

# 11 Enthusiasmus und Kooperation von Lehrkräften – zwei günstige Bedingungen professionellen Lehrerhandelns?

Christoph Helm

*Lehrerenthusiasmus und -kooperation gelten als zwei förderliche Bedingungen professionellen Lehrerhandelns. Unter anderem wurde gezeigt, dass sie mit Merkmalen guten Unterrichts im Zusammenhang stehen. Die berichteten Analysen gehen der Frage nach, inwiefern sich diese Befunde auf Basis der PIRLS-2016-Daten auch für Österreich bestätigen lassen. Konkret wird untersucht, ob die Professionsmerkmale Enthusiasmus und Kooperation von Lehrkräften mit der Unterrichtsqualitätsfacette „Lernunterstützung“ sowie den Leseleistungen der Schüler/innen positiv assoziiert sind. Darüber hinaus wird der Zusammenhang zwischen Lernunterstützung und Leseleistungen analysiert. Vor dem Hintergrund bestehender Studien, die zeigen, dass schülerorientierte Unterrichtsstrategien (vermutlich als remediale Maßnahme) eher in leistungsschwachen Klassen zu beobachten sind, wird von einem negativen Zusammenhang zwischen Lernunterstützung und der mittleren Leseleistung einer Schulklasse ausgegangen. Mehrebenen-Strukturgleichungsmodelle bestätigen diese Annahme, d. h., in leseschwachen Schulklassen wird von einem höheren Ausmaß an Lernunterstützung berichtet. Gleichzeitig zeigen die Analysen aber auch, dass der Lehrerenthusiasmus die Unterrichtsqualität im Sinne der Lernunterstützung negativ vorhersagt, d. h., das Unterrichtsangebot jener Lehrpersonen, die mit ihrer Arbeit positive Emotionen verbinden, wird entgegen den Erwartungen von Schülerinnen und Schülern als weniger stark lernförderlich wahrgenommen als das Angebot von Lehrpersonen, die weniger enthusiastisch über ihre Arbeit berichten. Diese erwartungswidrigen Befunde werden vor dem Hintergrund der Grenzen der Studie, die u. a. in der Operationalisierung der Unterrichtsqualität liegen, diskutiert.*

## Einleitung

Professionelles Handeln von Lehrkräften – etwa die Gestaltung einer kognitiv anregenden Lernumgebung – ist u. a. von Merkmalen der Lehrperson (z. B. Enthusiasmus, Lehrwissen, Lehrerselbstwirksamkeit) und dem Kontext (z. B. Kooperation im Kollegium, Lehrer-Schüler-Ratio, elterliche Unterstützung) abhängig. Die Forschung zur Frage, welche Merkmale erfolgreiche Lehrer/innen (Personenmerkmale) ausmachen, konzentriert sich aktuell auf die Messung und Analyse der Effekte zentraler Komponenten der professionellen Handlungskompetenz von Lehrkräften, wie sie im vielzitierten Modell von Baumert und Kunter (2006) beschrieben sind. Baumert und Kunter (2006, S. 505) heben neben den Formen des Lehrwissens (fachliches, fachdidaktisches, pädagogisches Wissen), den Lehrerüberzeugungen und den selbstregulativen Fähigkeiten auch die motivationale Orientierung von Lehrkräften als „individuelle Voraussetzungen für professionelles Lehrerhandeln“ hervor. Die Frage, welche Merkmale erfolgreiche Schulen (Kontextmerkmale) ausmachen, wird in zahlreichen School-Effectiveness-Studien zu beantworten versucht. Scheerens (2016, S. 109) argumentiert, es bestehe weitgehende Übereinstimmung, dass neben einer Reihe anderer Faktoren (wie die Leistungserwartungen an die Schüler/innen) auch die Lehrerkooperation (im Sinne von Professional Learning Communities) einen zentralen Faktor effektiver Schulen darstellt. Im vorliegenden Beitrag werden

daher der Lehrerenthusiasmus und die Lehrerkooperation, zwei Merkmale, die den eben dargestellten Ausführungen nach als förderliche Bedingungen professionellen Lehrerhandelns (insbesondere mit Bezug auf die Unterrichtsvorbereitung, -durchführung und -nachbereitung) angesehen werden können, untersucht.

Befragt man die Literatur zu den Effekten dieser beiden Bedingungen, so wird deutlich, dass bereits erste Studien zum Zusammenhang zwischen diesen Merkmalen und Merkmalen der Unterrichtsqualität (z. B. Klassenführung, kognitive Aktivierung, Lernunterstützung) vorliegen (Lehrerenthusiasmus: z. B. Frenzel, Goetz, Lüdtke, Pekrun & Sutton, 2009; Han, Yin & Wang, 2015; Holzberger, Philipp & Kunter, 2016; Kunter et al., 2008; Kunter et al., 2013; Schiefele & Schaffner, 2015; Warwas & Helm, 2017; Lehrerkooperation: z. B. Lomos, Hofman & Bosker, 2011; Louis & Marks, 1998; Warwas & Helm, 2018). Dagegen sind nur wenige Studien bekannt, die zusätzlich auch Schülerleistungen in einer bestimmten Domäne (z. B. dem Lesen) in den Blick nehmen. Für Österreich fehlen repräsentative Studien zu den skizzierten Zusammenhängen gänzlich. Die vorliegende Untersuchung geht daher auf Basis der österreichischen PIRLS-2016-Daten folgenden Forschungsfragen nach:

FF1. Sind Lehrerenthusiasmus und Lehrerkooperation miteinander einhergehende Lehrerprofessionsmerkmale?

- FF2. Sind der Lehrerenthusiasmus und die Lehrerkooperation mit der Unterrichtsqualität<sup>1</sup> assoziiert?
- FF3. Sind Lehrerenthusiasmus und Lehrerkooperation mit der Leseleistung der Schüler/innen assoziiert?
- FF4. Ist die Unterrichtsqualität mit der Leseleistung der Schüler/innen assoziiert?

## Theorie und Forschungsstand

### FF1: Zum Zusammenhang zwischen Lehrerenthusiasmus und Lehrerkooperation

Grundsätzlich ist davon auszugehen, dass Lehrerenthusiasmus positiv mit Lehrerkooperation korreliert. Eine für deutsche Sekundarstufe-I-Lehrkräfte repräsentative Studie (Richter & Pant, 2016, S. 24) zeigt, dass hoch enthusiastische Lehrkräfte intensiver kooperieren als niedrig und durchschnittlich enthusiastische Lehrkräfte. Auch tauschen enthusiastische Lehrkräfte deutlich häufiger Materialien mit Fachkolleginnen und Fachkollegen aus. Holzberger et al. (2016) konnten zudem zeigen, dass enthusiastische Lehrpersonen verstärkt die Kooperation mit Kolleginnen und Kollegen suchen, um diese Form der Lerngelegenheiten zu nutzen. Es wird daher auch für die vorliegende Studie von einem positiven Zusammenhang zwischen diesen beiden Professionsmerkmalen ausgegangen.

### FF2: Lehrerenthusiasmus und Lehrerkooperation als Prädiktoren der Unterrichtsqualität

#### Lehrerenthusiasmus

Nach Kunter, Frenzel, Nagy, Baumert und Pekrun (2011, S. 289) gibt es für den Begriff Enthusiasmus in der Psychologie *keine allgemein akzeptierte Definition*. Eine Forschungstradition, die auf Lehrerenthusiasmus fokussiert, ist jene zum Lehrerengagement (z. B. Bakker, 2005; Bakker & Bal, 2010; Han, Yin & Wang, 2016; Klassen, Perry & Frenzel, 2012; Klassen et al., 2012; Perera, Vosicka, Granziera & McIlveen, 2018). In dieser Forschungstradition wird Lehrerenthusiasmus über die Utrecht Work Engagement Scale (UWES) (Schaufeli, Bakker & Salanova, 2006) als eine von drei Facetten des Lehrerengagements konzeptualisiert: „Work engagement is defined as a positive, fulfilling work-related state of mind that is characterized by vigor, dedication, and absorption“ (Schaufeli et al. 2006, S. 702). Dabei beschreibt „dedication“ (dt. Einsatz, Hingabe, Enthusiasmus) das emotionale Lehrerengagement bzw. den Lehrerenthusiasmus: Personen,

die eine hohe „dedication“ aufweisen, fühlen sich enthusiastisch und stolz in Bezug auf ihre Arbeit (Schaufeli & Bakker, 2004, S. 6). Auch Kunter et al. (2011, 289 f.) beschreiben Lehrerenthusiasmus als ein Merkmal effektiver Lehrer/innen, das in Verhaltensweisen wie dem Hervorheben der Bedeutung des Unterrichtsstoffs und der eigenen Interessen am Unterrichtsfach oder in einem lebendigen, inspirierenden Unterrichtsstil zum Ausdruck kommt. Kunter et al. (2011) argumentieren, dass Lehrerenthusiasmus vieles mit intrinsischer Lehrermotivation gemeinsam hat (z. B. Freude, Begeisterung, Bedeutsamkeit, Flow), sodass Lehrerenthusiasmus aus Perspektive der Motivationstheorien (insbesondere der Selbstbestimmungstheorie) betrachtet werden kann. Tatsächlich existieren zahlreiche internationale Studien (z. B. Bakker, 2005; Bakker & Bal, 2010; Han et al., 2016; Klassen, Aldhafi et al., 2012; Klassen, Perry et al., 2012; Perera et al., 2018), die die UWE-Skala (und insbesondere die „dedication“-Subdimension), aber auch andere Operationalisierungen (z. B. Unterrichts- und Fachenthusiasmus; Holzberger et al., 2016; Kunter et al., 2011) in der Selbstbestimmungstheorie verorten. Im Folgenden werden daher über die Befunde zur UWE-Skala hinaus auch Befunde der Lehrermotivationsforschung dargestellt.

Die Erforschung der Effekte des Lehrerenthusiasmus im Sinne des (emotionalen) Lehrerengagements bzw. der selbstbestimmten/intrinsischen Lehrermotivation auf Unterrichtsmerkmale (z. B. Autonomieunterstützung) und Merkmale des Schülerlernens (z. B. Lernmotivation, Lernerfolg) stellt eine noch *junge Domäne* dar. Erst nach der Jahrtausendwende sind vermehrte Bemühungen beobachtbar (z. B. Bakker, 2005; Butler & Shibaz, 2008; Frenzel et al., 2009; Klassen et al., 2012; Kunter, 2011; Pelletier, Séguin-Lévesque & Legault, 2002; Tschannen-Moran & Hoy, 2007). Allerdings stellen jene Studien, die die UWE-Skala verwenden, primär Validierungsstudien dar (Klassen, Adhafi et al., 2012; Klassen, Perry et al., 2012; Perera et al., 2018; Schaufeli et al., 2006; Skaalvik & Skaalvik, 2014) oder belegen positive Zusammenhänge des UWES-Scores mit Lehrermerkmalen wie der Berufszufriedenheit (Klassen et al., 2012; Perera et al. 2018; Skaalvik & Skaalvik 2014) und dem selbsteingeschätzten Berufserfolg von Lehrkräften<sup>2</sup> (Bakker & Bal, 2010). Dagegen fehlen Studien, die den Zusammenhang zwischen dem mit der UWES erfassten Lehrerenthusiasmus und der Unterrichtsqualität sowie den Schülerleistungen analysieren. Die vorliegende Untersuchung trägt daher (in Bezug auf Leseleistungen) zur Schließung dieser Forschungslücke bei.

Dass *Lehrerenthusiasmus* positiv mit Unterrichtsqualität assoziiert ist, zeigen zahlreiche Studien, die Lehrerenthusias-

1 Im vorliegenden Beitrag wird die Unterrichtsqualität aufgrund der Datenlage (siehe Abschnitt „Unterrichtsqualität“) ausschließlich über die Qualitätsdimension (Lernunterstützung) erfasst. Die Dimensionen Klassenführung und kognitive Aktivierung bleiben dagegen unberücksichtigt.

2 Die Performance von Lehrkräften wurde über Selbsteinschätzung zu Items erfasst, die danach fragen, in welchem Ausmaß vergangene Woche berufliche Pflicht- und Zusatzanforderungen erfüllt werden konnten.

mus mit anderen Skalen erfassen. So verweisen mehrere Studien zum „Enthusiastic teaching“ (z. B. Frenzel et al., 2009; Keller, Goetz, Becker, Morger & Hensley, 2014; Warwas & Helm, 2017) darauf, dass Lehrer/innen, die positive Emotionen (z. B. Freude, Glück, Interesse) beim Unterrichten empfinden, ein stärker inspirierendes, animierendes und dynamisches Lehrerverhalten, das ebendiese Gefühle ausdrückt, an den Tag legen. Darüber hinaus zeigen die Studien, dass die Schüler/innen diese Lehreremotionen wahrnehmen und folglich den Unterricht als lernförderlicher beschreiben. Sehr ähnlich verhält es sich mit der Vermittlung der Wertigkeit eines Unterrichtsfachs (Keller et al., 2014): Lehrer/innen, die ihrem Unterrichtsfach eine höhere persönliche Bedeutung beimessen, zeigen dies auch in ihrem Unterrichtsverhalten, was wiederum von den Schülerinnen und Schülern als Vorbildwirkung wahrgenommen wird. Darüber hinaus zeigen Warwas und Helm (2017), dass Lehrerenthusiasmus mit höherer Klarheit von Lehrerklärungen, Lernunterstützung sowie Anwendungsorientierung im Unterricht einhergeht. Diese Unterrichtsmerkmale erweisen sich zudem prädiktiv für die Schülerfreude im Unterricht. Neben dem „Enthusiastic teaching“ wurde auch der Einfluss des Lehrerenthusiasmus auf die Basisdimensionen der Unterrichtsqualität untersucht. Im Rahmen der COACTIV-Studie konnten bspw. Kunter et al. (2008) und Kunter (2011) zeigen, dass der Lehrerenthusiasmus am Unterricht, nicht jedoch der Fachenthusiasmus für die Unterrichtsqualitätsdimensionen (kognitive Aktivierung, Klassenführung, konstruktive Unterstützung) prädiktiv ist. Auch Holzberger et al. (2016) bestätigen im Rahmen der BilWiss-Studie den positiven Einfluss von Unterrichts-enthusiasmus auf die Basisdimensionen der Unterrichtsqualität. Darüber hinaus zeigen die Analysen, dass enthusiastische Lehrpersonen sich mehr Mühe bei der Erledigung zentraler Aufgaben im Lehrerberuf (z. B. Unterrichtsvor- und -nachbereitung, Unterrichten, Erziehen, Innovieren, Beurteilen) geben, was wiederum mit den Basisdimensionen der Unterrichtsqualität assoziiert ist (ebd.). Zusammenfassend, so Kunter (2011) und Holzberger et al. (2016), ist davon auszugehen, dass höherer Lehrerenthusiasmus mit mehr Lehreranstrengung/-engagement in Aufgaben, die mit dem Lehrberuf verbunden sind, einhergeht, was sich wiederum positiv auf die Unterrichtsqualität niederschlägt.

In ähnlicher Weise argumentieren auch Roth, Assor, Kanat-Maymon und Kaplan (2007) vor dem Hintergrund der Selbstbestimmungstheorie (Deci & Ryan, 1985), dass *intrinsische Lehrermotivation* mit höherem Lehrerengagement einhergeht, was sich wiederum positiv auf die Unterrichtsqualität auswirkt. Konkret nennen Roth et al. (2007) drei Prozesse, die den Zusammenhang erläutern: (1) Intrinsisch motivierte Lehrkräfte verfügen aufgrund ihres höheren Weiterbildungsengagements über ein höheres fachdidaktisches Wissen, das als Voraussetzung für guten Unterricht gilt. (2) Zudem haben sie selbst gute Erfahrungen mit selbstbestimmtem Lernen gemacht und wollen diese auch ihren Schülerinnen und Schülern ermöglichen. (3) Intrinsisch motivierte Lehrkräfte be-

sitzen eine höhere Widerstandskraft gegenüber Druck von außen, wodurch sie mehr Kapazitäten für die Schaffung von Unterrichtsqualität frei haben. Der Forschungsstand stützt diese Annahmen indirekt: Intrinsisch/extrinsisch motivierte, stark/wenig selbstbestimmte Lehrpersonen unterstützen die Basic Needs ihrer Schüler/innen deutlich stärker/weniger (Müller, Hanfstingl & Andreitz, 2009; Pelletier et al., 2002; Roth et al., 2007; Warwas & Helm, 2017). Schließlich findet sich auch bei Schiefele und Schaffner (2015, S. 160) die Erklärung, dass didaktisch interessierte Lehrkräfte qualitativ hochwertigere Instruktionspraktiken anwenden, weil diese Lehrpersonen mehr Energie und Engagement in Fortbildung und Unterrichtsvorbereitung bzw. den Einsatz bestimmter Unterrichtsstrategien stecken: „interest in teaching methods enhances the motivation to learn more about efficient instructional practices and, thus, increases the use of mastery oriented and cognitively activating practices“ (Schiefele & Schaffner, 2015, S. 160).

Da die vorliegende Studie auf die Lernergebnisse im Sinne der Leseleistungen der Schüler/innen fokussiert, erscheinen insbesondere jene Studien interessant, die den *Zusammenhang zwischen dem Lehrerenthusiasmus/der intrinsischen Lehrermotivation und der kognitiven Aktivierung* im Unterricht – dem Unterrichtsmerkmal, welches das Schülerverständnis zum Ziel hat – untersuchen. Hierzu liegt bereits eine Reihe von Studien vor (z. B. Hein et al., 2012; Holzberger et al., 2016; Keller, Neumann & Fischer, 2017; Kunter et al., 2008; Kunter et al., 2013; Roth et al., 2007; Schiefele & Schaffner, 2015). Die Befundlage ist allerdings sehr heterogen; knapp die Hälfte aller identifizierten Korrelationen ist statistisch nicht signifikant. Dort, wo sie signifikant sind (Kunter, 2011; Kunter et al., 2013; Schiefele & Schaffner, 2015), fallen sie der Höhe nach eher gering bzw. moderat aus.

#### Lehrerkooperation

Die Forschung zur Lehrerkooperation wird insbesondere vor dem Hintergrund des Konzepts der *Professionellen Lerngemeinschaften* (PLGs) betrieben. PLGs beschreiben Gruppen von Lehrpersonen, die ihre (Unterrichts-)Praxis teilen und fortlaufend kritisch hinterfragen, um sich selbst und die eigenen Fähigkeiten weiterzuentwickeln (Rone, 2009, S. 4). Insbesondere für die Unterrichtsentwicklung spielen PLGs eine zentrale Rolle (Bonsen & Hübner, 2012). Die Arbeit in PLGs wird häufig über verschiedene Dimensionen beschrieben (für einen Überblick siehe Warwas & Helm, 2018). Lavié (2006) und Warwas, Helm und Schadt (2019) unterscheiden eine aktionale, eine ideelle und eine strukturelle Dimension:

- Die aktionale Dimension beschreibt die kooperative Entwicklung, d. h. alle gemeinschaftlichen Anstrengungen zur Entwicklung der professionellen Kompetenzen, die wiederum positiv auf das Schülerlernen wirken sollen. Hierunter fallen bspw. der kollegiale Austausch und Hospitationen.

- Die ideelle Dimension beschreibt die Entwicklung eines normativen Konsenses, der sich in geteilten pädagogischen Überzeugungen, Werthaltungen, Einstellungen widerspiegelt und sich in gemeinsamen Entwicklungszielen manifestiert.
- Die strukturelle Dimension meint die Infrastruktur, die kooperative Arbeitsprozesse unterstützt, z. B. klare Regeln und Zuständigkeiten.

Die in PIRLS 2016 eingesetzten Lehrerfragen beziehen sich ausschließlich auf die aktionale Dimension, was die Überprüfung der *Effekte der Lehrerverkooperation* nur bedingt ermöglicht, da Warwas und Helm (2018) zeigen konnten, dass insbesondere PLGs, die in allen drei Dimensionen eine hohe Ausprägung besitzen, auch mit höherer Unterrichtsqualität assoziiert sind. Grundsätzlich ist die Befundlage allerdings eher dünn, da nur wenige Studien existieren, die dieser Frage auf Basis von repräsentativen Stichproben bzw. Large-Scale-Daten nachgehen (bspw. Lee & Smith, 1996; Louis & Marks, 1998; Supovitz, 2002). Lehrpersonen, die stärker in Professionelle Lerngemeinschaften eingebunden sind, unterscheiden sich nicht nur in ihrer Unterrichtsplanung von anderen Lehrkräften (Supovitz, 2002), sondern nach Louis und Marks (1998) auch in ihrem Unterrichtsverhalten. Dabei sind sie authentischer, geben mehr soziale Unterstützung, fördern stärker tieferes Verständnis, indem u. a. lebensnahe Problemstellungen eingesetzt werden. Zudem wird ihr Unterricht als wertschätzender und freundlicher beschrieben. Auch Vescio, Ross und Adams (2008) kamen in ihrer Literaturreview zu dem Ergebnis, dass Mitglieder von PLGs schülerzentrierter unterrichten. Nach Penner-Williams, Díaz und Gonzales Worthen (2017) tragen PLGs zudem zu kulturell und sprachlich sensiblerem Lehrerverhalten bei. So stützen PLG-Lehrpersonen stärker gemeinschaftliches Lernen, gehen stärker auf das Hintergrundwissen von Lernenden ein und verwenden mehr Scaffolding-Techniken (z. B. Hilfestellungen, um die aktuelle Lernaufgabe mit dem Vorwissen zu verbinden). Darüber hinaus würden laut Fulton und Britton (2011) Lehrpersonen, die sich in PLGs engagieren, auch vermehrt forschungsbasierte Unterrichtsmethoden, wie z. B. Schülerexperimente, anwenden und Lernende zur Verwendung von Problemlösestrategien ermuntern. Zudem sind sie stärker um gründliches Verstehen bemüht. Warwas und Helm (2018) erweitern diese Forschungslage um den Befund, dass der Unterricht von Lehrpersonen, die intensiver kooperieren, als anwendungsorientierter und praxisnah erlebt wird. Darüber hinaus werden häufiger kooperative und aktivierende sowie problemorientierte und simulationsbasierte Lernaufgaben bzw. Unterrichtsmethoden eingesetzt. Es ist anzunehmen, dass dieser Unterricht über ein höheres kognitives Aktivierungspotenzial verfügt.

Diese internationale Befundlage zu den Effekten des Lehrerenthusiasmus und der Lehrerverkooperation auf die Unterrichtsqualität ist ermutigend. Der skizzierte Forschungsstand zeigt, dass bereits zahlreiche Studien vorliegen, die den Zusammen-

hang zwischen den Lehrerverprofessionen „Lehrerenthusiasmus“ und „Lehrerverkooperation“ und den Dimensionen der Unterrichtsqualität untersuchten. Die vorliegende Studie schließt an diese Arbeiten an. Da kaum Befunde aus Österreich vorliegen, verspricht die Analyse der österreichischen PIRLS-2016-Erhebung wichtige, weil bislang fehlende Erkenntnisse.

### **FF3: Lehrerenthusiasmus und Lehrerverkooperation als Prädiktoren der Schülerleistung**

Den Zusammenhang von Lehrerenthusiasmus mit Schüleroutcomes betreffend, beschränken sich die weiter oben berichteten Studien primär auf die erlebte Freude von Schülerinnen und Schülern im Unterricht, während fachliche Schülerleistungen häufig ausgeklammert bleiben. Zwei Ausnahmen bilden Kunter (2011) und Kunter et al. (2013). Sie zeigen im Rahmen der COACTIV-Studie, dass Unterrichtsenthusiasmus (nicht aber Fachenthusiasmus) die Mathematikleistungen in der zehnten Klasse positiv vorhersagt. Während Studien zur Vorhersage der Schülerleistung (insbesondere der Leseleistung) durch den Lehrerenthusiasmus rar sind, liegt wie bereits berichtet eine Reihe von Studien vor, die moderat positive Zusammenhänge zwischen dem Lehrerenthusiasmus/der intrinsischen Lehrermotivation und der kognitiven Aktivierung im Unterricht belegen (Kunter, 2011; Kunter et al., 2013; Schiefele & Schaffner, 2015). Es ist daher davon auszugehen, dass der Lehrerenthusiasmus positiv, aber schwach im Zusammenhang mit den Leseleistungen der Schüler/innen steht.

Den Zusammenhang zwischen Lehrerverkooperation und Schülerleistungen betreffend, weisen Literaturreviews und Metaanalysen darauf hin, dass sich leistungsstarke Schulen durch komplexe Formen der Lehrerverkooperation wie PLGs auszeichnen. So berichten Fulton und Britton (2011) in ihrem Literaturreview von sechs Studien, die positive Zusammenhänge zwischen der Lehrerverkooperation und den Schülerleistungen in Mathematik belegen konnten. Lomos et al. (2011) stützen ihre Metaanalyse auf fünf Studien, die zusammengefasst einen kleinen aber signifikanten Effekt ( $d = .25$ ) der PLGs auf die Schülerleistungen nahelegen. Vor dem Hintergrund dieser Befunde wird ein schwacher, positiver Zusammenhang zwischen Lehrerverkooperation und den Leseleistungen erwartet.

### **FF4: Unterrichtsqualität und Schülerleistungen in der Domäne Lesen**

Unterrichtsqualität wird in aktuellen Studien häufig über die drei Basisdimensionen Klassenführung, kognitive Aktivierung und Lernunterstützung operationalisiert. Dies gilt für Videostudien (z. B. Praetorius, Klieme, Herbert & Pinger, 2018) gleichermaßen wie für Fragebogenstudien (z. B. Baumert et al., 2010; Fauth, Decristan, Rieser, Klieme & Büttner, 2014). Traditionell sind die Domänen Mathematik und Naturwissenschaft besonders intensiv erforscht.

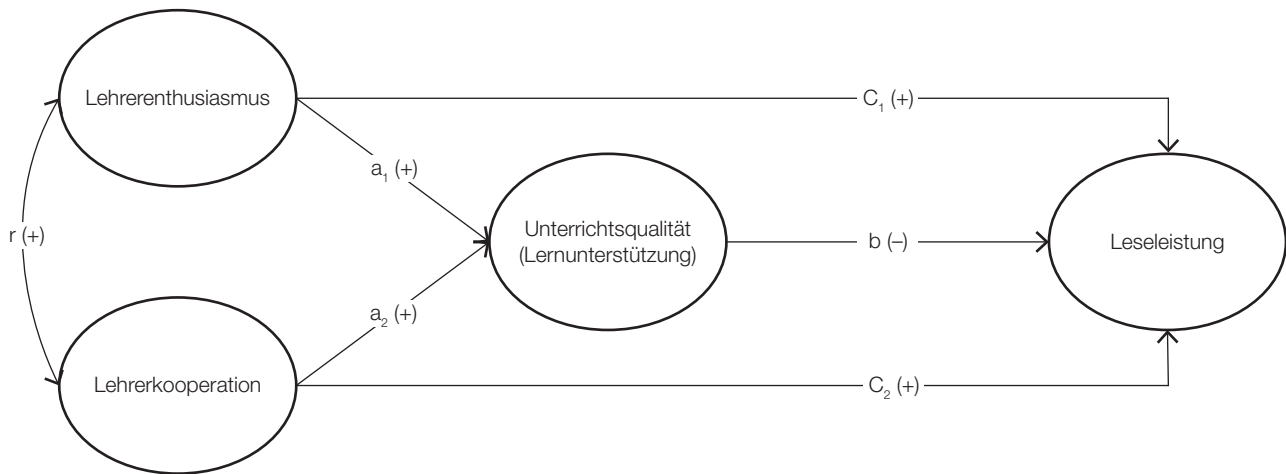


Abbildung 1: Theoriemodell zu den Zusammenhängen zwischen den Kontextmerkmalen und den Leseleistungen auf Klassenebene

Allerdings liegen auch für die Domäne Lesen bereits Studien zum Einfluss der Unterrichtsqualität vor. Eine aktuelle Review (Park, Brownell, Bettini & Benedict, 2017) von 28 Unterrichtsbeobachtungsstudien bestätigt die positiven Effekte der drei Basisdimensionen für die Domäne Lesen. Darüber hinaus wurden auf Basis der PIRLS-2011-Daten positive Zusammenhänge zwischen der mittels Schülerfragebogen erfassten Unterrichtsqualität und der Leseleistung über die teilnehmenden Länder hinweg beobachtet – allerdings nicht für alle Länder, so auch nicht für Österreich (Nortvedt, Gustafsson & Lehre, 2016, S. 107).<sup>3</sup> Bereits im Rahmen der PIRLS-2006-Untersuchung konnten für Österreich keine Zusammenhänge zwischen Merkmalen des Leseunterrichts (Häufigkeit und Dauer des Leseunterrichts, Einsatz von Lesestrategien und -methoden, Verständnisüberprüfung etc.) und den Leseleistungen der Schüler/innen beobachtet werden (Suchaň & Breitfuß-Muhr, 2009).<sup>4</sup> Zudem geben Large-Scale-Untersuchungen wie die BIJU-Studie des Max-Planck-Instituts oder die Bildungsstandardüberprüfung in Österreich Hinweise darauf, dass Merkmale fördernd-unterstützenden, schülerzentrierten Unterrichts eher von Schülerinnen und Schülern in leistungsschwächeren Klassen wahrgenommen werden (z. B. Gruehn, 2000; Helm, Kemethofer, Moosbrugger, Bröderbauer & Luthe, 2016; Klieme & Kuger, 2016; Schreiner et al., in Druck). Eine mögliche Erklärung kann darin liegen, dass „derartige Unterrichtsformen unter Betonung ihrer remedialen Funktion speziell in problematischen, lernschwachen Klassen zum Einsatz kommen“ (Clausen, 2002, S. 52). Ob positive oder

negative Zusammenhänge zwischen den Schülerleistungen und verschiedenen Unterrichtsmerkmalen zu erwarten sind, hängt damit stark von der inhaltlichen Ausrichtung bzw. Operationalisierung der Unterrichtsmerkmale ab. Die in der vorliegenden Studie verwendeten Items der PIRLS-2016-Erhebung beziehen sich insbesondere auf eine schülerzentrierte Lernunterstützung (siehe Tabelle 1). Auch die faktorenanalytische Modellierung der Unterrichtsqualität (siehe weiter unten) zeigt, dass insbesondere jene Items auf den Faktor „Unterrichtsqualität“ laden, die Ausdruck der Lernunterstützung durch die Lehrperson sind. Auch wenn die Lernunterstützung ein Merkmal qualitativen Unterrichts darstellt (etwa, weil die Lehrperson sich um die Lernentwicklung der einzelnen Schüler/innen bemüht), sind negative Zusammenhänge mit den mittleren Leseleistungen einer Klasse nicht ausgeschlossen und aus den genannten Gründen sogar zu erwarten.

### Hypothesen

Vor dem Hintergrund der dargestellten Argumentationslinien und Befunde werden folgende Hypothesen abgeleitet, die in Abbildung 1 in einem Wirkmodell zusammengefasst sind. Dieses Wirkmodell wird auf Basis der PIRLS-2016-Daten einer empirischen Überprüfung unterzogen.

H1: Lehrerenthusiasmus und Lehrerkooperation (jeweils aus Lehrersicht erfasst) hängen positiv zusammen (r-Pfad in Abbildung 1).

3 In der Studie von Nortvedt et al. (2016) wurde die Unterrichtsqualität ähnlich der vorliegenden Studie über die Items 2–4 in Tabelle 1 operationalisiert. Im Gegensatz zur vorliegenden Studie wurden die aggregierten individuellen Verarbeitungsprozesse der Schüler/innen nicht statistisch kontrolliert, sodass sie mit in das Konstrukt „Unterrichtsqualität“ einfließen.

4 Allerdings wurden die erwähnten Unterrichtsmerkmale auf Basis von Lehrerangaben erfasst. In der vorliegenden Studie werden Schülerurteile verwendet, die als valider gelten (Clausen, 2002; Kunter & Baumert, 2006) und sich stärker auf allgemeine, fachübergreifende Basisdimensionen beziehen.

- H2: Lehrerenthusiasmus und Lehrerkooperation (jeweils aus Lehrersicht erfasst) sagen die Unterrichtsqualität im Sinne einer stärker schülerorientierten Lernunterstützung (aus Schülersicht erfasst) signifikant positiv vorher ( $a_1/a_2$ -Pfade in Abbildung 1).
- H3: Die Unterrichtsqualität im Sinne einer stärker schülerorientierten Lernunterstützung hängt negativ mit der mittleren Leseleistung einer Klasse zusammen. Mit anderen Worten: Schüler/innen aus leseschwachen Klassen nehmen eine höhere Unterrichtsqualität im Sinne einer stärker schülerorientierten Lernunterstützung durch die Lehrperson wahr (b-Pfad in Abbildung 1).
- H4: Lehrerenthusiasmus und Lehrerkooperation (jeweils aus Lehrersicht erfasst) sagen die mittlere Leseleistung der Schüler/innen einer Klasse positiv vorher ( $C_1/C_2$ -Pfade in Abbildung 1).

## Methode

### Stichprobenbeschreibung

Die Hypothesen bzw. die in Abbildung 1 dargestellten Zusammenhänge werden auf Basis der in PIRLS 2016 erhobenen Schüler- und Lehrerdaten empirisch geprüft. 2016 nahmen 4360 Schüler/innen und deren Lehrer/innen ( $N = 256$ ) aus 150 Volksschulen an der PIRLS-Erhebung in Österreich teil. Das durchschnittliche Alter der Schüler/innen betrug 10,3 Jahre; jenes der Lehrpersonen 41,2 Jahre.

### Skalen

Es wurden Items aus dem Lehrer- und Schülerfragebogen herangezogen, die eine Erfassung der interessierenden Konstrukte erlauben. Tabelle 1 und 2 geben einen Überblick über die Konstruktoperationalisierung. Für Informationen zur Erfassung der *Leseleistungen* der Schüler/innen in PIRLS 2016 siehe Widauer und Wallner-Paschon (2017).

Der *Lehrerenthusiasmus* wurde über die Subdimension „dedication“ der UWES (Schaufeli et al., 2006) auf Basis von Lehrerselbsteinschätzungen erfasst. Die Subskala „dedication“ erfasst das emotionale Engagement von Lehrkräften im Sinne des Enthusiasmus. Personen, die eine hohe „dedication“ aufweisen, fühlen sich enthusiastisch und stolz in Bezug auf ihre Arbeit (Schaufeli & Bakker, 2004, S. 6). Mit anderen Worten: Dedication verweist auf eine starke Identifikation mit der Arbeit im Sinne des Erlebens von Bedeutung, Enthusiasmus, Stolz etc. Dedication wurde – wie das Work Engagement insgesamt – als stabiles Konstrukt (trait) in Abgrenzung zum Burnout definiert (Schaufeli et al., 2006,

S. 703). Zur Operationalisierung des Lehrerenthusiasmus wurden folgende drei Items der „dedication“-Subskala verwendet, die sich in Studien (Klassen et al., 2012; Schaufeli et al., 2006) als valide erwiesen haben: „Mein Beruf begeistert mich“, „Meine Arbeit inspiriert mich“, „Ich bin stolz auf meine Arbeit“.

Das Ausmaß an *Lehrerkooperation* wurde über fünf Items erfasst, die Ausdruck der aktionalen Dimension der Arbeit in PLGs (Lavié, 2006; Warwas & Helm, 2018) sind. Die Items geben bspw. Auskunft über die Häufigkeit der Hospitationen in anderen Klassen, um mehr über das Unterrichten zu lernen. Diese Form der Kooperation geht über den bloßen Austausch von bspw. Lernmaterialien und Informationen, wie sie in der ersten Qualitätsstufe nach Gräsel, Stark, Sparka und Herzmann (2007) beschrieben ist, hinaus. Mit den vorliegenden Items ist es allerdings nicht möglich, höhere Qualitätsstufen der Lehrerkooperation im Sinne der Arbeitsteilung und der Kokonstruktion nach Gräsel et al. (2006) abzubilden.

Um die beiden Lehrermerkmale „Lehrerenthusiasmus“ und „Lehrerkooperation“ genauer zu beschreiben, wurde geprüft, inwiefern sie mit dem Lehrergeschlecht, -alter und der Lehrererfahrung zusammenhängen. Die Analysen zeigen, dass das Ausmaß an Lehrerkooperation unabhängig vom Geschlecht ( $.052, p = .647$ ), Alter ( $-.016, p = .850$ ) und der Lehrererfahrung ( $.042, p = .596$ ) ist. Dagegen korreliert der Lehrerenthusiasmus signifikant negativ mit der Lehrererfahrung ( $-.164, p = .016$ ) und tendenziell negativ mit dem Lehreralter ( $-.120, p = .068$ ). Zudem berichten männliche Lehrpersonen über einen signifikant niedrigeren Lehrerenthusiasmus ( $-.146, p = .001$ ).

### Unterrichtsqualität (Lernunterstützung)

Während in PIRLS 2016 Lehrerenthusiasmus, Lehrerkooperation und die Leseleistungen mit explizit dafür entwickelten Messinstrumenten (Mullis & Martin, 2015; Schaufeli et al., 2006; Widauer & Paschon, 2017) erfasst wurden, wurden im Schülerfragebogen<sup>5</sup> der PIRLS-2016-Erhebung keine Items eingesetzt, die explizit – d. h. vor dem Hintergrund einer konkreten Theorie wie z. B. dem Modell der Basisdimensionen der Unterrichtsqualität – für die Erfassung von Unterrichtsmerkmalen konstruiert wurden. Im Folgenden wird daher ausführlich dargestellt, wie die verfügbaren Items im Rahmen der hier durchgeführten Unterrichtsanalysen zum Einsatz kommen.

#### Inhaltliche Skalenstruktur

Für die in PIRLS 2016 eingesetzten Items zur Erfassung von Merkmalen des Leseunterrichts (z. B. Abschnitt „Lesen im

5 Zwar sind auch im Lehrerfragebogen Items zu Unterrichtsmerkmalen enthalten, allerdings gelten Lehrerselbsteinschätzungen für die Erfassung von Unterrichtsmerkmalen als weniger valide (Clausen, 2003; Kunter & Baumert, 2006). Daher finden sie in der vorliegenden Arbeit keine Berücksichtigung.

Unterricht“ des PIRLS-2016-Schülerfragebogens) konnten in der zugänglichen PIRLS Dokumentation keine Hinweise auf zugrunde gelegte Theorien gefunden werden. Kapitel 2 des PIRLS 2016 Assessment Frameworks (Hooper, Mullis & Martin, 2015) bildet zwar das theoretische „Fundament“ für die Konstrukte, die im Rahmen der Kontextfragebögen erhoben wurden. Allerdings sind die Ausführungen relativ abstrakt gehalten, sodass keine explizite Zuordnung von Items oder Skalen zu bestimmten pädagogisch-psychologischen oder fachdidaktischen Theorien möglich ist. Aus diesem Grund wurden für die vorliegenden Analysen Items auf Basis inhaltlicher Überlegungen vor dem Hintergrund aktueller Literatur zu den Basisdimensionen der Unterrichtsqualität (z. B. Klieme, Pauli & Reusser, 2009) ausgewählt. Items I1 und I2 in Tabelle 1 sind Indikatoren der Basisdimension „Motivationale Lernunterstützung“ (ebd.), da sie Ausdruck davon sind, inwiefern es der Lehrperson im Unterricht gelingt, das Interesse der Schüler/innen am Lesen zu unterstützen. Die Items I3 und I4 stellen Indikatoren der Basisdimension „Klassenführung“ dar, weil Klarheit im Unterricht eine präventive Wirkung vor Unterrichtsstörungen zugeschrieben wird (ebd.). Gleichsam ist Klarheit im Unterricht (dies gilt insbesondere für verständliche Lehrerklärungen; Findeisen, 2017) auch eine förderliche Bedingung für die „kognitive Aktivierung im Unterricht“, sodass die Zuordnung hier nicht eindeutig ist. Schließlich stellen I5 bis I9 Indikatoren der Basisdimension „Konstruktive Lernunterstützung“ (Kunter & Voss, 2013) dar. Unter diese Dimension fallen alle Maßnahmen, die das Schülerverstehen und -lernen – nach der Phase der kognitiven Aktivierung im Sinne der Probleminduktion/-einsicht (Kunter & Voss, 2013) – betreffen. Diese Zuordnung der Items zu den Basisdimensionen ist in erster Linie theoretischer Natur. Die empirischen latenten Level-2-Korrelationen zwischen diesen Konstrukten sind derart hoch ausgeprägt ( $r > .90$ ), dass in den folgenden Analysen ein *Globalfaktor* der Unterrichtsqualität modelliert wird. Diese Vorgehensweise deckt sich mit anderen Studien zur Unterrichtsqualität (Morin, Marsh, Nagengast & Scalas, 2014; Wagner, Göllner, Helmke, Trautwein & Lüdtke, 2013) sowie der Validierungsstudie zu den PIRLS-„Unterrichtsqualitätsitems“ von Scherer und Gustafsson (2015), die die Unterrichtsqualität in PIRLS u. a. ebenfalls auf Basis eines Globalfaktors modellieren. An dieser Stelle ist allerdings hervorzuheben, dass die weiter unten erläuterte Modellierung des latenten Faktors „Unterrichtsqualität“ über ein sogenanntes „shared and configural construct“-Modell zeigt, dass insbesondere jene Items aus Tabelle 1 auf den Faktor „Unterrichtsqualität“ laden, die primär Ausdruck von motivationaler Unterstützung einerseits und konstruktiver Unterstützung (im Sinne der Ermutigung, Gelesenes verbal zu reflektieren) andererseits sind (I1: „Meine Lehrerin gibt mir interessante Dinge zu lesen“; I2: „Mich interessiert, was meine Lehrerin sagt“; I5: „Meine Lehrerin ermutigt mich, zu sagen, was ich über das Gelesene denke“; I9: „Meine Lehrerin möchte, dass wir im Unterricht über das reden, was wir gelesen haben“). Dagegen laden die Items zur Klarheit im Unterricht (I3, I4)

nicht signifikant auf den Faktor „Unterrichtsqualität“. Im Folgenden wird daher von Unterrichtsqualität im Sinne der Lernunterstützung gesprochen.

#### Angebots-Nutzungs-Logik

Die Itemformulierung gibt nicht nur Hinweise über die inhaltliche Struktur der Skala und über die inhaltliche Bedeutung des Faktors „Unterrichtsqualität“, sondern – wie im vorliegenden Fall – auch über die Analyseebene (Schüler- oder Klassenebene). Da sich die Itemformulierungen in vielen Fällen nicht ausschließlich auf die Lehrperson oder den Unterricht, sondern auch auf die Schüler/innen selbst beziehen, ist anzunehmen, dass sie nicht nur Informationen über Unterrichtsmerkmale, sondern auch über Merkmale des Schülerlernens widerspiegeln (siehe Tabelle 1). Diese Unterscheidung erinnert an die Angebots-Nutzungs-Logik nach Helmke (2015). Mit „Angebot“ meint Helmke u. a. Unterrichtsqualitätsmerkmale. Mit „Nutzung“ sind die Lernaktivitäten der Schüler/innen, d. h. die Merkmale der individuellen Verarbeitungsprozesse gemeint, die Helmke (2015, S. 80) einerseits in der Wahrnehmung und Interpretation des Unterrichtsangebots, insbesondere des Lehrerverhaltens, und andererseits in den Lern- und Denkprozessen, den Motivationen und Emotionen der Schüler/innen sieht. Diese Unterscheidung zwischen Unterrichtsangebot (im Sinne von Unterrichtsqualitätsmerkmalen) und Unterrichtsnutzung (im Sinne individueller Verarbeitungsprozesse) wird für die vorliegende Untersuchung übernommen. Die im nächsten Abschnitt dargestellten statistischen Analysen erlauben eine Modellierung dieser beiden Aspekte des Lehr-Lern-Prozesses auf Basis der PIRLS-Items (Tabelle 1).

#### Modellierung des Unterrichtsangebots und der Unterrichtsnutzung

Die in PIRLS 2016 gewählte Itemformulierung, die sich teilweise sowohl auf Lehrer- als auch auf Schülerinformationen bezieht (siehe Tabelle 1), macht auch aus forschungsmethodischer Sicht eine Klärung darüber notwendig, welche Informationen die analysierten Items auf den unterschiedlichen Analyseebenen (Klassenebene, Schülerebene) enthalten (Stapleton, Yang & Hancock, 2016). Für die Erfassung von Konstrukten auf Klassenebene (z. B. Unterrichtsmerkmale) durch Schülerratings werden in der Literatur die „shared constructs“ (oder auch reflective, climate construct) und die „configural constructs“ (oder auch formative, contextual construct) unterschieden (Lüdtke et al., 2008; Lüdtke, Marsh, Robitzsch & Trautwein, 2011; Marsh et al., 2012; Stapleton et al., 2016). Beide Konstrukte modellieren die Varianz auf Klassenebene (ICC, Intraklassenkorrelationskoeffizient). Mit anderen Worten: Sie geben Auskunft darüber, wie die Varianz der Schülerantworten zwischen den Klassen interpretiert werden kann. Nach Stapleton et al. (2016) können Klassen in ihrer durchschnittlichen Zustimmung zu einem Item aus mindestens zwei Gründen variieren:

(1) *shared constructs*: Das im Mittel wahrgenommene Lehrerverhalten bzw. *Unterrichtsangebot* unterscheidet sich zwischen den Klassen, da die Lehrpersonen tatsächlich unterschiedlich handeln. In diesem Fall würde der ICC Informationen über die Lehrperson widerspiegeln. Das beurteilte Objekt (hier die Lehrperson) liegt auf Klassenebene (Spalte 2 in Tabelle 1). Beispiele für derartige Konstrukte sind Merkmale des Unterrichts oder der Lehrperson, die per Definition nur auf Klassenebene existieren und für alle Schüler/innen objektiv gesehen gleich stark ausgeprägt sind (z. B. die Qualität von Lehrerklärungen im Rahmen von Plenumsvorträgen).

(2) *configural constructs*: Die im Klassendurchschnitt berichteten individuellen Verarbeitungsprozesse bzw. die individuelle *Nutzung des Unterrichtsangebots* variieren zwischen den Klassen, da – wie in der vorliegenden Studie aufgrund der Itemformulierungen zu vermuten ist – die Schüler/innen verschiedener Klassen ...

- a. im Klassenmittel tatsächlich unterschiedlich stark am Unterricht interessiert sind (I1, I2),
- b. die Lehrperson im Klassenmittel unterschiedlich gut verstehen können (I3, I4) und/oder
- c. im Klassenmittel unterschiedlich viel Lehrerunterstützung bekommen/wahrnehmen (I5–I9).

Das beurteilte Objekt (hier die Schülerin oder der Schüler bzw. das schülerspezifische Lehrerverhalten) liegt auf Schüler-ebene (Spalte 3 in Tabelle 1). Diese Schülerratings werden nicht nur auf Schüler-ebene analysiert, sondern auch auf Klassen-ebene. Indem die Schülereinschätzungen hochaggregiert werden, können Unterschiede in der klassendurchschnittlichen Nutzung des Unterrichtsangebots analysiert werden.

Beide Konstrukte (shared und configural) stellen Kontextmerkmale dar, die das Lernen allgemein und die Leseleistungen der Schüler/innen in PIRLS 2016 im Besonderen beeinflussen können.

Aufgrund der PIRLS-2016-Itemformulierungen ist anzunehmen, dass der ICC eine Mischung beider Konstrukte widerspiegelt, die es in statistischen Analysen zu trennen gilt (siehe den Methodenteil). Die Reliabilitäten der Konstrukte auf Klassenebene (ICC[2]) liegen bei .56 (shared construct) und .60 (configural construct) und damit unter dem Cut Off von .70. Die niedrigen Reliabilitäten sind aufgrund der niedrigen ICC(1)-Werte der Items (siehe Tabelle 2) zu erwarten. Da die Modelle für die Reliabilität sowohl Mess- als auch Stichprobenfehler kontrollieren (Lüdtke et al. 2008), ist nicht zu erwarten, dass die im Folgenden berichteten Parameterschätzungen der Kontexteffekte durch die niedrigen ICC(2)-Werte verzerrt sind.

### Auswertungsstrategie

Die Effekte des Lehrereнтуusiasmus und der Lehrerkoope-ration auf die Schülerleistungen im Leseunterricht werden

im Rahmen eines Mehrebenen-Strukturgleichungsmodells (z. B. Lüdtke et al., 2011) geprüft, wobei die Schüler- und Klassenebene modelliert werden. Die Auswertungen werden in Mplus 8 (Muthén & Muthén, 1998–2018) unter Verwendung der nötigen Spezifikationen für den korrekten Umgang mit komplexen Daten vorgenommen:

- *Berücksichtigung der Mehrebenenstruktur*: Um die hierarchisch geschichteten Daten (Schüler/innen in Klassen, Klassen in Schulen) im Rahmen der Auswertung zu berücksichtigen, wurde die Analyse-Methode TWOLEVEL COMPLEX in Mplus verwendet, wobei die Lehrpersonen (IDTEALIN) als Cluster und die Schulen (ID-SCHOOL) als Stratifikationsmerkmal eingesetzt wurden (Foy, 2018).
- *Modellierung des Unterrichtsangebots und der Unterrichts-nutzung*: Wie oben angeführt, spiegeln die untersuchten Items (Tabelle 1) unterschiedliche Informationen auf Individual- und Gruppenebene (individuelle Verarbeitungsprozesse der Schüler/innen) und Gruppenebene (Unterrichtsqualität im Sinne der Lernunterstützung, Unterrichts-nutzung im Sinne der auf Klassenebene aggregierten Verarbeitungsprozesse) wider. Um die Informationen über die Unterrichtsqualität (= Angebot) und die aggregierten Verarbeitungsprozesse (= Nutzung) zu extrahieren, wurden sogenannte „shared and configural constructs“ nach Stapleton et al. (2016) spezifiziert. Dabei wird wie bei traditionellen Multilevel latent confirmatory factor analyses (z. B. Lüdtke et al., 2011) die Varianz der Items auf die Individual- und Gruppenebene aufgeteilt. Auf Individual-ebene wird dann ein Messmodell für das Konstrukt „individuelle Verarbeitungsprozesse“ spezifiziert, das auf Gruppenebene gespiegelt wird, um das Messmodell (configural construct) für das Konstrukt „aggregierte Verarbeitungsprozesse“ zu erhalten. Dazu werden die Faktorladungen über die beiden Ebenen hinweg gleichgesetzt. Für das Konstrukt „Unterrichtsqualität“ wird simultan auf Gruppenebene ein Messmodell (shared construct) mit frei geschätzten Faktorladungen spezifiziert (siehe Stapleton et al., 2016).
- *Berücksichtigung der komplexen Stichprobenziehung*: Für PIRLS 2016 wurde ein zweistufiges Verfahren der Stichprobenziehung verwendet, bei dem die Schulen stratifiziert (in Österreich nach den Bundesländern) werden (Hafner, 2017). In einem ersten Schritt werden Schulen „systematisch mit Wahrscheinlichkeiten, die proportional zu ihrer Größe (gemessen an der Schülerzahl in der 4. Schulstufe, Anm.) sind (probability proportional to size – PPS), ausgewählt“ (ebd., S. 25). In einem zweiten Schritt werden zufällig eine oder mehrere Klassen dieser Schulen ausgewählt, deren Schüler/innen die Stichprobe bilden. Um die komplexe Stichprobenziehung bei den Auswertungen zu berücksichtigen, werden Sampling Weights verwendet (George, Oberwimmer & Itzlinger-Bruneforth, 2016; Joncas & Foy, 2012). Für die vorliegenden Analysen wurde die GewichtungsvARIABLE



Item	Lehrerinformation (= Unterrichtsangebot)	Schülerinformation (= Unterrichtsnutzung)
<b>Motivationale Lernunterstützung</b>		
I1: Meine Lehrerin gibt mir interessante Dinge zu lesen.	Die Lehrperson wählt und gibt interessantes Lesematerial.	Die einzelne Schülerin/der einzelne Schüler ist am Lesematerial interessiert.
I2: Mich interessiert, was meine Lehrerin sagt.	Die Lehrperson kommuniziert anregend und interessant (z. B. gibt interessante Lehrerklärungen).	Die einzelne Schülerin/der einzelne Schüler ist am Unterrichtsthema/-fach interessiert.
<b>Förderliche Bedingungen für die Klassenführung und die kognitive Aktivierung</b>		
I3: Ich weiß, was meine Lehrerin von mir erwartet.	Die Lehrperson stellt ihre Erwartungen klar dar.	Die einzelne Schülerin/der einzelne Schüler kann die Erwartungen der Lehrperson an sie/ihn erkennen.
I4: Ich kann meine Lehrerin leicht verstehen.	Die Lehrperson kann gut erklären.	Die einzelne Schülerin/der einzelne Schüler in kann die Lehrperson leicht verstehen.
<b>Konstruktive Lernunterstützung</b>		
I5: Meine Lehrerin ermutigt mich, zu sagen, was ich über das Gelesene denke.	Die Lehrperson ermutigt zum Nachdenken über das Gelesene.	Die einzelne Schülerin/der einzelne Schüler fühlt sich ermutigt und motiviert, über das Gelesene zu sprechen.
I6: Meine Lehrerin lässt mich zeigen, was ich gelernt habe.	Die Lehrperson lässt Schüler/innen zeigen, was sie gelernt haben.	Die einzelne Schülerin/der einzelne Schüler kann im Unterricht zeigen, was sie/er gelernt hat.
I7: Meine Lehrerin macht Verschiedenes, um uns beim Lernen zu helfen.	Die Lehrperson macht Verschiedenes, um Schülerinnen und Schülern beim Lernen zu helfen.	(keine Schülerinformation)
I8: Meine Lehrerin hilft mir, Fehler auszubessern.	Die Lehrperson hilft Schülerinnen und Schülern, Fehler auszubessern.	Die einzelne Schülerin/der einzelne Schüler nimmt Unterstützung der Lehrperson wahr, die ihr/ihm hilft, Fehler auszubessern.
I9: Meine Lehrerin möchte, dass wir im Unterricht über das reden, was wir gelesen haben.	Die Lehrperson möchte, dass Schüler/innen im Unterricht über das reden, was sie gelesen haben.	(keine Schülerinformation)
<i>Anmerkung: I = Item.</i>		

Table 1: Items zum Leseunterricht

Dimension	#	Beispielitem*	$\alpha$	$\omega_{L2}$	M	SD
Lehrerenthusiasmus (LF)	3	Mein Beruf begeistert mich.	.85	.86	2.46	.55
Lehrerkooperation (LF)	5	Hospitationen in anderen Klassen, um mehr über das Unterrichten zu lernen.	.73	.74	2.56	.46
UA: Unterrichtsqualität i. S. d. LU/ UN: Verarbeitungsprozesse (SF)	9	Meine Lehrerin gibt mir interessante Dinge zu lesen.	.82	.92	2.44	.56
Leseleistung (SF)		Siehe Widauer und Wallner-Paschon (2017)	-	-	-	-
<p><i>Anmerkungen:</i> LF = Lehrerfragebogen, SF = Schülerfragebogen, UA = Unterrichtsangebot, UN = Unterrichtsnutzung, LU = Lernunterstützung. Das Antwortformat aller Items reicht von 1 („stimme völlig zu“) bis 4 („stimme überhaupt nicht zu“). Zur leichteren Interpretation wurden die Variablen für die Zusammenhangsanalysen umcodiert, sodass höhere Werte auch eine höhere Konstruktausprägung implizieren. Die Items für die Konstrukte „Unterrichtsmerkmale“ und „aggregierte Verarbeitungsprozesse“ variieren auf Klassenebene (ICC[1]) zwischen 4,4 % (I4) und 7,5 % (I7). <math>\alpha</math> = Cronbachs Alpha, <math>\omega_{L2}</math> = Omega-Reliabilität auf Level 2 nach Geldhof, Preacher und Zyphur (2014), M = Mittelwert, SD = Standardabweichung.</p> <p>* Zur vollständigen Auflistung der Items siehe Anhang 1 &amp; 2.</p>						

Table 2: Operationalisierung der untersuchten Konstrukte

TCHWGT verwendet. TCHWGT gewährleistet korrekte Resultate für Analysen auf Basis von verlinkten Lehrer- und Schülerdatensets (Foy, 2018, S. 72), wie sie in der vorliegenden Studie zum Einsatz kommen.

Der Einsatz von Resampling-Methoden (Replicate Weights) ist im vorliegenden Fall nicht nötig, da Mplus 8 bei Spezifikation der Mehrebenenstruktur das mehrstufige Stichprobendesign explizit berücksichtigt und nicht verzerrte Schätzer der Stichprobenvarianz liefert (George, Oberwimmer & Itzlinger-Bruneforth, 2016).

- *Verwendung von Plausible Values:* Die Leseleistung wird aufgrund der Annahme, dass diese Werte in Abhängigkeit von Schülertestdaten (Item-Responses) und anderen Schülerdaten (Kovariaten) bedingt verteilt sind, über multiple Imputationen generiert (Foy, Brossman & Galia, 2012; Robitzsch, Pham & Yanagida, 2016). Die so in PIRLS 2016 zur Verfügung gestellten Plausible Values können dann mittels der Pooling-Methode (Rubin, 1987) ausgewertet werden. Dies wurde im vorliegenden Beitrag über den Mplus-Befehl TYPE = IMPUTATION realisiert.
- *Umgang mit fehlenden Werten:* Von den 4360 analysierten Schülerdatensätzen liegen je nach Analyse zwischen 3530 und 3877 vollständige Datensätze vor. Alle verwendeten Variablen mit Ausnahme der Leseleistung besitzen fehlende Werte. Mit 134 (3,1 %) fehlenden Werten stellt das Item „Meine Lehrerin lässt mich zeigen, was ich gelernt habe“ jene Variable dar, für die am häufigsten keine gültigen Werte vorliegen. Robitzsch, Pham und Yanagida (2016) stellen drei Verfahren zum Umgang mit fehlenden Werten in Large-Scale Assessments vor; darunter das Full-Information-Maximum-Likelihood-Verfahren (FIML), das im vorliegenden Beitrag zur Anwendung kommt. „Dabei werden alle vorliegenden Schüler/innen für eine Analyse aufgenommen und die fehlenden Daten werden ‚ausintegriert‘“ (ebd., S. 290 f.).<sup>6</sup> Eine zentrale Annahme des FIML-Verfahrens ist die Missing-at-Random-Annahme, d. h. es wird davon ausgegangen, dass die Werte aller anderen Variablen zufällig fehlen.

Zur Bestimmung des Modellfits werden übliche Cut-Off-Kriterien herangezogen (Hu & Bentler, 1999; Little, 2013): der Bentlers Comparative Fit Index (CFI  $\geq$  .90), der Tucker Lewis Index (TLI  $\geq$  .90) und der Root Mean Square Error of Approximation (RMSEA  $\leq$  .08) sowie der Standardized Root

Mean Square Residual auf Schüler- als auch auf Klassenebene (SRMR<sub>1/2</sub>  $\leq$  .10).

## Ergebnisse

### Deskriptive Befunde

Die latenten Korrelationen der untersuchten Variablen auf Klassenebene sind in Tabelle 3 dargestellt. H1: Erwartungskonform hängen Lehrerenthusiasmus und Lehrerkoope- ration schwach, aber statistisch signifikant positiv zusammen. H2: Entgegen den Erwartungen hängen weder der Lehrerenthusiasmus noch die Lehrerkoope- ration positiv mit der Unterrichtsqualität im Sinne der schülerzentrierten Lernunter- stützung zusammen. Im Gegenteil, der Lehrerenthusiasmus korreliert sogar schwach negativ mit der Unterrichtsqualität. H3: Ebenfalls entgegen den Erwartungen korrelieren beide Professionsmerkmale nicht signifikant mit der mittleren Lese- leistung einer Schulklasse. H4: Wie vermutet wird die Un- terrichtsqualität im Sinne der schülerorientierten Lernunter- stützung verstärkt in leeschwachen Klassen wahrgenommen. Darauf deutet die signifikante, negative Korrelation hin. Der überaus hohe negative Zusammenhang zwischen dem Konst- rukt „Unterrichtsmerkmale“ und den Leseleistungen ist u. a. auf die niedrige Klassenebenenvarianz der Items zum Lese- unterricht zurückzuführen (siehe die ICC(1)-Werte in den Anmerkungen zu Tabelle 2). Aus diesem Grund wird dieser Effekt im Folgenden – wie von Marsh et al. (2012) empfoh- len – an der Gesamtvarianz standardisiert.

### Pfadanalyse mittels Mehrebenen-Struktur- gleichungsmodellierung

Abbildung 2 enthält die Ergebnisse der Mehrebenen- Strukturgleichungsmodellierung. Die Fit-Werte für das in Abbildung 2 geschätzte Modell liegen über bzw. unter den in Abschnitt Auswertungsstrategie dargestellten Cut-off-Kri- terien (#Parameter: 81; CFI: .93; TLI: .92; RMSEA: .03; SRMR<sub>Level1</sub>: .04; SRMR<sub>Level2</sub>: .07), sodass von einer guten Pas- sung zwischen modellimplizierter und empirischer Daten- struktur ausgegangen werden kann.

*H1:* Lehrerenthusiasmus und Lehrerkoope- ration korrelieren signifikant positiv ( $r = .195$ ,  $p = .014$ ).

*H2:* Der Lehrerenthusiasmus, nicht aber die Lehrerkoope- ration, sagt die aus Schülersicht erfasste Unterrichtsqualität

<sup>6</sup> Während beim Verfahren der multiplen Imputation das Imputations- und das Analysemodell unterschiedlich spezifiziert werden, wird beim FIML-Ansatz das Imputationsmodell auf Basis aller in den Analysen verwendeten Variablen spezifiziert, was bei Modellen mit verschiedenen Mengen an Variablen unterschiedliche Imputationsmodelle impliziert und damit die Vergleichbarkeit der Modelle im Hinblick auf die modellimplizierten Mittelwerte erschwert (Robitzsch, Pham & Yanagida, 2016), die allerdings in den vorliegenden Analysen nicht interessieren.

#	Variablen	1	2	3	4	5
1	Lehrerenthusiasmus (LF)	1				
2	Lehrerkooperation (LF)	-.193*	1			
3	UA: Unterrichtsqualität (Lernunterstützung) (SF)	-.193**	-.088	1		
4	UN: aggregierte Verarbeitungsprozesse (SF)	.227*	.063	0 <sup>a</sup>	1	
5	Leseleistung	.027*	.072	-.671**	.119	1

Anmerkungen: N = 246, LF = Lehrerfragebogen, SF = Schülerfragebogen, UA = Unterrichtsangebot, UN = Unterrichtsnutzung; \* = .01 < p < .05, \*\* = p < .01; <sup>a</sup> auf 0 fixiert (siehe Stapleton et al., 2016).

Tabelle 3: Latente Level-2-Korrelationen zwischen den untersuchten Variablen

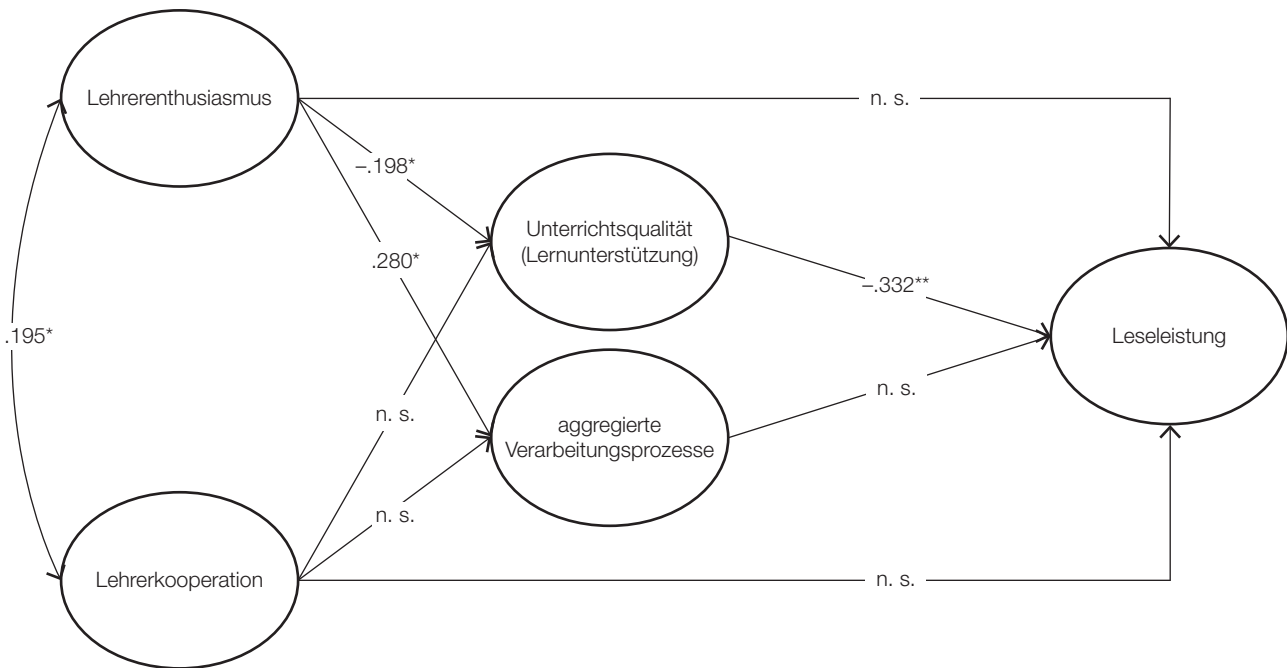


Abbildung 2: Standardisierte Zusammenhänge zwischen den Kontextmerkmalen und den Leseleistungen auf Klassenebene

im Sinne der schülerorientierten Lernunterstützung vorher; allerdings negativ ( $\beta = -.198, p = .033$ ;  $\beta = -.053, p = .659$ ).

H3: Die mittlere Leseleistung einer Klasse wird signifikant negativ durch die aus Schülersicht erfasste Unterrichtsqualität im Sinne der schülerorientierten Lernunterstützung vorhergesagt ( $\beta = -.332, p < .001$ ).

H4: Statistisch signifikante Zusammenhänge zwischen den Lehrerprofessionsmerkmalen (Lehrerenthusiasmus, Kooperation) und den mittleren Leseleistungen einer Klasse sind nicht beobachtbar ( $\beta = -.153, p = .104$ ;  $\beta = .031, p = .775$ ).

Zusammenfassend lassen sich damit lediglich die Hypothesen 1 und 3 bestätigen. Diese Befunde werden im nächsten Ab-

schnitt inhaltlich diskutiert. An dieser Stelle soll auf zwei „Zusatzergebnisse“ der vorgelegten Analysen hingewiesen werden:

(1) *Indirekte Effekte:* Aufgrund der fast ausschließlich nicht signifikanten Zusammenhänge zwischen den erfassten Lehrerprofessionsmerkmalen und allen anderen untersuchten Konstrukten sind keine indirekten Effekte der Lehrerprofessionsmerkmale über die Unterrichtsqualität auf die Leseleistungen zu beobachten. Eine Ausnahme bildet der Lehrerenthusiasmus, der über die Unterrichtsmerkmale schwach positiv mit der Leseleistung zusammenhängt (unstd.  $\beta = 9,417, p = .036$ ). Dabei geht ein höherer Lehrerenthusiasmus mit geringeren Werten in der Unterrichtsqualität (Lernunterstützung) einher, die vor allem in leistungsschwachen Schulklassen höher ausgeprägt ist.

(2) *Unterrichtsnutzung*: Die auf Klassenebene aggregierten Verarbeitungsprozesse waren nicht Gegenstand der a priori aufgestellten Hypothesen, da sie erst im Rahmen der statistischen Modellierung der Unterrichtsqualität mit in die Analysen aufgenommen wurden. Es hätte vermutet werden können, dass ein höheres Ausmaß an individuellen Verarbeitungsprozessen (z. B. das Verstehen der Lehrperson, wie es sich in den Items I3 und I4 widerspiegelt) einer Klasse mit höheren Leseleistungen einer Klasse einhergehen. Dieser positive Zusammenhang ist zwar zu beobachten (siehe Tabelle 3), allerdings ist er statistisch nicht signifikant. Die Befunde zeigen darüber hinaus, dass die Unterrichtsnutzung der Schüler/innen im Sinne der auf Klassenebene aggregierten Verarbeitungsprozesse durch den Lehrerenthusiasmus, nicht aber durch die Lehrerkooperation positiv vorhergesagt werden ( $\beta = .280$ ,  $p = .023$ ;  $\beta = .026$ ,  $p = .878$ ).

## Diskussion

Die Bildungsforschung hat wiederholt gezeigt, dass Lehrerenthusiasmus und Lehrerkooperation positiv mit Unterrichtsqualität assoziiert sind, welche wiederum höhere fachliche Schülerleistungen erwarten lässt. Im vorliegenden Beitrag wurde angenommen, dass sich diese Zusammenhänge auch in den Daten der PIRLS-2016-Erhebung für Österreich widerspiegeln.

Entgegen den Erwartungen zeigen die Analysen allerdings, dass sich diese Vermutung nicht bestätigen lässt. Der Grund dafür liegt u. a. in der vorhandenen Datenlage, die eine umfassende Abbildung differenzierter Unterrichtsqualitätsmerkmale (aus Schülersicht) nicht oder nur bedingt erlaubt. So zeigen fortgeschrittene Analysen, dass auf Klassenebene primär Informationen über die aus Schülersicht wahrgenommene, schülerorientierte Lernunterstützung (motivationale Unterstützung, Lernhilfen) vorliegen, für die angenommen werden kann, dass sie als remediale Unterrichtsmaßnahme vor allem in leseschwachen Klassen eingesetzt wird (z. B. Clausen, 2002; Gruehn, 2000; Helm et al., 2016). Die untersuchten Daten bestätigen diese Annahme. Die Unterrichtsqualität im Sinne (remedialer) Lernunterstützung ist moderat negativ mit den mittleren Leseleistungen einer Klasse assoziiert. Nach Gruehn (2000, S. 190) ist aber auch denkbar, dass schülerorientierter Unterricht – weil organisatorisch zeitaufwendiger – die verfügbare Unterrichtszeit für Lerngelegenheiten reduziert, worunter die Schülerleistungen leiden können.

Das Ausmaß der Lernunterstützung selbst wird entgegen den Erwartungen weder vom Ausmaß des Lehrerenthusiasmus noch vom Ausmaß der Lehrerkooperation positiv vorhergesagt. Vielmehr deuten die Analysen darauf hin, dass der Lehrerenthusiasmus schwach negativ mit der Lernunterstützung zusammenhängt, d. h., das Unterrichtsangebot von Lehrpersonen, die mit ihrer Arbeit positive Emo-

tionen verbinden, wird von Schülerinnen und Schülern als weniger stark lernunterstützend (im Sinne der Unterstützung der Schülerinteressen und des Diskutierens über Gelesenes im Unterricht) wahrgenommen als das Angebot von Lehrpersonen, die weniger enthusiastisch über ihre Arbeit berichten. Für diesen unerwarteten Befund sind mehrere Erklärungen denkbar. Genau genommen ist für die hier analysierten Daten noch nicht erforscht, welche Zusammenhänge überhaupt zu erwarten sind: (1) Dem Autor sind nur zwei Studien (Schiefele & Schaffner, 2015; Roth et al., 2007) bekannt, die diesen Zusammenhang auf Primarstufenebene untersuchen, während alle anderen identifizierten Studien entweder die Sekundarstufe in den Blick nehmen und/oder die Unterrichtsqualität nicht durch Schüler/innen, sondern Lehrpersonen einschätzen. Bei Schiefele und Schaffner (2015) sowie Roth et al. (2007) werden zudem andere Konstrukte der Unterrichtsqualität (Autonomieunterstützung und Zielorientierung) analysiert. (2) Eine weitere Erklärung könnte darin liegen, dass die schülerorientierte Lernunterstützung mit deutlich höherem Aufwand für Lehrpersonen (z. B. Diagnose des Lernstands und der Interessen der Schüler/innen, Vorbereitung und Organisation der inneren Differenzierung im Unterricht, Einsatz schüleraktiver Lernformen) einhergehen (z. B. Wang, 1982; Gruehn, 2000, S. 190; Helm, 2016, S. 67 ff.). Daher ist denkbar, dass Lehrpersonen, die mehr Lernunterstützung anbieten, auch mehr Belastung erleben und sich dieses Belastungserleben negativ auf das Erleben von positiven Emotionen wie Enthusiasmus, Stolz und Inspiration auswirkt. (3) Schließlich liegt die Vermutung nahe, dass Lehrerenthusiasmus mit der Klassengröße zusammenhängt. Lehrpersonen, die größere Klassen unterrichten, könnten weniger enthusiastisch sein, da für größere Klassen meist ein höherer Aufwand und eine höhere erlebte Belastung vermutet wird. Dies ist auch deshalb anzunehmen, da größere Klassen insbesondere in städtischen Schulen in meist schwierigeren Lagen (niedriger SES, bildungsferne Eltern) vorkommen. Entgegen diesen Annahmen zeigen hier nicht berichtete Zusatzanalysen allerdings, dass in größeren Klassen der Lehrerenthusiasmus sogar tendenziell höher ausgeprägt ist. Da die Schüler/innen in diesen Klassen gleichzeitig von einer geringeren Lernunterstützung berichten, ergibt sich ein leicht negativer Zusammenhang zwischen Lehrerenthusiasmus und Lernunterstützung. Dieser verschwindet nach Kontrolle der Klassengröße. Daher gilt grundsätzlich, dass der erwartungswidrige Zusammenhang zwischen Lehrerenthusiasmus und Lernunterstützung nicht zu stark interpretiert werden sollte.

Für die fehlenden Zusammenhänge zwischen den Lehrerprofessionsmerkmalen und der Unterrichtsqualität lässt sich die allgemeine Operationalisierung der Konstrukte als mögliche Ursache anführen: So können bspw. Kunter et al. (2008) zeigen, dass der Lehrerenthusiasmus insbesondere dann mit höherer Unterrichtsqualität einhergeht, wenn Lehrerenthusiasmus in Fach- und Unterrichtsenthusiasmus differenziert wird, da positive Effekte lediglich für den Unterrichtsenthusiasmus beobachtet wurden. Auch wenn diese Befunde pri-

mär für die Sekundarstufe gelten und unklar ist, ob die Differenzierung im Lehrereнтуhusiasmus auch für Primarstufenlehrkräfte (die ohnehin „alle“ Fächer unterrichten) relevant ist, so soll zumindest darauf hingewiesen werden, dass in der vorliegenden Studie der Lehrereнтуhusiasmus in Bezug auf den Lehrerberuf allgemein und nicht mit ausschließlicher Bezug auf die Unterrichtstätigkeit erfasst wurde. Denkbar ist auch, dass weniger der Lehrereнтуhusiasmus und vielmehr die Lehrselbstwirksamkeitserwartungen mit einem unterstützenden Lernangebot einhergehen (z. B. Zee & Koomen, 2016). Auch für die Lehrerkoooperation erweist sich die vorliegende Operationalisierung als nur bedingt für die untersuchten Hypothesen geeignet: Warwas und Helm (2018) zeigen, dass Lehrerkoooperation insbesondere dann mit Unterrichtsqualitätsmerkmalen einhergeht, wenn nicht nur die hier erfasste aktionale Dimension, sondern auch die strukturelle (z. B. Infrastruktur der Kooperation) und ideelle (z. B. gemeinsame Normen) Dimensionen hoch ausgeprägt sind. Gleichzeitig ist darauf hinzuweisen, dass in ca. der Hälfte aller österreichischen Volksschulen nur eine Klasse in der 4. Schulstufe (PIRLS-Klasse) existiert. An solchen Schulen arbeitet folglich auch weniger Lehrpersonal, was eine mögliche Kooperation erschwert. Inwiefern in kleinen Schulen Lehrerkoooperation überhaupt möglich ist und ob sich der Zusammenhang zur Unterrichtsqualität von größeren Schulen unterscheidet, muss künftig erst erforscht werden. Darüber hinaus ist fraglich, inwiefern die vorliegenden Items zur Lehrerkoooperation höhere Qualitätsstufen der Kooperation im Sinne der Arbeitsteilung und Kokonstruktion nach Gräsel et al. (2006) abbilden. Gleichzeitig gelang es mit der aktuellen Operationalisierung der Professionsmerkmale, die Annahme zu bestätigen, dass Volksschullehrpersonen, die stolz, inspiriert und enthusiastisch in Bezug auf ihre Arbeit sind, auch stärker kooperieren. Enthusiastische Lehrpersonen tauschen sich bspw. stärker mit anderen Kolleginnen und Kollegen aus, hospitieren häufiger und arbeiten bspw. stärker mit anderen Kolleginnen und Kollegen zusammen, um den Unterricht zu verbessern.

Die vorgelegten Analysen förderten einen weiteren diskussionswürdigen Befund, der allerdings nicht im Zentrum der Hypothesenbildung stand, zutage: Die im Durchschnitt einer Schulklasse berichteten individuellen Verarbeitungsprozesse des Unterrichtsangebots (z. B. das Verstehen der Lehrperson) hängen zwar positiv, aber statistisch nicht signifikant mit der mittleren Leseleistung einer Klasse zusammen. Eine mögliche Erklärung liegt darin, dass die verwendeten Items einerseits Ausdruck motivationaler Verarbeitungsprozesse (I1, I2) sind, die primär mit der Lesefreude assoziiert sein dürften und weniger mit den Leseleistungen. Andererseits sagen die Items (I5–I9) wenig über die fachdidaktische Qualität des Lernens aus. Vielmehr stehen allgemeine aktionale Handlungen (z. B. über das Gelesene reden) im Vordergrund. Es empfiehlt sich daher auch hier in künftigen PIRLS-Studien stärker fachdidaktisch validierte Instrumente zur Erfassung der Unterrichts- und Lernprozessqualität einzusetzen.

Die dargestellten Befunde sollten vor dem Hintergrund der Grenzen der vorliegenden Studie angemessen interpretiert werden.

(1) Das präsentierte Pfadmodell verleitet allzu schnell zur kausalen Interpretation der Ergebnisse. Die Voraussetzung dafür ist aber u. a. aufgrund des querschnittlichen Designs der PIRLS-Studien nicht gegeben. Daher ist auch die oben angeführte Interpretation des negativen Einflusses des Unterrichtsangebots auf die Leseleistung als umgekehrte Kausalität zulässig.

(2) Hinzu kommt, dass relevante alternative Prädiktoren der Schülerleistung, wie etwa das Vorwissen und die Bildungsnähe des Elternhauses, statistisch nicht als Kontrollvariablen berücksichtigt wurden. Da Informationen zur Bildungsnähe und zum sozioökonomischen Status des familiären Hintergrunds der Schüler/innen in den PIRLS-Daten enthalten sind, sollten zukünftige Studien diese berücksichtigen. Für die vorliegende Studie ist allerdings nicht zu erwarten, dass ihre Mitberücksichtigung die Ergebnisse beeinflussen würde, da zu vermuten ist, dass der negative Zusammenhang zwischen Lernunterstützung und Schülerleistung selbst dann bestehen bleibt, wie Helm et al. (2016) auf Basis von TIMSS-Daten zeigen konnten.

(3) Das in der vorliegenden Studie erfasste Merkmal von Unterrichtsqualität aus Schülersicht bezieht sich primär auf die wahrgenommene Lernunterstützung. Es fehlt eine adäquate Erfassung (etwa durch Videostudien) der fachdidaktischen Unterrichtsqualität, die insbesondere für die Vorhersage der Schülerleistungen als zentral angesehen werden kann (Baumert et al., 2010). Es ist fraglich, inwiefern Schüler/innen in der Lage sind, diese valide einzuschätzen. Es zeigt sich bereits bei allgemeineren Aspekten der Unterrichtsqualität wie den drei Basisdimensionen, dass die theoretisch postulierte 3-dimensionale Struktur auf Klassenebene oft nicht differenziert genug modelliert werden kann, sodass die Verwendung eines Globalfaktors eine nicht unübliche Vorgehensweise darstellt. Da aber von unterschiedlichen Basisdimensionen unterschiedliche Effekte auf die Leseleistung zu erwarten sind (z. B. Gruehn, 2000), sind Instrumente notwendig, die der dimensional Struktur gerecht werden.

(4) Aus methodischer Perspektive stellen die geringen Klassenebenenvarianzen der Items zum Unterrichtsangebot eine Herausforderung für die Analyse von Kontexteffekten dar. Lediglich 4,4 % bis 7,5 % der Variabilität des wahrgenommenen Unterrichtsangebots liegen auf Klassenebene. Dass das Unterrichtsangebot von den Schülerinnen und Schülern unterschiedlicher Klassen kaum unterschiedlich wahrgenommen wird, kann mehrere Gründe haben. So können die Itemformulierungen, die sich in fast allen Fällen nicht nur auf die Lehrperson, sondern auch auf die Schüler/innen beziehen, dazu beitragen. Allerdings weist auch Item I7 („Meine Leh-

rerin macht Verschiedenes, um uns beim Lernen zu helfen“), das sich ausschließlich auf die Klasse bezieht, mit 7,5 % eine niedrige Klassenebenenvarianz auf. Die Itemformulierung ist daher nur bedingt ausschlaggebend. Es ist zu hinterfragen, ob Schülerratings im Volksschulbereich valide Einschätzungen des Unterrichtsangebots erlauben.

(5) Die Lehrerprofessionsmerkmale wurden durch Lehrerselbsteinschätzungen erhoben, die einer sozial erwünschten Antworttendenz unterliegen können.

(6) Schließlich ist die Stichprobe aufgrund des Sampling-Designs in den PIRLS-Erhebungen zwar auf Individualebene (Schüler/innen der 4. Schulstufe in Österreich) repräsentativ, nicht jedoch auf Klassenebene (Foy, 2018), d. h., die hier berichteten Effekte auf Klassenebene lassen sich nur bedingt verallgemeinern.

Es lässt sich zusammenfassend festhalten, dass Bedarf daran besteht, sowohl Lehrerprofessionsmerkmale als auch Unterrichtsqualitätsmerkmale in künftigen PIRLS-Studien differenzierter (insbesondere mit Blick auf ihre qualitative, inhaltliche Struktur) zu erfassen.

Trotz dieser Einschränkungen bieten die PIRLS-Daten einen großen Vorteil gegenüber anderen Datensätzen, die in der Unterrichtsforschung häufig zum Einsatz kommen: Insbesondere für Österreich existieren kaum Studien, die eine vergleichbare Stichprobengröße auf Klassen- bzw. Lehrer- oder Schulebene vorweisen können. Eine derart große Stichprobe erlaubt es, komplexere Modelle, wie die hier durchgeführten Mehrebenen-Strukturgleichungsmodelle, die den Stichproben- und Messfehler kontrollieren und unterschiedli-

che Effekte der Lernumgebung vor dem Hintergrund aktueller methodischer Entwicklungen (shared and configural constructs; Stapleton et al., 2018) adäquat modellieren, zu schätzen. Zudem liegt aufgrund der Schüler- und Lehrerkontextfragebögen ein umfangreiches Set an erhobenen Variablen auf unterschiedlichen Ebenen des Lehr- und Lernprozesses vor. Aus diesen Gründen sollten die vorliegenden Large-Scale-Daten – trotz der Einschränkungen, die sie aufweisen – stärker für Fragen der Unterrichts- und Bildungsforschung in Österreich genutzt werden.

Auch wenn in der vorliegenden Studie die postulierten Zusammenhänge mehrheitlich nicht beobachtet werden konnten, so wurde in anderen Studien, die auf diese Forschungsfragen zugeschnitten sind – bspw. Pelletier et al. (2002), Roth et al. (2007) und Müller et al. (2009) – wiederholt gezeigt, dass von Lehrpersonen erlebter Druck durch Bildungsadministrationen, Schulleitungen und Eltern die Lehrermotivation einschränkt und negativ auf die Unterrichtsqualität wirkt. Die Schaffung von Rahmenbedingungen, die die Lehrermotivation und -kooperation fördern, sollte daher im Sinne der Förderung der Schüler/innen ein Anliegen aller Akteure im Bildungswesen sein. Bildungspolitik, -behörden, Schulleitungen, Lehrkörper, Eltern etc. können einen Beitrag zur Unterstützung der hier untersuchten Determinanten professionellen Lehrerhandelns leisten. Auch für die Lehrerkoooperation gilt, dass ein unterstützendes Schulleiterverhalten die Arbeit in professionellen Teams bedeutend fördern kann (Warwas, Helm & Schadt, 2019), was wiederum mit Unterrichtsqualität einhergeht (Warwas & Helm, 2018). Hier bedarf es insbesondere in Österreich weiterer Forschung, die den motivations- und kooperationsförderlichen Rahmenbedingungen im Lehrerberuf nachgehen.

---

## Literatur

- Bakker, A. B. (2005). Flow among music teachers and their students: The crossover of peak experiences. *Journal of Vocational Behavior*, 66 (1), 26–44. doi:10.1016/j.jvb.2003.11.001
- Bakker, A. B. & Bal, M. P. (2010). Weekly work engagement and performance: A study among starting teachers. *Journal of Occupational and Organizational Psychology*, 83 (1), 189–206. doi:10.1348/096317909X402596
- Baumert, J. & Kunter, M. (2006). Stichwort: Professionelle Kompetenz von Lehrkräften. *Zeitschrift für Erziehungswissenschaft*, 9 (4), 469–520. doi:10.1007/s11618-006-0165-2
- Baumert, J., Kunter, M., Blum, W., Brunner, M., Voss, T., Jordan, A. et al. (2010). Teachers' Mathematical Knowledge, Cognitive Activation in the Classroom, and Student Progress. *American Educational Research Journal*, 47 (1), 133–180. doi:10.3102/0002831209345157
- Bonsen, M. & Hübner, C. (2012). Unterrichtsentwicklung in Professionellen Lerngemeinschaften. In K.-O. Bauer & N. Logemann (Hrsg.), *Effektive Bildung. Zur Wirksamkeit und Effizienz pädagogischer Prozesse* (S. 55–76). Münster u.a.: Waxmann.
- Butler, R. & Shibaz, L. (2008). Achievement goals for teaching as predictors of students' perceptions of instructional practices and students' help seeking and cheating. *Learning and Instruction*, 18 (5), 453–467. doi:10.1016/j.learninstruc.2008.06.004
- Clausen, M. (2002). *Unterrichtsqualität: eine Frage der Perspektive? Empirische Analysen zur Übereinstimmung, Konstrukt- und Kriteriumsvalidität* (Pädagogische Psychologie und Entwicklungspsychologie). Münster: Waxmann.
- Deci, E. L. & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior* (Perspectives in social psychology). New York, NY et al.: Plenum Press.
- Fauth, B., Decristan, J., Rieser, S., Klieme, E. & Büttner, G. (2014). Student ratings of teaching quality in primary school: Dimensions and prediction of student outcomes. *Learning and Instruction*, 29, 1–9. doi:10.1016/j.learninstruc.2013.07.001
- Findeisen, S. (2017). *Fachdidaktische Kompetenzen angehender Lehrpersonen. Eine Untersuchung zum Erklären im Rechnungswesen* (Economics Education und Human Resource Management). Wiesbaden: Springer.
- Foy, P. (Ed.). (2018). *PIRLS 2016. User Guide for the International Database*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
- Foy, P., Brossman, B. & Galia, J. (2012). TIMSS and PIRLS Achievement Scaling Methodology. In M. O. Martin & I. V. S. Mullis (Eds.), *Methods and procedures in TIMSS and PIRLS 2011* (pp. 1–28). Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
- Frenzel, A., Goetz, T., Lüdtke, O., Pekrun, R. & Sutton, R. E. (2009). Emotional transmission in the classroom: Exploring the relationship between teacher and student enjoyment. *Journal of Educational Psychology*, 101 (3), 705–716. doi:10.1037/a0014695
- Fulton, K. & Britton, T. (2011). *STEM Teachers in Professional Learning Communities: From Good Teachers to Great Teaching*. Washington DC: National Commission on Teaching and America's Future.
- Geldhof, G. J., Preacher, K. J. & Zyphur, M. J. (2014). Reliability estimation in a multilevel confirmatory factor analysis framework. *Psychological methods*, 19 (1), 72–91. doi:10.1037/a0032138
- George, A. C., Oberwimmer, K. & Itzlinger-Bruneforth, U. (2016). Stichprobenziehung. In S. Breit & C. Schreiner (Hrsg.), *Large-Scale Assessment mit R. Methodische Grundlagen der österreichischen Bildungsstandardüberprüfung* (S. 51–81). Wien: Facultas.

- Gräsel, C., Stark, R., Sparka, A. & Herzmann, P. (2007). Schulische Kooperationsmuster und die Implementation eines Programms zur Förderung der Lesekompetenz. *Zeitschrift für Berufs- und Wirtschaftspädagogik* (Beiheft 21), 93–107.
- Gruehn, S. (2000). *Unterricht und schulisches Lernen. Schüler als Quellen der Unterrichtsbeschreibung*. Münster: Waxmann.
- Hafner, J. (2017). Stichproben-Design und Stichprobenziehung. In C. Wallner-Paschon & U. Itzlinger-Bruneforth (Hrsg.), *PIRLS 2016 Technischer Bericht* (S. 23–28). Salzburg: Bundesinstitut für Bildungsforschung, Innovation & Entwicklung des österreichischen Schulwesens (BIFIE).
- Han, J., Yin, H. & Wang, W. (2015). Exploring the relationship between goal orientations for teaching of tertiary teachers and their teaching approaches in China. *Asia Pacific Education Review*, 16 (4), 537–547. doi:10.1007/s12564-015-9392-7
- Han, J., Yin, H. & Wang, W. (2016). The effect of tertiary teachers' goal orientations for teaching on their commitment: the mediating role of teacher engagement. *Educational Psychology*, 36 (3), 526–547. doi:10.1080/01443410.2015.1044943
- Hein, V., Ries, F., Pires, F., Caune, A., Heszteráné Ekler, J., Emeljanovas, A. et al. (2012). The Relationship Between Teaching Styles and Motivation to Teach Among Physical Education Teachers. *Journal of Sports Science & Medicine*, 11 (1), 123–130.
- Helm, C., Kemethofer, D., Moosbrugger, R., Bröderbauer, S. & Luthe, S. (2016). Effekte der aus Schulleitersicht wahrgenommenen Belastungen des schulischen Lernens auf die Unterrichtsgestaltung und die Schülerkompetenzen in Mathematik sowie auf die Schulzufriedenheit. *Zeitschrift für Bildungsforschung*, 6 (3), 245–264. doi:10.1007/s35834-016-0165-0
- Helmke, A. (2015). *Unterrichtsqualität und Lehrerprofessionalität. Diagnose, Evaluation und Verbesserung des Unterrichts. Franz Emanuel Weinert gewidmet* (Schule weiterentwickeln, Unterricht verbessern, 6. Aufl.). Seelze-Velber: Klett/Kallmeyer.
- Holzberger, D., Philipp, A. & Kunter, M. (2016). Ein Blick in die Black-Box. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 48 (2), 90–105. doi:10.1026/0049-8637/a000150
- Hooper, M., Mullis, I. V. S. & Martin, M. O. (2015). PIRLS 2016 Context Questionnaire Framework. In I. V. S. Mullis & M. O. Martin (Eds.), *PIRLS 2016 Assessment Framework* (S. 31–54). Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
- Hu, L.-t. & Bentler, P. M. (1999). Cutoff Criteria for Fit Indexes in Covariance Structure Analysis: Conventional Criteria versus New Alternatives. *Structural Equation Modeling*, 6 (1), 1–55. doi:10.1080/10705519909540118
- Joncas, M. & Foy, P. (2012). Sample Design in TIMSS and PIRLS. In M. O. Martin & I. V. S. Mullis (Eds.), *Methods and procedures in TIMSS and PIRLS 2011* (S. 1–21). Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
- Keller, M. M., Goetz, T., Becker, E. S., Morger, V. & Hensley, L. (2014). Feeling and showing: A new conceptualization of dispositional teacher enthusiasm and its relation to students' interest. *Learning and Instruction*, 33, 29–38. doi:10.1016/j.learninstruc.2014.03.001
- Keller, M. M., Neumann, K. & Fischer, H. E. (2017). The impact of physics teachers' pedagogical content knowledge and motivation on students' achievement and interest. *Journal of Research in Science Teaching*, 54 (5), 586–614. doi:10.1002/tea.21378
- Klassen, R. M., Aldhafri, S., Mansfield, C. F., Purwanto, E., Siu, A. F. Y., Wong, M. W. et al. (2012). Teachers' Engagement at Work: An International Validation Study. *The Journal of Experimental Education*, 80 (4), 317–337. doi:10.1080/00220973.2012.678409
- Klassen, R. M., Perry, N. E. & Frenzel, A. (2012). Teachers' relatedness with students: An underemphasized component of teachers' basic psychological needs. *Journal of Educational Psychology*, 104 (1), 150–165. doi:10.1037/a0026253
- Klieme, E. & Kuger, S. (2016). *Umgang mit Heterogenität an Schulen des Sekundarbereichs im internationalen Vergleich*, Berlin. Verfügbar unter [http://www.geb2016.de/aktuelles/Symposien\\_Abstracts\\_240216.pdf](http://www.geb2016.de/aktuelles/Symposien_Abstracts_240216.pdf)



- Klieme, E., Pauli, C. & Reusser, K. (2009). The Pythagoras Study. Investigating effects of teaching and learning in Swiss and German mathematics classrooms. In T. Janik & T. Seidel (Eds.), *The power of video studies in investigating teaching and learning in the classroom* (S. 137–160). Münster et al.: Waxmann.
- Kunter, M. (2011). Motivation als Teil der professionellen Kompetenz – Forschungsbefunde zum Enthusiasmus von Lehrkräften. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss & M. Neubrand (Hrsg.), *Professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV* (S. 259–275). Münster u.a.: Waxmann.
- Kunter, M., Frenzel, A., Nagy, G., Baumert, J. & Pekrun, R. (2011). Teacher enthusiasm: Dimensionality and context specificity. *Contemporary Educational Psychology*, 36 (4), 289–301. doi:10.1016/j.cedpsych.2011.07.001
- Kunter, M., Klusmann, U., Baumert, J., Richter, D., Voss, T. & Hachfeld, A. (2013). Professional competence of teachers: Effects on instructional quality and student development. *Journal of Educational Psychology*, 105 (3), 805–820. doi:10.1037/a0032583
- Kunter, M., Tsai, Y.-M., Klusmann, U., Brunner, M., Krauss, S. & Baumert, J. (2008). Students' and mathematics teachers' perceptions of teacher enthusiasm and instruction. *Learning and Instruction*, 18 (5), 468–482. doi:10.1016/j.learninstruc.2008.06.008
- Kunter, M. & Voss, T. (2013). The Model of Instructional Quality in COACTIV: A Multicriteria Analysis. In M. Kunter (Ed.), *Cognitive activation in the mathematics classroom and professional competence of teachers. Results from the COACTIV project* (Mathematics teacher education, Vol. 8, pp. 97–124). New York: Springer. doi:10.1007/978-1-4614-5149-5\_6
- Lavié, J. M. (2006). Academic Discourses on School-Based Teacher Collaboration: Revisiting the Arguments. *Educational Administration Quarterly*, 42 (5), 773–805. doi:10.1177/0013161X06290647
- Lee, V. E. & Smith, J. B. (1996). Collective Responsibility for Learning and Its Effects on Gains in Achievement for Early Secondary School Students. *American Journal of Education*, 104 (2), 103–147. doi:10.1086/444122
- Little, T. D. (2013). *Longitudinal structural equation modeling* (Methodology in the social sciences). New York NY: Guilford Press.
- Lomos, C., Hofman, R. H. & Bosker, R. J. (2011). Professional communities and student achievement – a meta-analysis. *School Effectiveness and School Improvement*, 22 (2), 121–148. doi:10.1080/09243453.2010.550467
- Louis, K. S. & Marks, H. M. (1998). Does Professional Community Affect the Classroom? Teachers' Work and Student Experiences in Restructuring Schools. *American Journal of Education*, 106 (4), 532–575. doi:10.1086/444197
- Lüdtke, O., Marsh, H. W., Robitzsch, A. & Trautwein, U. (2011). A 2 × 2 taxonomy of multilevel latent contextual models. Accuracy-bias trade-offs in full and partial error correction models. *Psychological methods*, 16 (4), 444–467. doi:10.1037/a0024376
- Lüdtke, O., Marsh, H. W., Robitzsch, A., Trautwein, U., Asparouhov, T. & Muthén, B. (2008). The multilevel latent covariate model: a new, more reliable approach to group-level effects in contextual studies. *Psychological methods*, 13 (3), 203–229. doi:10.1037/a0012869
- Marsh, H. W., Lüdtke, O., Nagengast, B., Trautwein, U., Morin, A. J. S., Abduljabbar, A. S. et al. (2012). Classroom Climate and Contextual Effects: Conceptual and Methodological Issues in the Evaluation of Group-Level Effects. *Educational Psychologist*, 47 (2), 106–124. doi:10.1080/00461520.2012.670488
- Morin, A. J. S., Marsh, H. W., Nagengast, B. & Scalas, L. F. (2014). Doubly Latent Multilevel Analyses of Classroom Climate: An Illustration. *The Journal of Experimental Education*, 82 (2), 143–167. doi:10.1080/00220973.2013.769412
- Müller, F. H., Hanfstingl, B. & Andreitz, I. (2009). Bedingungen und Auswirkungen selbstbestimmter Lehrermotivation. *Erziehung & Unterricht*, 159 (1–2), 142–152.

- Mullis, I. V. S. & Martin, M. O. (Hrsg.). (2015). *PIRLS 2016 Assessment Framework*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
- Muthén, L. K. & Muthén, B. O. (1998–2018). *Mplus User's Guide* (8th ed.). Los Angeles, CA: Muthén & Muthén.
- Nortvedt, G. A., Gustafsson, J.-E. & Lehre, A.-C. W. (2016). The Importance of Instructional Quality for the Relation Between Achievement in Reading and Mathematics. In T. Nilsen & J.-E. Gustafsson (Eds.), *Teacher quality, instructional quality and student outcomes. Relationships across countries, cohorts and time* (IEA Research for Education, Vol. 2, pp. 97–113). Switzerland: SpringerOpen. doi:10.1007/978-3-319-41252-8\_5
- Park, Y., Brownell, M. T., Bettini, E. F. & Benedict, A. E. (2017). Multiple Dimensions of Instructional Effectiveness in Reading: A Review of Classroom Observation Studies and Implications for Special Education Classrooms. *Exceptionality*, 42, 1–17. doi:10.1080/09362835.2017.1283628
- Pelletier, L. G., Séguin-Lévesque, C. & Legault, L. (2002). Pressure from above and pressure from below as determinants of teachers' motivation and teaching behaviors. *Journal of Educational Psychology*, 94 (1), 186–196. doi:10.1037//0022-0663.94.1.186
- Penner-Williams, J., Díaz, E. I. & Gonzales Worthen, D. (2017). PLCs: Key PD component in learning transfer for teachers of English learners. *Teaching and Teacher Education*, 65, 215–229. doi:10.1016/j.tate.2017.03.014
- Perera, H. N., Vosicka, L., Granziera, H. & McIlveen, P. (2018). Towards an integrative perspective on the structure of teacher work engagement. *Journal of Vocational Behavior*, 108, 28–41. doi:10.1016/j.jvb.2018.05.006
- Praetorius, A.-K., Klieme, E., Herbert, B. & Pinger, P. (2018). Generic dimensions of teaching quality: the German framework of Three Basic Dimensions. *ZDM*, 50 (3), 407–426. doi:10.1007/s11858-018-0918-4
- Richter, D. & Pant, H. A. (2016). *Lehrerkooperation in Deutschland. Eine Studie zu kooperativen Arbeitsbeziehungen bei Lehrkräften der Sekundarstufe I*. Gütersloh: Bertelsmann Stiftung.
- Robitzsch, A., Pham, G. & Yanagida, T. (2016). Fehlende Daten und Plausible Values. In S. Breit & C. Schreiner (Hrsg.), *Large-Scale Assessment mit R. Methodische Grundlagen der österreichischen Bildungsstandardüberprüfung* (S. 259–293). Wien: Facultas.
- Rone, B. C. (2009). *The Impact of the Data Team Structure on Collaborative Teams and Student Achievement*. Lindenwood: University Missouri.
- Roth, G., Assor, A., Kanat-Maymon, Y. & Kaplan, H. (2007). Autonomous motivation for teaching: How self-determined teaching may lead to self-determined learning. *Journal of Educational Psychology*, 99 (4), 761–774. doi:10.1037/0022-0663.99.4.761
- Rubin, D. B. (1987). *Multiple Imputation for nonresponse in surveys* (Wiley series in probability and mathematical statistics Applied probability and statistics). New York, NY: Wiley.
- Schaufeli, W. B. & Bakker, A. B. (2004). *UWES Utrecht Work Engagement Scale. Preliminary Manual*. Utrecht: Occupational Health Psychology Unit, Utrecht University.
- Schaufeli, W. B., Bakker, A. B. & Salanova, M. (2006). The Measurement of Work Engagement With a Short Questionnaire. *Educational and Psychological Measurement*, 66 (4), 701–716. doi:10.1177/0013164405282471
- Scheerens, J. (2016). *Educational Effectiveness and Ineffectiveness. A Critical Review of the Knowledge Base* (1st ed.). Dordrecht: Springer. doi:10.1007/978-94-017-7459-8
- Scherer, R. & Gustafsson, J.-E. (2015). Student assessment of teaching as a source of information about aspects of teaching quality in multiple subject domains: An application of multilevel bifactor structural equation modeling. *Frontiers in psychology*, 6, 15–50. doi:10.3389/fpsyg.2015.01550

- Schiefele, U. & Schaffner, E. (2015). Teacher interests, mastery goals, and self-efficacy as predictors of instructional practices and student motivation. *Contemporary Educational Psychology, 42*, 159–171. doi:10.1016/j.cedpsych.2015.06.005
- Schreiner, C., Wiesner, C., Kemethofer, D., Helm, C., Freller-Töglhofer, M., Illetschko, M. et al. (im Druck). *Zyklusbericht zur Bildungsstandarderhebung 2017 – Charakterisierung des fachlichen Unterrichts*. Salzburg: Bundesinstitut für Bildungsforschung, Innovation und Entwicklung des österreichischen Schulwesens (BIFIE).
- Skaalvik, E. M. & Skaalvik, S. (2014). Teacher self-efficacy and perceived autonomy: relations with teacher engagement, job satisfaction, and emotional exhaustion. *Psychological reports, 114* (1), 68–77. doi:10.2466/14.02.PRO.114k14w0
- Stapleton, L. M., Yang, J. S. & Hancock, G. R. (2016). Construct Meaning in Multilevel Settings. *Journal of Educational and Behavioral Statistics, 41* (5), 481–520. doi:10.3102/1076998616646200
- Suchań, B. & Breitfuß-Muhr, G. (2009). Leseunterricht an Österreichs Volksschulen. In B. Suchan, C. Wallner-Paschon & C. Schreiner (Hrsg.), *PIRLS 2006. Die Lesekompetenz am Ende der Volksschule. Österreichischer Expertenbericht* (S. 17 – 92). Graz: Leykam.
- Supovitz, J. A. (2002). Developing Communities of Instructional Practice. *Teachers College Record, 104* (8), 1591–1626.
- Tschannen-Moran, M. & Hoy, A. W. (2007). The differential antecedents of self-efficacy beliefs of novice and experienced teachers. *Teaching and Teacher Education, 23* (6), 944–956. doi:10.1016/j.tate.2006.05.003
- Vescio, V., Ross, D. & Adams, A. (2008). A review of research on the impact of professional learning communities on teaching practice and student learning. *Teaching and Teacher Education, 24* (1), 80–91. doi:10.1016/j.tate.2007.01.004
- Wagner, W., Göllner, R., Helmke, A., Trautwein, U. & Lüdtke, O. (2013). Construct validity of student perceptions of instructional quality is high, but not perfect. Dimensionality and generalizability of domain-independent assessments. *Learning and Instruction, 28*, 1–11. doi:10.1016/j.learninstruc.2013.03.003
- Wang, M. C. (1982). Entwicklung und Förderung von Kompetenzen zur Selbststeuerung und zum Selbst-Management bei Schülern. *Unterrichtswissenschaft, 2*, 129–139.
- Warwas, J. & Helm, C. (2017). Enjoying working and learning in vocational education: A multilevel investigation of emotional crossover and contextual moderators. *Empirical Research in Vocational Education and Training, 9* (1), 1–50. doi:10.1186/s40461-017-0055-2
- Warwas, J. & Helm, C. (2018). Professional learning communities among vocational school teachers: Profiles and relations with instructional quality. *Teaching and Teacher Education, 73*, 43–55. doi:10.1016/j.tate.2018.03.012
- Warwas, J., Helm, C. & Schadt, C. (2019). Unterstützendes Führungsverhalten schulischer Führungskräfte für die Arbeit Professioneller Lerngemeinschaften im Kollegium. *Zeitschrift für Bildungsforschung*. Online First doi:10.1007/s35834-019-00230-w
- Widauer, K. & Wallner-Paschon, C. (2017). Entwicklung und Aufbau der Testinstrumente und Kontextfragebögen. In C. Wallner-Paschon & U. Itzlinger-Bruneforth (Hrsg.), *PIRLS 2016 Technischer Bericht* (S. 9–21). Salzburg: Bundesinstitut für Bildungsforschung, Innovation und Entwicklung des österreichischen Schulwesens (BIFIE).
- Zee, M. & Koomen, H. M. Y. (2016). Teacher Self-Efficacy and Its Effects on Classroom Processes, Student Academic Adjustment, and Teacher Well-Being. *Review of Educational Research, 86* (4), 981–1015. doi:10.3102/0034654315626801

