenüber dem Kind – erbracht hat (vgl. Metaanalyse Roorda et al., 2020). Wir vermuten, dass Lehrkräfte durch eine zur in seinen Noten reflektierten fachlichen und fachübergreifenden Kompetenz des Kindes komplementäre Agency nicht nur das Erleben von Autonomie des Kindes stärken, sondern auch sein subjektives Erleben von Kompetenz, das ebenso wie Autonomieerleben einen Prädiktor autonom-selbstbestimmter Formen von Motivation darstellt (Deci & Ryan, 2000). Ein Autonomieunterstützungsstil der Lehrkraft (Aelterman et al., 2018), der sich durch besonders geringe Agency auszeichnet und der Selbstbestimmungstheorie zufolge einen günstigen Stil darstellt, bedeutet unseren Ergebnisse zufolge nicht für jedes Kind die optimale Unterstützung seiner Motivation (vgl. auch Jang et al., 2010). Eine direkte Prüfung dieser

Annahme ist mit unseren Daten jedoch nicht möglich, da wir keine Angaben der Kinder zu ihrer Motivation erhoben haben.

## Unterschiede im Lehrkraftverhalten zwischen verschiedenen Gruppen von Kindern

Wir haben weitergehend untersucht, ob sich Lehrkräfte unterschiedlich gegenüber Gruppen von Kindern verhalten, die sich in ihren – durch Noten erfassten – mittleren Kompetenzen voneinander unterscheiden.

### Kinder mit oder ohne sonderpädagogischen Förderbedarf

Die Unterschiede im Verhalten der Lehrkräfte gegenüber Kindern mit und ohne SPF, die wir bei Nichtberücksichtigung der Noten der Kinder auf vier der acht Facetten („unbestimmt-abwartend“, „ermahnend“, „fordernd-streng“, „anleitend“) gefunden hatten, waren mit Ausnahme der geringeren Ausprägung auf der Facette „unbestimmt-abwartend“ nicht mehr sichtbar, wenn für die Noten kontrolliert wurde. Dies bedeutet, die Unterschiede auf den drei Facetten, die durch relativ starke Agency charakterisiert sind, konnten durch die Komplementarität des Lehrkraftverhaltens zur durch die Note erfassten Kompetenz des jeweiligen Kindes erklärt werden. Der für die Facette „unbestimmt-abwartend“ verbleibende Effekt kann bedeuten, dass Lehrkräfte für Kinder mit SPF besonders wahrscheinlich ein Konzept oder einen Plan haben, wie sie sie am Unterrichtsgeschehen beteiligen, und zwar in stärkerem Ausmaß, als dies allein durch die (im Mittel) noch geringere Kompetenz dieser Kinder gerechtfertigt ist. Diese Kinder stellen in einer inklusiven Schulklasse eine Minorität dar und erhalten deshalb möglicherweise bei Planung und Monitoring des Unterrichtsverlaufs von der Lehrkraft besondere Aufmerksamkeit.

### Kinder mit deutscher oder anderer Erstsprache

Die nach Erstsprache der Kinder getrennte Analyse ergab, dass nach Kontrolle des Einflusses der Noten einer der drei beobachteten Unterschiede im Lehrkraftverhalten bestehen blieb, nämlich der Effekt für die Facette „anleitend“, demnach Kinder mit nichtdeutscher Erstsprache von der Lehrkraft wahrscheinlicher solche durch starke Agency (und relativ starke Communion) charakterisierten Verhaltensweisen erfuhren als Kinder mit deutscher Erstsprache. Dies bedeutet, die Lehrkräfte lassen sich in ihrem Verhalten gegenüber diesen Kindern noch von anderen mit ihrer Gruppenzugehörigkeit assoziierten Merkmalen als ihrer in den Noten reflektierten Kompetenz leiten. Solche assoziierten Merkmale könnten in Deutschland beste-

hende negative Erwartungen oder Stereotype betreffend die Kompetenz dieser Kinder sein (Appel, Weber & Kronberger, 2015; Froehlich & Schulte, 2019). So fanden z. B. Tobisch und Dresel (2017), dass Grundschullehrkräfte, die auf der Grundlage eines fiktiven Zeugnisses Prognosen für die Noten machen sollten, die die Kinder im nächsten Test erzielen würden, für SuS ohne Migrationshintergrund eine Verbesserung gegenüber ihren bisherigen Noten, für SuS mit Migrationshintergrund hingegen gleichbleibende Noten erwarteten. Möglicherweise erachten Lehrkräfte es also bei Kindern mit nichtdeutscher Erstsprache für eher erforderlich, selbst stark lenkend und überwachend zu sein (hohe Agency), um bei diesen Kindern dieselben Lernzuwächse hervorzurufen wie bei Kindern mit deutscher Erstsprache. Problematisch kann dabei jedoch sein, dass eine stärkere als durch die Kompetenz der Kinder gerechtfertigte Agency des Lehrkraftverhaltens ihr Autonomieerleben einschränken (Alivernini, Cavicchiolo, Manganeli, Chirico & Lucidi, 2019) und vermittelt darüber autonom-selbstbestimmte Formen von Motivation abschwächen kann (Deci & Ryan, 2000).

Eine alternative Erklärung für die stärkere Agency, die wir gegenüber Kindern mit nichtdeutscher Erstsprache gefunden haben, ist, dass Lehrkräfte diesen SuS bessere Noten geben als durch ihre Kompetenzen gerechtfertigt wäre, um sie in ihrer Motivation zu stärken (Motivierungsfunktion von Noten; vgl. Krampen, 1984; Ingenkamp, 1975). Dies würde implizieren, dass die Agency der Lehrkraft nur deshalb nach Kontrolle der Noten noch signifikant stärker ausgeprägt war als gegenüber Kindern mit deutscher Erstsprache, weil die Benotung der Kinder mit nichtdeutscher Erstsprache gemessen an ihren tatsächlichen fachlichen Kompetenzen pädagogisch motiviert „zu wohlwollend“ ausfiel. Gegen diese Annahme spricht jedoch, dass Kinder mit Migrationshintergrund in der deutschen Grundschule keine besseren, sondern sogar schlechtere Noten erhalten, wie Kiss (2013) anhand der PIRLS-Daten 2001 und der PISA-Daten 2003 zeigte: nach Kontrolle mathematischer Fähigkeiten und mündlicher Beteiligung im Mathematikunterricht wurden Kinder mit Zuwanderungshintergrund im Mittel 0.27 Noten schlechter beurteilt als Kinder ohne Zuwanderungshintergrund (für ähnliche Befunde für zwei Schulfächer, Deutsch und Mathematik, siehe Lüdemann & Schwerdt, 2013).

Eine weitere Erklärung des stärker agentischen Verhaltens der Lehrkräfte gegenüber Kindern mit nichtdeutscher Erstsprache können Unterschiede im wahrgenommenen oder tatsächlichen Verhalten der Kinder sein. Unter Verwendung des Datensatzes aus der TIMSS-Transition Study fanden Brandmiller, Dumont und Becker (2020), dass nach Kontrolle der durch standardisierte Tests gemessenen Kompetenz der Viertklässlerinnen und Viertklässler Lehrkräfte Kindern mit Migrationshintergrund ein weni-

ger angepasstes Verhalten und eine geringere Lernmotivation attestierten als Kindern ohne Migrationshintergrund. Unsere Ergebnisse könnten also bedeuten, dass die Lehrkräfte auf ein weniger angepasstes und motiviertes Verhalten der Kinder mit stärkerer Agency reagierten. Eine Entscheidung, ob die stärkere Agency der Lehrkräfte tatsächliche Unterschiede im Verhalten der Kinder reflektiert oder auf Voreingenommenheiten der Lehrkräfte zurückgeht, kann aber weder auf der Grundlage der Daten von Brandmiller et al. (2020), noch der Daten aus unserer Studie entschieden werden.

### Jungen und Mädchen

Gegenüber Jungen beschrieben die Lehrkräfte ihr Verhalten auf Facetten mit hoher Agency („fordernd-streng“, „anleitend“) als stärker und auf der Facette mit geringer Agency „beteiligend-tolerierend“ als geringer ausgeprägt als gegenüber Mädchen. Dies kann mit dem gut belegten Befund erklärt werden, dass Mädchen relativ zu Jungen stärker selbstreguliert lernen, also weniger Anleitung oder Kontrolle durch die Lehrkraft benötigen als Jungen (z. B. Duckworth & Seligman, 2006; Kuhl & Hannover, 2012). Auffällig war jedoch, dass sich auch auf anderen Facetten Unterschiede im Lehrkraftverhalten zeigten. So verhielten sich die Lehrkräfte gegenüber Jungen auch stärker „ignorierend“ und „ermahnend“ (Facetten mit geringer Communion) und weniger „verständnisvoll-interessiert“ (Facetten mit hoher Communion) als gegenüber Mädchen. Drei der insgesamt sechs Unterschiede blieben bestehen, nachdem für den Einfluss der Noten der Kinder kontrolliert worden war: nämlich das stärker fordernd-streng, das stärker ermahnende und das stärker ignorierende Verhalten der Lehrkräfte gegenüber Jungen als gegenüber Mädchen. Auffällig ist, dass diese drei Verhaltensfacetten im linken Teil des Zirkumplex verortet sind, d. h. sämtlich durch eine (sehr) geringe Communion charakterisiert sind.

Eine Erklärung kann darin liegen, dass Jungen ein weniger zu den schulischen Anforderungen passendes Arbeits- und Sozialverhalten zeigen als Mädchen. So fanden Gentrup, Rjosk, Stanat und Lorenz (2018), dass Lehrkräfte Jungen sowohl im Fach Deutsch als auch im Fach Mathematik ein weniger günstiges Arbeitsverhalten und eine signifikant geringere Motivation attestierten als Mädchen. Glock und Kleen (2017) fanden bei Master-Lehramtsstudierenden im Praktikum, dass sie in einer impliziten Messung (Implicit Association Test) Jungen mit negativem Verhalten (erfasst über Adjektive wie z. B. loud, insubordinate) und Mädchen mit positivem Verhalten (z. B. attentive, well-behaved) assoziierten. Umso stärker die Lehramtsstudierenden Jungen mit negativem Verhalten assoziierten, desto eher wollten sie gegenüber einem fiktiven Schüler, der den Unterricht durch Schwatzen störte,

mit harschen Interventionen bestrafen (z.B. den Schüler vor die Tür schicken, eine Schulkonferenz einberufen). Die Assoziationen mit Mädchen zeigten keinen Zusammenhang zu Interventionen, die die Versuchsteilnehmenden ergreifen wollten. Han, Elsässer, Lang und Ditton (2017) fanden, dass Lehrkräfte Jungen aus zweiten, dritten und vierten Klassen hinsichtlich ihres Arbeitsverhaltens und insbesondere hinsichtlich ihres Sozialverhaltens hochsignifikant schlechter bewerteten als Mädchen und dass eine um eine Standardabweichung verbesserte Bewertung des Arbeitsverhaltens zu einer um 0.18 besseren Deutschnote und zu einer um .016 besseren Mathematiknote führte (unter Kontrolle des Sozialverhaltens und der Deutsch- bzw. Mathematikkompetenz). Brandmiller et al. (2020) fanden, dass Lehrkräfte Viertklässlerinnen ein deutlich stärker angepasstes Verhalten im Klassenzimmer attestierten als Viertklässlern, auch nachdem für den Einfluss der Leistung und der von den Kindern und ihren Eltern angegebenen Lernmotivation kontrolliert worden war. Weder auf Grundlage der Daten aus den zitierten Studien noch unserer eigenen Daten kann jedoch geschlossen werden, ob die Beschreibungen der Lehrkräfte auf Unterschiede im tatsächlichen Verhalten der Mädchen und Jungen zurückgehen oder eine Folge negativer Stereotype sind, von denen Jungen in Bezug auf ihr Sozialverhalten im Kontext Schule betroffen sind: von ihren Peers und Lehrkräften werden sie als unangepasst, sozial auffällig und in der Schule nicht leistungsbereit stigmatisiert (Heyder & Kessels, 2013, 2015; Latsch & Hannover, 2014).

Eine andere mögliche Erklärung des weniger kommunalen Verhaltens der Lehrkräfte gegenüber Jungen bieten Geschlechtsunterschiede in der Communion, die Jungen versus Mädchen ihrerseits gegenüber der Lehrkraft zeigen. So fanden viele Studien, dass Mädchen eine positive Einstellung zur Schule, mehr Schulfreude und ein stärkeres Zugehörigkeitsgefühl zur Schule berichten als Jungen (Arens & Niepel, 2019; van Ophuysen, 2008; Wang & Eccles, 2012) und dass Mädchen sich von ihrer Lehrkraft stärker akzeptiert und unterstützt fühlen als Jungen dies tun (Arens & Niepel, 2019; Lietaert, Roorda, Laevers, Verschueren De Fraine, 2015; Makri-Botsari, 2015; Wentzel, Battle, Russell & Looney, 2010), was möglicherweise in stärker kommunalem Verhalten von Mädchen relativ zu Jungen gegenüber ihren Lehrkräften seinen Ausdruck findet. Dem Komplementaritätsprinzip der Interpersonalen Theorie folgend (z.B. Sadler et al., 2011) ist Verhalten auf der Communion-Dimension komplementär, wenn es *gleichgerichtet* ist (im Unterschied zur Agency-Dimension, auf der komplementäre Verhaltensweisen entgegengesetzt sind). So wird in einer interpersonalen Beziehung warmes, herzliches Verhalten wahrscheinlicher mit warmem und herzlichem Verhalten (als mit kaltem oder distanzierendem Verhalten) beantwortet. Dies könnte be-

deuten, dass Lehrkräfte komplementär zur mittleren Communion der Kinder gegenüber Jungen stärkere Verhaltensausprägungen auf Facetten mit geringer Communion zeigen als gegenüber Mädchen.

Auch wenn eine geringere Communion des Verhaltens von Jungen (relativ zu dem von Mädchen) eine Erklärung dafür bieten sollte, warum Lehrkräfte gegenüber Jungen stärkere Ausprägungen auf Verhaltensfacetten mit geringer Communion berichteten als gegenüber Mädchen, ist das stärker fordernd-strenge, ermahnende und ignorierende Verhalten gegenüber Jungen, das sich auch nach Kontrolle des Einflusses der Noten zeigte, als kritisch zu bewerten. Professionelles Lehrkraftverhalten sollte stets durch hohe Communion gekennzeichnet sein, und zwar auch gegenüber Kindern, denen es (noch) nicht gelingt, selbst Verhalten zu zeigen, das durch hohe Communion charakterisiert ist.

Verhaltensfacetten mit geringer Communion haben sich als besonders abträglich für Motivation und Lernen von SuS erwiesen (Metaanalysen Roorda et al., 2011, 2017). Gleichzeitig sind stark agentische Verhaltensweisen von Lehrkräften mit kontrollierten Formen von Motivation, mit Reaktanz und Amotivation auf Seiten der Lernenden assoziiert (Aelterman et al., 2018). Die stärkeren Verhaltensausprägungen auf Facetten mit hoher Agency und geringer Communion, die die Lehrkräfte angaben, gegenüber Jungen zu zeigen, könnten somit zu einer Erklärung des geringeren Bildungserfolgs von Jungen beitragen (vgl. z.B. Hannover & Kessels, 2011; Hannover & Ollrogge, 2021): Jungen fühlen sich weniger akzeptiert und unterstützt von der Lehrkraft (z.B. Arens & Niepel, 2019; Lietaert et al., 2015) und erleben in der Schule weniger autonom-selbstbestimmte Formen von Motivation als Mädchen (z.B. Alivernini, Manganelli, Cavicchiolo, Girelli, Biasi & Lucidi, 2018; Lietaert et al., 2015; Vantieghem & Van Houtte, 2018), mit entsprechenden Auswirkungen auf ihren schulischen Lernerfolg.

Ein erster Schritt in Richtung einer Veränderung der Verhaltensweisen von Lehrkräften gegenüber Jungen könnte darin bestehen, sie über unsere Befunde aufzuklären. Während Lehrkräfte eine hohe Sensibilität gegenüber negativen Stereotypen über Kinder mit SPF oder Kinder mit nichtdeutscher Erstsprache und potentiellen Einflüssen auf ihr Verhalten haben dürften, sind sie möglicherweise gegenüber potentiellen stereotypbasierten Einflüssen auf ihr Verhalten gegenüber Jungen weniger aufmerksam – gelten Jungen doch noch immer als Mitglieder des „starken Geschlechts“. Indem sie in Bezug auf Benachteiligungen von Jungen weniger sensibel sind als in Bezug auf Benachteiligungen von Kindern, die von einem negativen leistungsbezogenen Stereotyp betroffen sind, tragen Lehrkräfte möglicherweise, ohne dass sie dies beabsichtigen, zur Aufrechterhaltung von Geschlechtsunter-

schieden im schulischen Verhalten, in Motivation und Leistungen bei (vgl. Gansen, 2019).

## Limitationen

Da wir das jeweilige Kind nicht zu seinem eigenen Verhalten oder seiner Wahrnehmung des Verhaltens der Lehrkraft in dyadischen Interaktionen befragt haben, liegen keine dyadischen Daten vor, wechselseitige Einflüsse oder Interdependenzen können wir also nicht untersuchen. Weil jede Lehrkraft mit mehreren Kindern in der Klasse interagiert, haben wir angenommen, dass jede Dyade nicht nur von idiosynkratischen Merkmalen der beiden beteiligten Personen beeinflusst ist, sondern auch durch Gruppenmerkmale des Kindes.

Eine bedeutsame Einschränkung unserer Studie besteht darin, dass wir die Kompetenzen der Kinder nicht über standardisierte Leistungsmessungen, sondern über Noten erfasst haben. Denn Noten reflektieren nicht nur fachliche, sondern auch fachübergreifende Kompetenzen des Kindes, die Lehrkräfte bei der Notengebung miteinbeziehen. Darüberhinausgehend sind Noten auch von Gruppenzugehörigkeiten (z.B. Kiss, 2013; Lüdemann & Schwerdt, 2013) oder von Stereotypen über diese Gruppen (z.B. Bonefeld & Dickhäuser, 2018; Hofer, 2015) sowie von Bezugsgruppeneffekten (z.B. Südkamp & Möller, 2009) beeinflusst. Schließlich variiert die an Testleistungen gemessene Kriteriumsvalidität von Schulnoten auch zwischen Lehrkräften (z.B. Thiel & Valtin, 2002). Da Noten nicht nur eine Berichts-, Rückmelde- und Selektionsfunktion, sondern auch eine Motivierungs-, Anreiz- und Disziplinarfunktion haben, also pädagogisch motiviert eingesetzt werden (z.B. Krampen, 1984; Ingenkamp, 1975), bleibt unklar, was genau für die Veränderungen in den Effekten verantwortlich ist, die wir beim Vergleich des Lehrkraftverhaltens zwischen verschiedenen Gruppen von Kindern gefunden hatten, wenn entweder für die Noten der Kinder kontrolliert worden war oder aber nicht.

Auch wurden die Noten in vielen Fällen von den befragten Lehrkräften selbst vergeben, d.h., sowohl die Verhaltensbeschreibungen als auch die Noten können in gleicher Weise von Merkmalen der Lehrkraft – wie z.B. ihren Erwartungen – und / oder des Kindes beeinflusst worden sein (common-method bias). Gleichzeitig ist denkbar, dass Lehrkräfte ihr Verhalten als zur Ausprägung der selbst vergebenen Noten passend beschreiben. Deshalb sollten unsere Befunde unter Nutzung von Leistungsdaten der Kinder statt ihrer Noten repliziert werden.

Wir hatten die Lehrkräfte gebeten, ihr Verhalten gegenüber fünf von ihnen gewählten Kindern ihrer Klasse zu beschreiben, die sich in verschiedenen Merkmalen deut-

lich voneinander unterscheiden sollten. Es kann nicht ausgeschlossen werden, dass die Lehrkräfte sich durch diese Instruktion dazu aufgefordert fühlten, besonders auffällige, statt repräsentative Kinder der jeweiligen Gruppe auszuwählen. Dadurch, dass wir die Lehrkräfte nicht direkt gefragt haben, wie sie sich z.B. gegenüber Jungen oder gegenüber Kindern mit SPF verhalten, haben wir versucht Tendenzen zu verhindern, die Beschreibungen des eigenen Verhaltens dahingehend zu korrigieren, dass es möglichst „vorurteilsfrei“, also unbeeinflusst von Gruppenzugehörigkeiten der Kinder, erscheint (z.B. Fehr, Sassenberg & Jonas, 2012). Dennoch kann nicht ausgeschlossen werden, dass die Lehrkräfte ihr Verhalten in sozial erwünschter Weise dargestellt haben. Dies könnte mit dazu beigetragen haben, dass die Lehrkräfte sich insbesondere auf Facetten mit starker Communion hohe und auf Facetten mit schwacher Communion geringe Ausprägungen attestiert haben. Die systematischen Unterschiede im Lehrkraftverhalten gegenüber verschiedenen Gruppen von Kindern, die sichtbar wurden, nachdem wir die auf einzelne Kinder bezogenen Verhaltensbeschreibungen der Lehrkräfte für verschiedene Gruppen von Kindern zusammengefasst hatten, können hingegen eher nicht mit sozialem Erwünschtheitsstreben erklärt werden. Denn es ist unwahrscheinlich anzunehmen, dass die Verhaltensbeschreibungen der Lehrkräfte je nach Gruppenzugehörigkeiten des Kindes in unterschiedlicher oder unterschiedlich starker Weise sozial erwünscht verzerrt waren.

Zukünftige Studien sollten nicht nur Selbstangaben von Lehrkräften erheben, sondern ihr Verhalten direkt beobachten oder von den Kindern selbst beschreiben lassen. Beispielsweise ist bekannt, dass Lehrkräfte, Schülerinnen und Schüler und externe Beobachtungspersonen in ihren Beurteilungen von Unterrichtsqualität einer Lehrkraft nur schwach übereinstimmen (Fauth, Göllner, Lenske, Praetorius & Wagner, 2020). Ähnlich fanden Begrich, Fauth und Kunter (2020), dass diejenigen Einschätzungen von Unterrichtsqualität, die sich als für die Lernergebnisse der Schülerinnen und Schülern prädiktiv erwiesen, auf unterschiedliche Aspekte der Unterrichtsqualität bezogen, in Abhängigkeit davon, ob sie von Schülerinnen und Schülern oder von Bildungsforschungsexperten und -expertinnen abgegeben worden waren. Vor diesem Hintergrund ist es denkbar, dass sich unsere Ergebnisse anders dargestellt hätten, wären die Verhaltensweisen der Lehrkraft durch andere Personen als die Lehrkräfte selbst beschrieben worden. Schließlich bleibt es eine Aufgabe für zukünftige Forschung, die Selbstangaben von Lehrkräften zu ihrem Verhalten auf den Dimensionen Agency und Communion mit Angaben der Kinder zu ihrem Erleben von Autonomie, Kompetenz und sozialer Unterstützung durch die Lehrkraft sowie zu ihrer Motivation und ihren Lernergebnissen zusammenzuführen, um unsere Annahme empirisch zu

prüfen, dass die Selbstbeschreibungen der Lehrkräfte mit ihrem tatsächlichen Verhalten im Klassenzimmer korrespondieren, das dann auf die Entwicklung der Kinder Einfluss nimmt.

## Elektronisches Supplement (ESM)

Das elektronische Supplement ist mit der Online-Version dieses Artikels verfügbar unter <https://doi.org/10.1024/1010-0652/a000327>

**ESM 1.** Ergebnisse der Regressionsanalysen zur Vorhersage des Lehrkraftverhaltens (Tabelle)

## Literatur

- Abele, A. E., Hauke, N., Peters, K., Louvet, E., Szymkow, A. & Duan, Y. (2016). Facets of the fundamental content dimensions: Agency with competence and assertiveness – Communion with warmth and morality. *Frontiers in Psychology*, 7, Article 1810. <https://doi.org/10.3389/fpsyg.2016.01810>
- Aelterman, N., Vansteenkiste, M., Haerens, L., Soenens, B., Fontaine, J. R. J. & Reeve, J. (2018). Toward an integrative and fine-grained insight in motivating and demotivating teaching styles: The merits of a circumplex approach. *Journal of Educational Psychology*, 111, 497 – 521.
- Alivernini, F., Cavicchiolo, E., Manganelli, S., Chirico, A. & Lucidi, F. (2019). Support for autonomy at school predicts immigrant adolescents' psychological well-being. *Journal of Immigrant and Minority Health*, 21, 761 – 766.
- Alivernini, F., Manganelli, S., Cavicchiolo, E., Girelli, L., Biasi, V. & Lucidi, F. (2018). Immigrant background and gender differences in primary students' motivations toward studying. *The Journal of Educational Research*, 111, 603 – 611.
- Anders, Y., McElvany, N. & Baumert, J. (2010). Die Einschätzung lernrelevanter Schülermerkmale zum Zeitpunkt des Übergangs von der Grundschule auf die weiterführende Schule: Wie differenziert urteilen Lehrkräfte? In Y. Anders & K. Maaz (Hrsg.), *Der Übergang von der Grundschule in die weiterführende Schule. Leistungsgerechtigkeit und regionale, soziale und ethnisch-kulturelle Disparitäten* (S.313–330). Bundesministerium für Bildung und Forschung.
- Appel, M., Weber, S. & Kronberger, N. (2015). The influence of stereotype threat on immigrants: Review and meta-analysis. *Frontiers in Psychology*, 6, 900.
- Arens, A. K. & Niepel, C. (2019). School attitude and perceived teacher acceptance: Developmental trajectories, temporal relations, and gender differences. *British Journal of Educational Psychology*, 89, 689 – 706.
- Baumrind, D. (1968). Authoritarian vs authoritative parental control. *Adolescence*, 3, 255 – 272.
- Bayer, M. & Zinn, S. (2018). *Zur Entwicklung des Zusammenhangs und der Erklärbarkeit von Leistungen und Kompetenzen bei Schülerinnen und Schülern der Sekundarstufe*. Universität Bamberg: LfBi-Working Papers 77.
- Begrich, L., Fauth, B. & Kunter, M. (2020). Who sees the most? Differences in students' and educational research experts' first impressions of classroom instruction. *Social Psychology of Education*, 23, 673 – 699.
- 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>
- Bortz, J., & Döring, N. (2016). *Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler* (5., überarb. Aufl.). Heidelberg: Springer.
- Brandmiller, C., Dumont, H. & Becker, M. (2020). Teacher perceptions of learning motivation and classroom behavior: The role of student characteristics. *Contemporary Educational Psychology*. <https://doi.org/10.1016/j.cedpsych.2020.101893>
- Deci, E. & Ryan, R. (2000). The “what” and “why” of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry*, 11, 227.
- Duckworth, A. & Seligman, M. (2006). Self-discipline gives girls the edge: Gender in self-discipline, grades, and achievement test scores. *Journal of Educational Psychology*, 98, 198 – 208.
- Fauth, B., Göllner, R., Lenske, G., Praetorius, A.-K. & Wagner, W. (2020). Who sees what?: conceptual considerations on the measurement of teaching quality from different perspectives. *Zeitschrift für Pädagogik*, 66 (Beiheft 1), 138 – 155.
- Fehr, J., Sassenberg, K. & Jonas, K. J. (2012). Willful stereotype control: The impact of internal motivation to respond without prejudice on the regulation of activated stereotypes. *Zeitschrift für Psychologie*, 220, 180 – 186.
- Fiske, S. T. (2018). Stereotype content: Warmth and competence endure. *Current Directions in Psychological Science*, 27, 67 – 73.
- Froehlich, L. & Schulte, I. (2019). Warmth and competence stereotypes about immigrant groups in Germany. *PLoS ONE*, 14 (9), e0223103. Verfügbar unter <https://doi.org/10.1371/journal.pone.0223103>
- Gansen, H. M. (2019). Push-ups versus clean-up: Preschool teachers' gendered beliefs, expectations for behavior, and disciplinary practices. *Sex Roles*, 80, 393 – 408.
- Gentrup, S., Rjosk, C., Stanat, P. & Lorenz, G. (2018). Einschätzungen der schulischen Motivation und des Arbeitsverhaltens durch Grundschullehrkräfte und deren Bedeutung für Verzerrungen in Leistungserwartungen. *Zeitschrift für Erziehungswissenschaft*, 21, 867 – 891.
- Glock, S. & Kleen, H. (2017). Gender and student misbehavior: Evidence from implicit and explicit measures. *Teaching and Teacher Education*, 67, 93 – 103.
- Girard, J., Zimmermann, J. & Wright, A. (2020). *Circumplex: Analysis and visualization of circular data* (R package version 0.3.8). Verfügbar unter <https://CRAN.R-project.org/package=circumplex>
- Gurtman, M. B. & Pincus, A. L. (2003). The circumplex model: Methods and research applications. In I. B. Weiner, I. A. Schinka & W. F. Velicer (Hrsg.), *Handbook of Psychology: Vol. 2. Research methods in psychology* (S.407 – 428). New York: Wiley.
- Han, M., Elsäßer, S., Lang, V. & Ditton, H. (2017). Geschlechtsspezifische Benotung? Der Einfluss der von Lehrkräften eingeschätzten Verhaltensmerkmale auf die Notengebung. *Zeitschrift für Soziologie der Erziehung und Sozialisation*, 37, 174 – 194.
- Hannover, B. & Kessels, U. (2011). Sind Jungen die neuen Bildungsverlierer? Empirische Evidenz für Geschlechterdisparitäten zugunsten von Jungen und Erklärungsansätze. *Zeitschrift für Pädagogische Psychologie*, 25, 89 – 103.
- Hannover, B. & Ollrogge, K. (2021). Bildungsungleichheiten zwischen den Geschlechtern. *Bundeszentrale für Politische Bildung*. Verfügbar unter <https://www.bpb.de/gesellschaft/bildung/zukunftsbildung/315992/geschlechterungleichheiten>
- Heyder, A. & Kessels, U. (2013). Is school feminine? Implicit gender stereotyping of school as a predictor of academic achievement. *Sex Roles*, 69, 605 – 617.

- Heyder, A. & Kessels, U. (2015). Do teachers equate male and masculine with lower academic engagement? How students' gender enactment triggers gender stereotypes at school. *Social Psychology of Education: An International Journal*, 18, 467–485.
- Hofer, S. I. (2015). Studying gender bias in physics grading: The role of teaching experience and country. *International Journal of Science Education*, 37, 2879–2905.
- Ingenkamp, K. (1975). *Pädagogische Diagnostik. Ein Forschungsbericht über Schülerbeurteilung in Europa. Trendbericht im Auftrag des Europarats in Straßburg*. Weinheim: Beltz.
- Ingenkamp, K. (Hrsg.) (1995). *Die Fragwürdigkeit der Zensurenggebung. Texte und Untersuchungsberichte* (9. Aufl.). Weinheim: Beltz.
- Jang, H., Reeve, J. & Deci, E. L. (2010). Engaging students in learning activities: It is not autonomy support or structure but autonomy support and structure. *Journal of Educational Psychology*, 102, 588–600.
- Klassen, R. M., Perry, N. E. & Frenzel, A. C. (2012). Teachers' relatedness with students: An underemphasized component of teachers' basic psychological needs. *Journal of Educational Psychology*, 104, 150–165.
- KMK – Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2015). *Übergang von der Grundschule in Schulen des Sekundarbereichs I und Förderung, Beobachtung und Orientierung in den Jahrgangsstufen 5 und 6 (sog. Orientierungsstufe)*. Verfügbar unter [http://www.kmk.org/fileadmin/veroeffentlichungen\\_beschluesse/2015/2015\\_02\\_19-Uebergang\\_Grundschule-SI-Orientierungsstufe.pdf](http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2015/2015_02_19-Uebergang_Grundschule-SI-Orientierungsstufe.pdf)
- Knierim, B., Raufelder, D. & Wettstein, A. (2017). Die Lehrer-Schüler-Beziehung im Spannungsfeld verschiedener Theorieansätze. *Psychologie in Erziehung und Unterricht*, 64, 35–48.
- Krampen, G. (1984). *Welche Funktionen haben Zensuren in der Schule? Eine empirische Untersuchung zu Funktionswahrnehmungen von Lehrern, Lehramtskandidaten und Schülern*. Verfügbar unter [https://www.uni-trier.de/fileadmin/fb1/prof/PSY/KPW/1984\\_welche\\_Funktionen\\_haben\\_Zensuren.pdf](https://www.uni-trier.de/fileadmin/fb1/prof/PSY/KPW/1984_welche_Funktionen_haben_Zensuren.pdf)
- Kuhl, P. & Hannover, B. (2012). Differentielle Benotungen von Mädchen und Jungen? Der Einfluss der von der Lehrkraft eingeschätzten Kompetenz zum selbstgesteuerten Lernen. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 44, 153–162.
- Latsch, M. & Hannover, B. (2014). Smart girls, dumb boys!? How the discourse on "failing boys" impacts performances and motivational goal orientation in German school students. *Social Psychology*, 45, 112–126.
- Leary, T. (1957). *Interpersonal diagnosis of personality*. New York: Ronald.
- Lietaert, S., Roorda, D., Laevers, R., Verschueren, K. & De Fraine, B. (2015). The gender gap in student engagement: The role of teachers' autonomy support, structure, and involvement. *British Journal of Educational Psychology*, 85, 498–518.
- Lintorf, K. (2012). *Wie vorhersagbar sind Grundschulnoten? Prädiktionskraft individueller und kontextspezifischer Merkmale*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Locke, K. (2011). Circumplex measures of interpersonal constructs. In M. Horowitz & S. Strack (Eds.), *Handbook of interpersonal psychology* (pp. 313–324). Hoboken, NJ: Wiley.
- Lüdemann, E. & Schwerdt, G. (2013). Migration background and educational tracking: Is there a double disadvantage for second-generation immigrants? *Journal of Population Economics*, 26, 455–481.
- Makri-Botsari, E. (2015). Adolescents' unconditional acceptance by parents and teachers and educational outcomes: A structural model of gender differences. *Journal of Adolescence*, 43, 50–62.
- R Core Team (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. Verfügbar unter <https://www.r-project.org/index.html>
- Retelsdorf, J. & Möller, J. (2012). Grundschule oder Gymnasium? Zur Motivation ein Lehramt zu studieren. *Zeitschrift für Pädagogische Psychologie*, 26, 5–17.
- Roorda, D. L., Jak, S., Zee, M., Oort, F. J. & Koomen, H. M. Y. (2017). Affective teacher-student relationships and students' engagement and achievement: A meta-analytic update and zest of the mediating role of engagement. *School Psychology Review*, 46, 239–261.
- Roorda, D., Koomen, H., Spilt, J. & Oort, F. (2011). The influence of affective teacher-student relationships on students' school engagement and achievement: A meta-analytic approach. *Review of Educational Research*, 81, 493–529.
- Roorda, D. L., Zee, M. & Koomen, H. M. Y. (2020). Don't forget student-teacher dependency! A meta-analysis on associations with students' school adjustment and the moderating role of student and teacher characteristics. *Attachment & Human Development*. <https://doi.org/10.1080/14616734.2020.1751987>
- Rothland, M. (2014). Warum entscheiden sich Studierende für den Lehrerberuf? Berufswahlmotive und berufsbezogene Überzeugungen von Lehramtsstudierenden. In E. Terhart, H. Bennewitz & M. Rothland (Hrsg.), *Handbuch der Forschung zum Lehrerberuf* (S. 349–385, 2. Aufl.). Münster: Waxmann.
- Sadler, P., Ethier, N. & Woody, E. (2011). Interpersonal complementarity. In L. Horowitz & S. Strack (Eds.), *Handbook of interpersonal psychology: Theory, research, assessment and therapeutic interventions* (pp. 123–142). Hoboken, NJ: Wiley.
- Spinath, B., Eckert, C. & Steinmayr, R. (2014). Gender differences in school success: What are the roles of students' intelligence, personality and motivation? *Educational Research*, 56, 230–243.
- Spörer, N., Maaz, K., Vock, M., Schröder-Lenzen, A., Luka, T., Bosse, S., et al. (2015). Lernen in der inklusiven Grundschule. Zusammenhänge zwischen fachlichen Kompetenzen, Sozialklima und Facetten des Selbstkonzepts. *Unterrichtswissenschaft*, 43, 22–35.
- Stubbe, T. C., Bos, W. & Schurig, M. (2017). Der Übergang von der Primar- in die Sekundarstufe. In A. Hußmann, H. Wendt, W. Bos, A. Bremerich-Vos, D. Kasper, E.-M. Lankes, N. McElvany, T. C. Stubbe & R. Valtin (Hrsg.), *IGLU 2016. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 235–250). Münster: Waxmann.
- Südkamp, A. & Möller, J. (2009). Referenzgruppeneffekte im Simulierten Klassenraum. Direkte und indirekte Einschätzungen von Schülerleistungen. *Zeitschrift für Pädagogische Psychologie*, 23, 161–174.
- Taxer, J. L., Becker-Kurz, B. & Frenzel, A. C. (2018). Do quality teacher-student relationships protect teachers from emotional exhaustion? The mediating role of enjoyment and anger. *Social Psychology of Education: An International Journal*, 22, 209–226.
- Thiel, O. & Valtin, R. (2002). Eine Zwei ist eine Drei ist eine Vier. Oder: Sind Zensuren aus verschiedenen Klassen vergleichbar? In R. Valtin (Hrsg.), *Was ist ein gutes Zeugnis? Noten und verbale Beurteilungen auf dem Prüfstand* (S. 67–76). Weinheim: Juventa.
- Thoren, K., Hannover, B. & Brunner, M. (2019). Welche Schulen machen mit beim Jahrgangübergreifenden Lernen? Demographische und leistungsbezogene Merkmale unterschiedlich reformbereiter Grundschulen in Berlin. *Psychologie in Erziehung und Unterricht*, 66, 19–32.
- 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: An International Journal*, 20, 731–752.
- Vantieghem, W. & Van Houtte, M. (2018). Differences in study motivation within and between genders: An examination by gender typicality among early adolescents. *Youth & Society*, 50, 377–404.

- van Ophuysen, S. (2008). Zur Veränderung der Schulfreude von Klasse 4 bis 7: Eine Längsschnittanalyse schulformspezifischer Effekte von Ferien und Grundschulübergang. *Zeitschrift für Pädagogische Psychologie*, 22, 293–306.
- Voyer, D. & Voyer, S. D. (2014). Gender differences in scholastic achievement: A meta-analysis. *Psychological Bulletin*, 140, 1174–1204.
- Wang, M.-T. & Eccles, J. S. (2012). Social support matters: Longitudinal effects of social support on three dimensions of school engagement from middle to high school. *Child Development*, 83, 877–895.
- Wendt, H., Schwippert, K. & Stubbe, T. C. (2016). Mathematische und naturwissenschaftliche Kompetenzen von Schülerinnen und Schülern mit Migrationshintergrund. In H. Wendt, W. Bos, C. Selter, O. Köller, K. Schwippert & D. Kasper (Hrsg.): *TIMSS 2015. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich*. Münster; New York: Waxmann.
- Wendt, H. & Schwippert, K. (2017). Lesekompetenzen von Schülerinnen und Schülern mit und ohne Migrationshintergrund. In A. Hußmann, H., Wendt, W., Bos, A., Bremerich-Vos, D. Kasper, E.-M. Lankes, N. McElvany, T. C. Stubbe & R. Valtin (Hrsg.): *IGLU 2016. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich*. Münster; New York: Waxmann.
- Wentzel, K. R., Battle, A., Russell, S. L. & Looney, L. B. (2010). Social support from teachers and peers as predictors of academic and social motivation. *Contemporary Educational Psychology*, 35, 193–202.
- Wright, A. G., Pincus, A. L., Conroy, D. E. & Hilsenroth, M. J. (2009). Integrating methods to optimize circumplex description and comparison of groups. *Journal of Personality Assessment*, 91, 311–322.
- Wubbels, T. & Brekelmans, M. (2005). Two decades of research on teacher–student relationships in class. *International Journal of Educational Research*, 43, 6–24.
- Wubbels, T., Créton, H., Levy, J. & Hooyman, H. (1993). The model for interpersonal teacher behavior. In T. Wubbels & J. Levy (Eds.), *Do you know what you look like? Interpersonal relationships in education* (pp. 13–28). Oxford: Falmer Press/Taylor & Francis.

#### Historie

Manuskript eingereicht: 13.12.2020  
 Manuskript nach Revision angenommen: 07.07.2021  
 Onlineveröffentlichung: 29.07.2021

#### Förderung

Open-Access-Veröffentlichung ermöglicht durch die Freie Universität Berlin.

#### ORCID

Bettina Hannover  
<https://orcid.org/0000-0003-1916-8455>

#### Prof. Bettina Hannover

Freie Universität Berlin  
 Arbeitsbereich Schul- und Unterrichtsforschung  
 Habelschwerdter Allee 45  
 14195 Berlin  
 Deutschland  
[bettina.hannover@fu-berlin.de](mailto:bettina.hannover@fu-berlin.de)