Prüfungsangst in Mathematik am Ende der Grundschulzeit: Entwicklung und Interaktionen mit Leistung und Selbstkonzept

Anna Jonberg · Claudia Pereira Kastens · Frank Lipowsky

Eingegangen: 10. Januar 2020 / Überarbeitet: 4. Dezember 2020 / Angenommen: 9. März 2021 / Online publiziert: 31. März 2021
© Der/die Autor(en) 2021, corrected publication 2022

Zusammenfassung Dieser Beitrag prüft die Entwicklung der Prüfungsangst und deren Effekt auf die Leistungsentwicklung unter zusätzlicher Berücksichtigung des akademischen Selbstkonzepts anhand einer Stichprobe von $N = 725$ Grundschulkindern vom Ende des dritten bis Ende des vierten Schuljahres. Die Prüfungsangst fällt im untersuchten Zeitraum leicht ab und weist dabei eine mittlere korrelative Stabilität auf. CLP Analysen zeigen einen reziproken Zusammenhang von Prüfungsangst und Leistung. Werden diese Analysen für anhand einer latenten Profilanalyse identifizierten Klassen (relatives hohes, mittleres, geringes Selbstkonzept) durchgeführt, zeigen sich differenziertere Ergebnisse: Prüfungsangst ist bei Schüler*innen mit geringen Selbstkonzepten am höchsten und steigt außerdem an, während sie bei jenen mit hohen Selbstkonzepten geringer ausfällt und zusätzlich sinkt. Bei kompetenzüberzeugten Schüler*innen erweist sich die Prüfungsangst als negativer Prädiktor für Leistung, während sich die Effekte für diejenigen mit geringen Selbstkonzepten umkehren. Für Schüler*innen mit mittleren Selbstkonzepten zeigen sich keine

Die Autorinnen Anna Jonberg und Claudia Pereira Kastens haben zu gleichen Teilen zum Manuskript beigetragen.

A. Jonberg (✉) · Jun. Prof. Dr. C. P. Kastens
Institut für Bildungsforschung, Grundschulforschung, Bergische Universität Wuppertal,
Gaußstr. 20, 42119 Wuppertal, Deutschland
E-Mail: jonberg@uni-wuppertal.de

Jun. Prof. Dr. C. P. Kastens
E-Mail: kastens@uni-wuppertal.de

Prof. Dr. F. Lipowsky
Institut für Erziehungswissenschaft, Empirische Schul- und Unterrichtsforschung, Universität Kassel,
Nora-Platiel-Straße 1, 34127 Kassel, Deutschland
E-Mail: lipowsky@uni-kassel.de

Veränderungen in der Ausprägung der Prüfungsangst und keine Effekte mit der Leistung.

Schlüsselwörter Cross-Lagged-Panel · Latente Profilanalyse · Prüfungsangst · Selbstkonzept

Test anxiety in mathematics for fourth grade elementary school children: development and interaction with achievement and self-concept

Abstract This article reviews the relationship between test anxiety and achievement from the end of third to the end of fourth grade of German elementary school classes in math, and whether academic self-concept predicts the relationship between these two factors. The sample consists of $N=725$ students. While correlational stability of test anxiety was moderate, we found a slight decline within fourth grade. Cross-lagged-panel analysis demonstrates a reciprocal effect between test anxiety and performance. When replicating this analysis for a priori identified latent classes (relatively low, medium, and high self-concept), the results revealed a more complex picture: children with lower academic self-concepts develop higher test anxiety which increases over time, and in this group, achievement is a negative predictor for test anxiety. Children with higher levels of academic self-concept however report lower test anxiety, which decreases over time and test anxiety is a negative predictor of future achievement. No changes in test anxiety and longitudinal relationships with achievement occur for the group reporting medium academic self-concepts.

Keywords Latent profile analysis · Path model · Self-concept · Test anxiety

1 Einleitung

Lern- und leistungsbezogene Ängste gelten als motivationshinderlicher und leistungshemmender Faktor (Hembree 1990; Jonberg und Porsch 2017; Pekrun und Götz 2006; Sorvo et al. 2019). Zudem beeinträchtigen sie die kindliche Lebensqualität und emotionale Befindlichkeit (Pixner und Kaufmann 2013). In der Literatur werden die Begriffe *Leistungsangst* und *Prüfungsangst* (PA, im Engl. beides *test anxiety*) oft synonym verwendet und umfassen im Gegensatz zu allgemeinen schul- oder fachbezogenen Ängstlichkeitsmaßen (siehe auch Sorvo et al. 2019) kognitive und affektive Angst-Reaktionen (*Worry*- und *Emotionality-Komponente*, Zeidner 1998) von Lernenden in ausgewiesenen schulischen Leistungssituationen. Die PA kann auch als eine Facette der schulischen Leistungsangst beschrieben werden und bezieht sich dann explizit auf Test- und Prüfungssituationen im Unterricht, die mit Leistungsbewertungen verbunden sind (Pekrun und Götz 2006).

Die Mechanismen, die den Zusammenhang zwischen PA und Leistungsentwicklung erklären, sind komplex. Die *Defizittheorie* besagt, dass Leistungsdefizite in einem Fach langfristig die Entstehung von Ängsten begünstigen (Krinzinger et al. 2009), während andere Modelle Effekte einer durch Ängste induzierten Leistungs-

minderung beschreiben (Sorvo et al. 2019). Gemeinhin wird jedoch angenommen, dass PA und Leistung in einem reziprok negativen Zusammenhang stehen (u. a. Carey et al. 2016). Gepaart mit der Entwicklung von Vermeidungstendenzen kann langfristig ein Teufelskreis entstehen: Durch weniger intensive Auseinandersetzung mit dem Lernstoff wird weniger gelernt, was – bei nachfolgender ungünstiger Lernkurve – einen Anstieg leistungsbezogener Ängste begünstigen kann, wiederum gefolgt von negativen Effekten auf die Leistung (Hembree 1990; Pekrun et al. 2017).

Bislang gibt es wenige längsschnittliche Studien, die das Zusammenspiel von PA und Leistungsentwicklung am Ende der Grundschulzeit betrachten. Dabei besteht insbesondere im deutschsprachigen Raum Forschungs- und möglicherweise Handlungsbedarf, wenn man berücksichtigt, dass sich die Kinder zu diesem Zeitpunkt mit der Einführung von Noten im dritten und der anstehenden Übergangsentscheidung im vierten Schuljahr oft einem erhöhten (Leistungs-)Druck ausgesetzt sehen (Pixner und Kaufmann 2013; Valtin und Wagner 2004).

Der *Kontroll-Wert-Theorie* zu Bedingungen und Wirkungen von Prüfungsangst von Pekrun und Götz (2006) zufolge, wird PA neben Leistungserfahrungen auch von Überzeugungen im Sinne eines Kontrollerlebens und der Valenz von Misserfolg und dessen Konsequenzen beeinflusst. Das akademische Selbstkonzept (aSK) ist definiert als die Wahrnehmung und Bewertung der eigenen akademischen Leistungsfähigkeit (Shavelson et al. 1976). Es zeigt bereits in der Grundschulzeit Zusammenhänge mit schulischer Leistung (Ehm et al. 2019; Faber 2012; Helmke und van Aken 1995; Kastens et al. 2013; Lipowsky et al. 2011). In der vorliegenden Untersuchung wird das Zusammenspiel der Entwicklung von PA und Leistung im Fach Mathematik von Kindern am Ende des dritten und vierten Schuljahres analysiert. Es wird überprüft, wie stabil sich das Konstrukt PA von der dritten zur vierten Klasse darstellt und welche reziproken Zusammenhänge zwischen den beiden Konstrukten bestehen. Ferner untersuchen wir, inwiefern die Ausprägung des aSK Einfluss auf diese Beziehung hat.

2 Theoretische und empirische Einbettung

2.1 Prüfungsangst am Ende der Grundschulzeit

Der *Kontroll-Wert-Theorie* nach (Pekrun und Götz 2006) beeinflusst PA als negativ-aktivierende Emotion das Lernen und die schulische Leistung auf drei Ebenen: Erstens werden bereits in der Vorbereitung auf Prüfungssituationen und in der Stoffaufarbeitung eher starre Lernstrategien, wie Auswendiglernen, effektiveren Strategien, wie Anstrengungsmanagement (z. B. sich Mühe geben, den Lernstoff zu verstehen) oder Überwachungsstrategien (z. B. sich selbst Fragen zu Lerninhalten stellen) vorgezogen. Zweitens beeinträchtigt PA die Motivation. So ist diese weniger intrinsisch und das Interesse am Stoff geht verloren. Drittens kann die Angst zu versagen bei starker Ausprägung kognitive Ressourcen binden, was insbesondere bei herausfordernden Aufgaben dazu führt, dass diese weniger effektiv gelöst werden. Misserfolg führt dann zur Entstehung oder Verstärkung von PA, wenn er bzw. seine Konsequenzen eine gewisse Wertigkeit (Valenz) für die Person haben. Zusätzlich spielt

subjektiver Kontrollmangel, auch im Sinne einer Kompetenzüberzeugung, eine zentrale Rolle bei der Genese von Angst.

Während bei älteren Lernenden in der Regel negative Zusammenhänge zwischen PA und der Leistung zu finden sind, gibt es einige Untersuchungen, die diese Zusammenhänge in den ersten Schuljahren nicht bestätigen. So fanden Hembree (1988) und Krininger et al. (2009) keine Zusammenhänge zwischen PA und Leistung, obwohl die Lernenden Anzeichen von PA bei sich wahrnahmen. Wenn in Leistungssituationen komplexe Problemlösekompetenzen gefordert werden, das Arbeitsgedächtnis also stark belastet wird, konnten jedoch auch bei jüngeren Lernenden Zusammenhänge zwischen PA und Leistung nachgewiesen werden (Ramirez et al. 2013; Wu et al. 2012). Insgesamt lassen sich im Primarbereich Korrelationen zwischen $r = -0,11$ und $r = -0,37$ aufzeigen (Helmke 1997; Pixner und Kaufmann 2013) und tendenziell werden die querschnittlichen Zusammenhänge zwischen PA und Leistungsmaßen im Laufe der Schulzeit enger (Hembree 1988; Pixner und Kaufmann 2013).

Bisherige Studien, die den Zusammenhang von PA und Leistung querschnittlich korrelativ untersuchten, erlauben allerdings keine Rückschlüsse auf die Wirkungsrichtung des Zusammenhangs der Konstrukte. Gemeinhin werden diese Korrelationen derart interpretiert, dass die PA als Prädiktor der Leistung betrachtet wird. Wie bereits erwähnt, ist die Entwicklung von PA jedoch nicht als unabhängig von gemachten Lernerfahrungen zu sehen. Längsschnittliche Erhebungen, die reziproke Zusammenhänge zwischen PA und Leistungsmaßen am Ende des Primarbereichs untersuchen, sind uns dennoch nicht bekannt.

Ähnlich dünn ist die Befundlage bei der Frage nach der Stabilität der PA während der Grundschulzeit (z. B. Eden et al. 2013). Ältere Studien berichten bezüglich der mittleren Ausprägung der PA einen Anstieg in den ersten (Grund-)Schuljahren bis hin zu einer Stagnation zu Beginn der Sekundarstufe (Hembree 1988). Zur korrelativen Stabilität von PA fanden Martschinke und Kammermeyer (2006) einen Zusammenhang von $r = 0,50$ ($p < 0,05$), gemessen je am Anfang und Ende des ersten Schuljahres. PA scheint sich erst über die Grundschulzeit hinweg zu entwickeln und ebenso festigen sich Zusammenhänge mit Leistungsmaßen über die Zeit (siehe zusammenfassend Cortina 2008). Es ist anzunehmen, dass die PA – ähnlich wie andere leistungsbezogene personen- und selbstbezogene Überzeugungen – bei Grundschulkindern noch kein habitualisiertes oder persönlichkeitspezifisches Merkmal abbildet (Asendorpf 2005). Sie entwickelt sich sowohl durch das erhaltene leistungsbezogene Feedback, als auch durch soziale Vergleichsprozesse und die subjektiven Leistungs- und Erfolgsbewertungen der Lernenden (*appraisal*; Pekrun und Götz 2006).

Aufgrund der vorliegenden Studienlage zur Entwicklung der PA zum Ende der Grundschule bleibt unklar, ab wann sich die Ausprägung der PA stabilisiert und inwieweit in der Grundschule (a) vorherige Leistung die Entwicklung der PA, (b) die PA die Leistungsentwicklung beeinflusst, oder aber (c) sich keine bedeutsamen Unterschiede in der Stärke der Zusammenhänge zeigen. Dabei sind auch Besonderheiten des deutschen Schulsystems zu beachten, die in der internationalen Forschung keine oder andere Bewandnis haben: Da im vierten Schuljahr die Übergangsempfehlungen anstehen und somit oft der Leistungsdruck (Wahrnehmung einer hohen Valenz bei gleichzeitigem Gefühl geringer Kontrolle des Ergebnisses) durch die Eltern und/oder die Schule ansteigt, ist hier eher eine Zunahme der PA zu erwarten,

die nicht ausschließlich auf einen *appraisal* tatsächlicher Leistungsrückmeldungen zurückzuführen ist. Mit einer so begründeten Zunahme der PA wäre zu vermuten, dass weniger der Effekt der erbrachten Leistung auf die PA relevant ist, sondern vor allem die PA negative Effekte auf die zukünftige Leistungsentwicklung hat.

2.2 Zur Bedeutung des akademischen Selbstkonzepts im Kontext von Prüfungsangst

Das aSK gilt als bedeutendes motivationales Merkmal schulischer Leistung und hat, vermittelt über die Anstrengungsbereitschaft (Wigfield und Eccles 2000), positive Effekte auf die Leistungsentwicklung (Marsh 1990; Marsh und Craven 2006; Valentine et al. 2004). Für das aSK ist bekannt, dass zu Beginn der Schulzeit die Leistung das aSK beeinflusst (*skill-development* Ansatz), während später das aSK selbst ein Prädiktor akademischer Leistungen ist (*self-enhancement*) und sich im Laufe der Zeit reziproke Effekte festigen (Marsh und Craven 2006). Ab wann genau und inwiefern diese Verschiebungen bereits während der Grundschulzeit stattfinden, ist nicht eindeutig geklärt (z.B. Ehm et al. 2019; Helmke und van Aken 1995; Martschinke und Kammermeyer 2006). Welche Effekte sich nachweisen lassen, scheint laut Ehm und Kollegen (2019) vom methodologischen Ansatz sowie der Operationalisierung der Konstrukte und dem Abstand der Messzeitpunkte abzuhängen (Helmke und van Aken 1995). Als gesichert gilt jedoch ein korrelativer Zusammenhang beider Konstrukte bereits im Grundschulalter (Ehm et al. 2019; Faber 2012; Helmke und van Aken 1995; Kastens et al. 2013; Lipowsky et al. 2011). Selbst wenn das aSK im Primärbereich hauptsächlich im Sinne des *skill-development* Ansatzes auf Leistungsrückmeldungen aufbaut, gilt die Ausbildung eines positiv-realistischen aSK, gerade weil es sich im Laufe der Zeit stabilisiert (Trautwein und Möller 2016), als wünschenswert und zu Beginn der Schulzeit leistungsförderlich (vgl. auch Praetorius et al. 2016).

Verschiedene Studien konnten zeigen, dass Selbstkonzeptmaße negativ mit PA korreliert sind (Bandalos et al. 1995; Mammarella et al. 2018; Marsh 1990). In einer Metaanalyse berichtet Hembree (1988) einen mittleren Zusammenhang von $r = -0,42$ zwischen PA und Maßen zur Erfassung des *self-esteem* bei Grundschulkindern. Für die Skala *Besorgtheit* bzw. *Worry* („Beim Diktat muss ich oft daran denken wie viele Fehler ich mache“) berichtet Faber (2012) sogar einen Zusammenhang von $r = -0,67$ mit dem rechtschreibbezogenen aSK (siehe auch Schoreit 2016). Auf Individualebene scheinen ein hohes aSK und PA sich jedoch nicht zwangsläufig auszuschließen. So konnte Schoreit (2016) anhand latenter Klassenanalysen Schüler*innen identifizieren (ca. 20 % der untersuchten Stichprobe), die sich als kompetent einschätzten und sich dennoch ängstlich zeigten.

Fraglich ist, ob das aSK auch einen Einfluss auf die Beziehung zwischen PA und Leistung haben kann, oder sogar als Schutzfaktor fungiert. Nach Pekrun und Götz (2006) wäre zu erwarten, dass ein hohes aSK einerseits mit hohen Leistungen und hohem Kontrollerleben einhergeht und somit vor der Ausbildung von PA schützt. Andererseits können die Folgen von Misserfolgen für diese Schüler*innen eine hohe Valenz aufweisen, was die Entstehung von PA wiederum begünstigen würde. Vor allem aber könnte dies dazu führen, dass gerade bei diesen Kindern negative Effekte

auf die Leistung durch leistungshemmende Prozesse hervorgerufen werden könnten. Das sollte insbesondere dann der Fall sein, wenn die Leistungssituationen (hier die anstehende Übergangsempfehlung) mit hohen Erwartungen, bspw. von Seiten der Eltern, verbunden sind. Kinder mit hohem aSK – und hoher Wahrscheinlichkeit eine Gymnasialempfehlung zu erhalten – könnten dann anfällig für leistungshemmende Wirkungen der PA sein, weil der Leistung eine besonders hohe Wertigkeit beigemessen wird und Leistungssituationen auch als selbstwertbedrohlich wahrgenommen werden („Meine Leistungen in den letzten Klassenarbeiten sollten die Erwartungen, die an mich gestellt werden, bestätigen“). Bei Schüler*innen, die über geringe Kompetenzüberzeugungen verfügen, kann eine Leistungssituation hingegen an Valenz verlieren, das potenzielle Scheitern ist ohnehin absehbar. Die fehlende Valenz des Misserfolgs würde dann dazu führen, dass vorhandene Prüfungsängste nicht unbedingt mit einer zusätzlichen Reduktion der Leistung einhergehen.

Bisher dominieren hier variablenzentrierte Zugänge. Für die Rechtschreibleistung hat Faber (2012) bereits Hinweise auf einen moderierenden Effekt des rechtschreibbezogenen aSK auf die Ausprägung der Leistungsangst in Abhängigkeit der Leistungsstärke finden können. Martschinke und Kammermeyer (2006) konstatierten für die ersten beiden Grundschulklassen eine Mediatorfunktion von Emotionen über das aSK auf die mathematische Leistung. Dass das aSK die Beziehung zwischen Leistung und PA mediiert, zeigt beispielsweise auch die Studie von Bandalos et al. (1995).

Mammarella et al. (2018) haben einen personenzentrierten Zugang gewählt und mit Hilfe latenter Klassenanalysen Schüler*innen mit unterschiedlichen Angstprofilen identifiziert. Sie konnten zeigen, dass Schüler*innen mit so genannten *high-risk* Profilen u. a. signifikant geringere aSK aufwiesen. Werden aSK mit solchen personenzentrierten Ansätzen untersucht, zeigt sich, dass in erster Linie Niveauunterschiede ausschlaggebend für die Zuweisung zu latenten Gruppen sind. Die Entwicklung des aSK ist am Ende der Grundschulzeit bereits vergleichsweise stabil (Praetorius et al. 2016) und Effekte des aSK unterscheiden sich in erster Linie in Abhängigkeit davon, ob das aSK hoch oder gering ausgeprägt ist (Marsh et al. 2009). Mit dieser Studie wählen wir ebenfalls einen personenzentrierten Ansatz, mit dem wir Schüler*innen mit verschiedenen Ausprägungen des aSK identifizieren und prüfen, inwiefern diese Unterschiede Auswirkungen auf die Entwicklung der PA und den Zusammenhang zwischen PA und Leistung haben.

3 Forschungsfragen und Hypothesen

Anhand der dargestellten Forschungslage ergeben sich folgende Fragestellungen:

1 Wie stabil ist die PA gemessen vom Ende des dritten bis zum Ende des vierten Schuljahres?

Ausgehend von bisherigen Studien zur Entwicklung von PA erwarten wir eine mittlere korrelative Stabilität. Aufgrund der unklaren Befundlage zur Entwicklung der PA am Ende der Grundschulzeit bleibt offen, ob diese im untersuchten Zeitraum eher zu- oder abnimmt.

2 Gibt es einen reziproken Zusammenhang zwischen PA und Leistung im vierten Grundschuljahr oder ist eine Wirkungsrichtung dominant?

Kinder am Ende der Grundschulzeit haben bereits (negative) Leistungsrückmeldungen erhalten. Es ist also davon auszugehen, dass der Kreislauf von Rückkopplungsprozessen zwischen der Entwicklung der PA und der Leistung im vierten Schuljahr bereits in Gang gesetzt ist. Wir erwarten daher reziproke Zusammenhänge zwischen den beiden Variablen. Gleichzeitig ist jedoch anzunehmen, dass in diesem Schuljahr external begründete Leistungserwartungen den Leistungsdruck auf die Lernenden erhöhen und somit der negative Effekt der PA auf die Leistungsentwicklung stärker ausfallen könnte.

3 Welchen Effekt hat das aSK auf den Zusammenhang zwischen Prüfungsangst und Leistung am Ende des vierten Schuljahres?

Vor dem Hintergrund der *Kontroll-Wert-Theorie* der Leistungsemotionen (Pekrun und Götz 2006) stellt sich die Frage, inwiefern die Ausprägung kompetenzbezogener Überzeugungen Effekte auf die Zusammenhänge zwischen PA und Leistung haben. Wenngleich bei Kindern mit hohen Kompetenzüberzeugungen im Mittel geringere Werte bei der Ausprägung der PA zu erwarten sind, ist zu vermuten, dass Angst die kognitiven Fähigkeiten dieser Kinder in Leistungssituationen dennoch beeinträchtigt. Die Angst vor Versagen in Leistungssituationen könnte für diese Kinder eine besonders hohe Valenz haben, da Versagen gerade bei leistungsüberzeugten Kindern als selbstwertbedrohlich wahrgenommen werden könnte. Gleichzeitig könnte ein hohes aSK davor schützen, dass schlechte Leistungen dazu beitragen, PA zu steigern (siehe Abschn. 2.2).

4 Methodik

4.1 Stichprobe

Als Datengrundlage werden Daten der PERLE-Studie, aus dem dritten und vierten Schuljahr herangezogen (t1 und t2). Die Erhebung fand jeweils am Ende des Schuljahres in 38 Schulen in Sachsen, Berlin, Mecklenburg-Vorpommern und Thüringen statt. Bezüglich des sozialen Hintergrundes handelt es sich um eine positiv selektierte Stichprobe mit einem mittleren HISEI von $M=62,56$ ($SD=15,43$; $\min=20$; $\max=90$); der mittlere HISEI des IQB-Ländervergleichs liegt bei $M=48,5$; $SD=15,5$ (Richter et al. 2012). 61,6 % der Schüler*innen haben am Ende des vierten Schuljahres eine Gymnasial- und 23,7 % eine Realschulempfehlung erhalten. Zu den für unsere Analysen genutzten MZP liegen Daten von $N=725$ Kindern vor (50,90 % Mädchen). Vollständige längsschnittliche Daten für PA und Leistung in Klassen 3 und 4 sind von 622 Schüler*innen vorhanden. Bei der Erhebung im dritten Schuljahr waren diese durchschnittlich 9,3 Jahre alt, bei der zweiten Erhebung im vierten Schuljahr 10,3 Jahre alt.

4.2 Durchführung und Testinstrumente

Zur Beantwortung der oben genannten Forschungsfragen werden Daten herangezogen, die mit Schülerfragebögen erfasst wurden. Nach Angaben zur Messinvarianz (anhand der Kriterien von Chen 2007) ist die Annahme strenger Messinvarianz für PA und aSK zulässig. Eine Listung aller Items und Hinweise zur Datenerhebung finden sich im Anhang.

Zur Erfassung der *Prüfungsangst Mathematik* wurde die Skala von Schwarzer und Jerusalem (1999) verwendet. Diese besteht aus fünf Items. Das Antwortformat der Items war vierstufig (1 = ‚stimmt überhaupt nicht‘ bis 4 = ‚stimmt genau‘). Höhere Ausprägungen stellten dabei für alle Items höhere PA dar. Zwischen *Worry* und *Emotionality* wurde nicht unterschieden. Die Skala wies zu beiden MZP eine gute interne Konsistenz auf (Cronbachs Alpha: $t_1 = 0,85$; $t_2 = 0,88$). Die durchschnittliche Ausprägung liegt im dritten Schuljahr bei $M = 2,56$ ($SD = 0,94$) und bei $M = 2,47$ ($SD = 0,88$).

Mathematikleistung Die mathematische Leistung der Lernenden wurde mittels eines standardisierten Schulleistungstests jeweils am gleichen Tag wie die mathematischen Selbstkonzepte erfasst. Die Tests wurden im Rahmen der PERLE-Studie entwickelt und stellen Adaptionen der für die jeweilige Klassenstufe konzipierten DEMAT-Versionen (u. a. Krajewski et al. 2002) sowie des TEDI-MATH (Kaufmann et al. 2009) und des LEst 4–7 (Moser et al. 2004) dar. Darüber hinaus wurden eigenentwickelte Aufgaben ergänzt (Greb et al. 2011). Der Fokus der Tests lag zu allen MZP auf arithmetischen Aufgaben (n Items $t_1 = 33$; $t_2 = 18$). Der Skalierung liegt ein dichotomes Rasch-Modell zugrunde, die Mittelwerte der WLE-Schätzer sind zu $t_1 = 3,99$ ($SD = 1,37$) und zu $t_2 = 5,23$ ($SD = 1,36$). Zur längsschnittlichen Skalierung der Skala dienten 18 Items aus t_1 als Ankeritems, die Reliabilitäten der Skala sind gut (t_1 : EAP/PV = 0,71; t_2 : EAP/PV = 0,60).

Mathematisches Fähigkeitsselbstkonzept Im Rahmen der PERLE-Studie wurde ein Fragebogen entwickelt, der es erlauben sollte, das aSK von Beginn bis Ende der Grundschulzeit längsschnittlich zu erfassen (s. Greb et al. 2011). Der Fokus lag dabei auf kognitiven Komponenten des aSK. Das Antwortformat der Items war dreistufig (1 = ‚nicht so gut‘, 2 = ‚gut‘, 3 = ‚sehr gut‘, für das genannte Beispielitem). Höhere Ausprägungen stellten dabei für alle Items höhere Selbstkonzeptausprägungen dar. Die interne Konsistenz des Konstruktes ist zu beiden Messzeitpunkten gut (Cronbachs Alpha: $t_1 = 0,89$; $t_2 = 0,90$). Der Skalenmittelwert liegt Ende des dritten Schuljahres bei $M = 2,24$ ($SD = 0,47$), Ende des vierten Schuljahres bei $M = 2,23$ ($SD = 0,48$). Die korrelative Stabilität des aSK liegt bei $r = 0,71$; $p < 0,01$.

4.3 Analysen

Um reziproke Zusammenhänge zwischen der PA und der Leistung zu untersuchen, wurden beide Variablen mit einem Cross-Lagged-Panel (CLP) in ein längsschnittliches Verhältnis gesetzt (Selig und Little 2012). So konnte neben den Zusammenhängen beider Konstrukte auch die Stabilität (als Autoregression) zwischen den MZP

(Ende drittes Schuljahr und Ende viertes Schuljahr) aufgezeigt werden. Die Berechnung des Designs wurde mit dem Softwarepaket Mplus 8.4 (Muthén und Muthén 2012) vorgenommen. Die Skalen zur PA und dem aSK wurden z-standardisiert.

In der vorliegenden Auswertung beträgt der Anteil fehlender Werte auf Itemebene (häufig als Indikator für systematische Fehler herangezogen, siehe auch Lüdtke et al. 2007) weniger als 2 %. Lediglich auf Klassenebene ergeben sich höhere Werte. Fehlende Werte auf Klassenebene sind auf Nichtteilnahme einzelner Kinder oder ganzer Klassen (wegen organisatorischer Hindernisse oder Krankheit) zurückzuführen. Daher gehen wir von zufälligen Effekten (MAR, *Missing At Random*) zur Entstehung der fehlenden Werte aus. Für den Umgang mit fehlenden Werten wurde die Full Information Maximum Likelihood (FIML)-Methode gewählt (siehe Lüdtke et al. 2007). Für alle Variablen liegen für Individuen (Schüler*innen) in Cluster (Klassen) geschachtelte Daten vor, es wurden alle CLP Modelle mit einem MLR Schätzer unter der Option *type = complex* spezifiziert.

Effekte des aSK auf den Zusammenhang zwischen PA und Leistung sind insbesondere für unterschiedliche Niveaueausprägungen des aSK zu erwarten (siehe Abschn. 3). Um sowohl Zusammenhänge zwischen PA und Leistung in Abhängigkeit der Ausprägung des aSK längsschnittlich zu untersuchen und Unterschiede in der Entwicklung des aSK im vierten Schuljahr simultan berücksichtigen zu können, greifen wir auf den personenzentrierten Ansatz der latenten Profilanalyse (LPA) zurück. Dieses Verfahren erlaubt es, Gruppen von Schüler*innen mit unterschiedlichen Ausprägungen ihres aSK am Ende des dritten und vierten Schuljahres zu identifizieren. Bei einer korrelativen Stabilität des aSK im vierten Schuljahr von $r = 0,71$ ist anzunehmen, dass sich Schüler*innen vor allem in der intraindividuellen Ausprägung ihrer aSK zu den untersuchten Messzeitpunkten unterscheiden (vgl. Marsh et al. 2009). Unklar ist, ob sich auch unterschiedliche Entwicklungen in Abhängigkeit der Ausprägung zeigen.

LPAs stellen eine Sonderform latenter Klassenanalysen dar und eignen sich zur Ermittlung latenter Profilklassen bei kontinuierlichen Variablen (DiStefano 2012; Nylund-Gibson und Choi 2018). Die Annahme ist, dass die Ausprägungen eines Merkmals durch die Zugehörigkeit zu einer Klasse erklärt werden. Die Klassen spiegeln in diesem Fall Gruppen von Schüler*innen wieder, deren aSK-Ausprägung vom Ende des dritten Schuljahres bis zum Ende des vierten Schuljahres innerhalb einer Gruppe eine geringere Varianz (oder höhere Ähnlichkeit) aufweisen als zwischen Gruppen (longitudinal LPA; siehe Brinkman et al. 2013). Wir treffen damit keine Aussagen über Verlaufsformen, Kovariationen oder Stärken von Veränderung und nutzen die Gruppenzuweisungen als deskriptives Maß. Der Vorteil dieses personenzentrierten Verfahrens gegenüber clusteranalytischen Verfahren ist, dass sowohl die Zuordnung einer Person zu einer Gruppe als auch die Bestimmung der Anzahl extrahierter Gruppen auf einem Schätzverfahren basiert, dessen Ergebnisgüte anhand verschiedener Parameter darstellbar ist (DiStefano 2012; Ferguson et al. 2019). Da die in den Modellen als Indikatoren aufgenommen Variablen (aSK Ende drittes und Ende viertes Schuljahr) wie auch bei querschnittlichen Daten häufig der Fall, nicht unabhängig voneinander sind, wurden über den Befehl „%overall%“ Korrelationen zwischen eben diesen für alle extrahierten Klassen spezifiziert (siehe

Marsh et al. 2009). Die LPAs wurden in MPlus 8.4 durchgeführt, die Ergebnisse sind in Abschn. 5.3 dargestellt.

5 Ergebnisse

5.1 Entwicklung und Stabilität von mathematischer Prüfungsangst

Insgesamt weisen im dritten Schuljahr 53,3 % der befragten Lernenden und 49,5 % im vierten Schuljahr einen Wert oberhalb des theoretischen Mittels von 2,5 auf. Eine Varianzanalyse mit Messwiederholung zeigt, dass die überdurchschnittlich ausgeprägte PA Ende des dritten Schuljahres ($M=2,56$; $SD=0,94$) zum Ende des vierten Schuljahres ($M=2,47$; $SD=0,88$) sinkt; $F(1; 510)=8,44$; $p<0,05$. Dieser Abfall ist jedoch mit $\eta^2=0,02$ als eher schwach zu bezeichnen. Zudem können wir eine mittlere bis hohe korrelative Stabilität der PA ($r=0,56$) verzeichnen.

5.2 Wirkungszusammenhang von mathematischer Prüfungsangst und Leistung über ein Schuljahr

Zu beiden Messzeitpunkten zeigen sich negative Korrelationen zwischen PA und Leistung von $r=-0,30$; $p<0,05$ für das dritte und $r=-0,32$; $p<0,05$ für das vierte Schuljahr. Das CLP (Abb. 1) zeigt, dass der Effekt der PA auf die Leistung mit $\beta=-0,21$ wie erwartet etwas stärker ausfällt, als der Effekt der Leistung auf die PA: $\beta=-0,14$. Allerdings ist der Unterschied zwischen den Kreuzpfaden statistisch nicht bedeutsam. Durch das Gleichsetzen der Kreuzpfade verschlechtert sich der Modellfit nicht ($\Delta\text{Chi}^2/\text{df}=1,29/1$; $\Delta\text{RMSEA}=0,00$; $\Delta\text{CFI}=0,00$; $\Delta\text{TLI}=0,00$;

Abb. 1 Cross-Lagged-Panel Modell von Prüfungsangst und Leistung Ende des dritten Schuljahres (t1) und Ende des vierten Schuljahres (t2). Abgebildet sind bivariate Korrelationen und Pfadkoeffizienten mit Standardfehlern in Klammern. ($N=622$; *n. s.* nicht signifikant. Modellfit für das restringierte Modell (strikte Invarianz): $\text{Chi}^2/\text{df}=185,050/59$; $\text{RMSEA}=0,059$; $\text{CFI}=0,949$; $\text{TLI}=0,943$; $\text{SRMR}=0,037$; Modellfit für Modell ohne Restriktionen: $\text{Chi}^2/\text{df}=178,025/50$; $\text{RMSEA}=0,064$; $\text{CFI}=0,948$; $\text{TLI}=0,932$; $\text{SRMR}=0,036$. Chi-Quadrat Differenzentest für MLR Schätzer: $\text{Chi}^2/\text{df}=15,013/9$; $p=0,091$)

$\Delta\text{SRMR}=0,00$). PA und Leistung bedingen sich für die Gesamtstichprobe längsschnittlich reziprok: Lernende, die am Ende des dritten Schuljahres geringe mathematische Leistung aufweisen, entwickeln im vierten Schuljahr eher eine PA als Lernende mit höheren Leistungen. Gleichzeitig hat eine ausgeprägte PA am Ende des dritten Schuljahres negative Effekte auf die Leistungsentwicklung.

5.3 Effekte des aSK auf den Zusammenhang zwischen Prüfungsangst und Leistung

Als bedeutsame Parameter zur Bestimmung der Modellgüte der latenten Profilanalysemodelle lassen sich der AIC, BIC und adjustierte BIC sowie der Lo-Mendell-Rubin (LMR) Test heranziehen. Das Entropy Maß gibt ergänzend die Güte der Klassenzuweisung an: Werte um 1 deuten auf eine große, Werte nahe 0 auf eine sehr geringe Klassifikationsgüte hin (Ferguson et al. 2019; Geiser 2011).

Verglichen werden die Parameter für drei Modelle mit 2- bis 5-Klassen. Die Ergebnisse dieser Analysen sind in Tab. 1 dargestellt.

Während der adjustierte LMR likelihood Test für alle Modelle mit mehr als zwei Klassen signifikant wird, zeigen sich für deskriptive Indices des Modellvergleichs (AIC, BIC und aBIC) kaum Unterschiede zwischen der 3- oder 4-Klassenlösung. Trotz hoher Startwertesets von 5000 und 500 und 50 Iterationen, die zur Modellschätzung angesetzt wurden, lassen unterschiedliche loglikelihood-Werte für eine 4- und 5-Klassenlösung jedoch das Vorhandensein lokaler Maxima vermuten. Dieser Umstand und auch die Betrachtung der anderen Modellparameter lassen den Schluss zu, dass eine 3-Klassenlösung die robustesten Schätzer der Klassenzugehörigkeit liefert. Auch das Entropy Maß ist für diese Lösung am höchsten. In Tab. 2 sind die Wahrscheinlichkeiten für die Klassifikation der Personen und die Anzahl

Tab. 1 Modellparameter der LPA Lösungen

	AIC	BIC	aBIC	<i>p</i> LMR	Entropy
2-Klassenlösung	2866,402	2901,878	2876,479	0,209	0,425
3-Klassenlösung	2756,904	2805,684	2770,761	0,000	0,854
4-Klassenlösung	2742,367	2804,451	2760,003	0,016	0,815
5-Klassenlösung	2724,784	2800,171	2746,199	0,010	0,845

AIC Akaike's Information Criterion, BIC Bayesian Information Criterion, aBIC sample size adjusted BIC, *p* LMR *p*-Werte des Lo-Mendell-Rubin adjusted likelihood tests

Tab. 2 Wahrscheinlichkeiten zur Güte der Klassenzuordnung, Anzahl an Fällen und Mädchen für jede der drei in der LPA ermittelten Klassen

	<i>p</i> Classification	<i>N</i> Fälle	<i>N</i> Mädchen
Klasse 1	0,847	58 (10%)	44 (14%)
Klasse 2	0,941	361 (57%)	194 (62%)
Klasse 3	0,943	204 (33%)	75 (24%)

Die Unterschiede zwischen den Gruppen in der Verteilung des Geschlechts sind statistisch bedeutsam $\chi^2/\text{df}=30,86/2$
 $p<0,01$

Tab. 3 Deskriptive Kennwerte, korrelative Stabilität innerhalb der Klassen und Ergebnisse einer MANOVA zur Prüfung von Klassenunterschieden im Querschnitt und Varianzanalyse mit Messwiederholung zwischen Klassenzugehörigkeit und Messzeitpunkt, für alle drei Variablen

	Klasse 1	Klasse 2	Klasse 3	MANOVA	
	M (SD)	M (SD)	M (SD)	<i>F</i> (2,486)	η^2
<i>Selbstkonzept</i>					
t1	-1,37 (0,78)	-0,22 (0,75)	0,76 (0,78)	156,18*	0,39
t2	-1,96 (0,39)	-0,34 (0,37)	1,16 (0,40)	1414,83*	0,85
Zeit	<i>F</i> (1, 512)=6,80*; $\eta^2=0,01$			—	—
Klasse $\times$ Zeit	<i>F</i> (2, 512)=49,52*; $\eta^2=0,16$			—	—
<i>r</i> _{t1t2}	0,20	0,40*	0,55*	—	—
<i>Prüfungsangst</i>					
t1	0,51 (0,88)	0,18 (0,91)	-0,47 (1,00)	32,36*	0,12
t2	0,79 (0,76)	0,17 (0,87)	-0,59 (0,97)	56,21*	0,19
Zeit	<i>F</i> (1, 508)=0,79, $\eta^2=0,00$			—	—
Klasse $\times$ Zeit	<i>F</i> (2, 508)=3,45*, $\eta^2=0,01$			—	—
<i>r</i> _{t1t2}	0,26	0,50*	0,49*	—	—
<i>Leistung</i>					
t1	-1,00 (0,83)	-0,15 (0,82)	0,69 (0,87)	81,72*	0,25
t2	-0,86 (0,84)	-0,16 (0,83)	0,64 (0,97)	62,65*	0,20
Zeit	<i>F</i> (1, 491)=0,18, $\eta^2=0,00$			—	—
Klasse $\times$ Zeit	<i>F</i> (2, 491)=0,69, $\eta^2=0,00$			—	—
<i>r</i> _{t1t2}	0,43*	0,44*	0,63*	—	—

* $p < 0,01$

der Fälle pro Klasse für die 3-Klassenlösung dargestellt. In Klasse 2 befinden sich mit 57 % etwas mehr als die Hälfte der Schüler*innen.

Wie erwartet beschreiben die drei Klassen wie bei Marsh et al. (2009), Schüler*innen mit unterschiedlichen Niveaueausprägungen des aSK: niedrige, mittlere und hohe aSK (siehe Tab. 3). Schüler*innen in Klasse 1 (geringes aSK zu beiden Messzeitpunkten) weisen geringere Leistungen und höhere PA auf, in Klasse 2 (mittleres aSK zu beiden Messzeitpunkten) sind die Schüler*innen mit mäßiger PA und mittelmäßigen Leistungen geclustert, während sich in Klasse 3 (hohes aSK zu beiden Messzeitpunkten) leistungsstarke Schüler*innen mit geringer PA befinden. Für alle drei Variablen zeigt eine MANOVA, dass diese Unterschiede in den mittleren Ausprägungen zwischen den Klassen zu beiden Messzeitpunkten signifikant sind. Post-Hoc Bonferroni Vergleiche weisen lediglich für die PA am Ende des dritten Schuljahres keinen signifikanten Unterschied zwischen Klasse 1 und Klasse 2 auf (mittlere Differenz=0,29; $SE=0,16$; $p=0,20$).

ANOVAs mit Messwiederholung und dem Faktor Zeit zeigen für das aSK signifikante Interaktionseffekte, die insbesondere auf den deutlichen Abfall des aSK in Klasse 1 und deutlichen Anstieg des aSK in Klasse 3 über die Zeit zurückzuführen sind (für alle post-hoc Bonferroni korrigierten Kontrastvergleiche $p < 0,00$). Schüler*innen, die am Ende des dritten Schuljahres bereits unterdurchschnittliche aSK aufwiesen, berichten am Ende der Grundschulzeit nochmals reduzierte aSK, für Schüler*innen mit hohen aSK finden sich gegenteilige mittlere Verläufe. Wie

Tab. 4 Pfadkoeffizienten des Cross-Lagged-Panels für alle drei Klassen

	L3 → PA4	PA3 → L4	PA3 ↔ L3	PA4 ↔ L4	PA3 → PA4	L3 → L4
Klasse 1	−0,18*	0,02	0,03	0,03	0,15	0,43*
Klasse 2	0,00	−0,08	−0,10*	−0,00	0,61*	0,44*
Klasse 3	−0,10	−0,28*	−0,12 ⁺	−0,02	0,54*	0,65*

Model Fit $\chi^2/\text{df} = 423,210^*/191$

CFI = 0,909

TLI = 0,906

SRMR = 0,075

RMSEA = 0,077

* $p < 0,01$; ⁺ $p < 0,05$

vermutet, zeigt sich kein signifikanter Haupteffekt für die PA über die Zeit, jedoch ein signifikanter Interaktionseffekt ($\eta^2 = 0,01$): Schüler*innen mit geringen aSK zu beiden Messzeitpunkten (Klasse 1) weisen eine Zunahme der PA zum Ende der Grundschulzeit auf, für Schüler*innen mit hohen aSK (Klasse 3) zeigt sich ein Abfall der PA im vierten Schuljahr (für alle post-hoc Bonferroni korrigierten Kontrastvergleiche $p < 0,00$). Die PA für Schüler*innen mit mittleren aSK (Klasse 2) verändert sich nicht. Es bestehen keine Unterschiede zwischen den Klassen in den mittleren Verläufen der Leistungsentwicklung, weder der Haupteffekt noch die Interaktion der Varianzanalyse mit Messwiederholung sind signifikant. Die korrelative Stabilität des aSK und der PA ist in Klasse 1 besonders gering mit $r = 0,20$ und $0,26$. Für die Klassen 2 und 3 zeigen sich höhere korrelative Stabilitäten aller drei Variablen mit $r > 0,40$.

Werden die Zusammenhänge zwischen der Entwicklung der PA und der Leistung in einem Mehrgruppen Cross-Lagged-Panel berechnet (Tab. 4), lassen sich die zuvor berichteten reziproken Effekte zwischen PA und Leistung nicht nachweisen. Für Kinder mit geringem aSK (Klasse 1) hat die Leistung Ende des dritten Schuljahres einen geringen, aber signifikanten negativen Effekt ($\beta = -0,18$) auf die PA am Ende des vierten Schuljahres. Es zeigt sich kein Effekt der PA auf die Leistung. Für Kinder mit hohen aSK (Klasse 3) zeigt sich der umgekehrte Effekt: PA am Ende des dritten Schuljahres ist ein bedeutsamer negativer Prädiktor ($\beta = -0,28$) für die Leistung Ende des vierten Schuljahres. Für Klasse 2 hingegen (Kinder mit mittleren aSK), zeigen sich gar keine Zusammenhänge zwischen PA und Leistung.

6 Diskussion

6.1 Zusammenfassung und Interpretation der Ergebnisse

Die PA wies in dieser Studie für die Gesamtstichprobe eine mittlere bis hohe korrelative Stabilität auf, es ließ sich ein leichter mittlerer Abfall der PA nachweisen und es zeigten sich negative reziproke Zusammenhänge mit PA und Leistung.

Die Ergebnisse der LPA zeigen, dass sich drei Gruppen von Schüler*innen basierend auf der Ausprägung ihrer aSK zu zwei benachbarten Messzeitpunkten unterscheiden lassen: Schüler*innen mit zu beiden Zeitpunkten unterdurchschnittlichem

aSK, welches innerhalb des vierten Schuljahres abfällt, Schüler*innen, die mittlere aSK zu beiden Zeitpunkten aufweisen und Schüler*innen mit überdurchschnittlich ausgeprägtem aSK, welches im Verlauf des vierten Schuljahres weiter zunimmt. Die Relevanz einer personenzentrierten Betrachtung wird deutlich, wenn man sich die klassenbezogenen Analysen zum Zusammenhang und der Entwicklung von PA und Leistungen ansieht: Eine Zunahme der PA ist vor allem für Schüler*innen mit zu beiden Zeitpunkten geringen aSK (Klasse 1) zu beobachten, während sie für diejenigen mit zu beiden Zeitpunkten hohen aSK (Klasse 3) sogar sinkt. Auch die Zusammenhänge zwischen PA und Leistung unterscheiden sich zwischen Schüler*innen mit hohen, mittleren und geringeren aSK und zeigen überraschende Ergebnisse: Während sich für Schüler*innen mit zu beiden Messzeitpunkten mittleren aSK (Klasse 2) keine Zusammenhänge zwischen PA und Leistung am Ende der Grundschulzeit zeigen, lassen sich für Schüler*innen mit hohen aSK negative Effekte der PA am Ende des dritten Schuljahres auf die Leistungsentwicklung nachweisen. Für Schüler*innen in Klasse 1 hingegen finden sich negative Effekte der Leistung auf die PA. Während bei Schüler*innen mit geringen Kompetenzüberzeugungen PA nicht unbedingt zu einer zusätzlichen Beeinträchtigung ihrer Leistungsentwicklung führt, zeigen sich für kompetenzüberzeugte Schüler*innen, dass das aSK nicht etwa zu einer höheren Resilienz beiträgt oder als Schutzfaktor fungiert, sondern Prüfungsangst hier bedeutsame negative Auswirkungen auf die Leistungsentwicklung haben kann. Hier zeigen sich, wenn auch nur bedingt vergleichbare, Parallelen zur Studie von Mammarella et al. (2018). Dort lag der Fokus auf der Entwicklung von Angstprofilen, welche zu einem Messzeitpunkt gebildet wurden, während hier nach aSK über zwei Messzeitpunkte (aufgrund der Stabilität des Merkmals) gruppiert wurde. Durch den unterschiedlichen Fokus zeigen sich bei uns differenziertere Ergebnisse: Nämlich, dass das aSK nicht grundsätzlich ein Protektor vor PA ist, wie bei Mammarella und Kolleg*innen (siehe auch Schoreit 2016).

Nach der *Kontroll-Wert-Theorie* zu Bedingungen und Wirkungen von PA (Pekrun und Götz 2006) könnten zur Erklärung dieser Befunde die Valenz von Misserfolg eine bedeutsame Rolle spielen: Leistungsstarke Schüler*innen mit hohen Kompetenzüberzeugungen haben zwar weniger Angst, fürchten aber möglicherweise Misserfolg stärker. Denkbare wäre zudem, dass unsere Befunde spezifischen, mit dem Erhalt von Übergangsempfehlungen verbundenen, Prozessen geschuldet sind: Schüler*innen mit hohem aSK und hohen Leistungen sollten zwar insgesamt eine hohe Kontrollierbarkeit in Leistungssituationen wahrnehmen, andererseits sind dies möglicherweise auch die Kinder, die hohe Erwartungen an sich selbst stellen oder an die mit Erhalt einer Gymnasialempfehlung hohe Ansprüche für den weiteren Schulverlauf kommuniziert werden (siehe Kastens und van Wickeren in Vorbereitung). Ist eine Leistungssituation jedoch mit hoher Valenz verbunden, ist auch die Voraussetzung für PA geschaffen (Pekrun und Götz 2006). So wäre denkbar, dass die Leistungen in Klassenarbeiten, die im untersuchten Zeitraum geschrieben werden, nicht nur an bisherigen Leistungen, sondern auch an der erteilten Übergangsempfehlung gemessen werden. Die Angst davor, dass es nicht gelingt, die an sie gestellten Erwartungen zu bestätigen, wären dann für diese Schüler*innen größer. In diesen Fällen ist die leistungshemmende Wirkung der PA dann bedeutsam (*skill-impairment*), während sich für Schüler*innen mit geringen aSK eine geringere Valenz

der Übergangsempfehlung ergibt und dann geringe Leistungen eher Unsicherheit verstärken und negative Effekte auf die PA aufweisen (*self-decrease*).

Die Befunde unserer Studie können als Hinweis darauf gesehen werden, dass die Wirkrichtung zwischen PA und Leistung durch die Wahrnehmung von Leistungserwartungen oder Bewertungsangst beeinflusst werden. Sie zeigen zudem, dass personenzentrierte Ansätze wie die hier eingesetzte LPA nicht nur weitere Erklärungen zum Zusammenspiel von Leistung, PA und aSK liefern, sondern auch neue Fragen aufwerfen.

6.2 Grenzen und Ausblick

Die vorliegenden Analysen bieten durch das längsschnittliche Design die Möglichkeit, bislang nicht untersuchte Wandlungsprozesse zwischen PA, Leistung und aSK am Ende der Grundschulzeit zu betrachten. Zudem konnten wir zeigen, dass die Beachtung von Niveauunterschieden in der Ausprägung des aSK bedeutsam ist, um Unterschiede in der Entwicklung der PA zu erklären (siehe Marsh et al. 2009; Praetorius et al. 2016; Trautwein und Möller 2016).

Eine Einschränkung unserer Studie ist, dass bei der Operationalisierung der PA nicht zwischen *Worry* und *Emotionality* differenziert wird, obwohl sich beide Facetten phänomenologisch voneinander unterscheiden und sie unterschiedlich mit Leistung zusammenhängen (Hembree 1990; Sparfeldt et al. 2005; Zeidner 1998). Ferner hätte ein weiterer Messzeitpunkt zum Schulhalbjahr des vierten Schuljahrs die Möglichkeit geboten, anhand latenter Wachstumskurvenmodelle bspw. mögliche nicht-lineare Verläufe der untersuchten Variablen in diesem Schuljahr zu untersuchen und auch Gruppen mit unterschiedlichen Entwicklungsverläufen über die Zeit zu identifizieren (Kastens et al. 2013). Konkret stellt sich die Frage, ob der Erhalt des Halbjahreszeugnisses mit der Übergangsempfehlung einen Einfluss auf die Entwicklung der PA hat. Auch wenn Übergangsempfehlungen in den meisten Bundesländern nicht mehr bindend sind, stellen sie eine bedeutsame Leistungsrückmeldung an die Schüler*innen dar, oftmals verbunden mit erhöhten Leistungsanforderung in der Schule und elterlichen Erwartungshaltungen (Pixner und Kaufmann 2013; Valtin und Wagner 2004; Kastens und van Wickeren in Vorbereitung).

Anzumerken ist an dieser Stelle auch, dass es an zusätzlichen Maßen zur Erfassung von bspw. Valenzen, Attributionsmustern oder Zielorientierungen der Schüler*innen fehlt. Denkbar wäre, dass Entwicklungen der PA und deren Zusammenhänge mit der Leistungsentwicklung neben der Ausprägung des aSK auch davon abhängig sind, ob die Schüler*innen leistungs- oder lernzielorientiert sind (Putwain und Daniels 2010). Wir stellen mit dieser Studie einige Überlegungen an, wie differentielle Effekte der im Modell von Pekrun und Götz (2006) genannten Prozesse zur Entstehung von PA für den Übergang erklärbar sein könnten. Weiterführende Studien könnten diese anhand einer Stichprobe mit größerer Heterogenität bezüglich der erhaltenen Übergangsempfehlung, sozioökonomischem Hintergrund und Migrationshintergrund und unter Berücksichtigung der Bildungsaspirationen der Eltern überprüfen.

Die Ergebnisse haben eine hohe praktische Relevanz. Lehrkräften sollte bewusst sein, dass hohe und niedrige Fähigkeitsüberzeugungen in Leistungssituationen un-

terschiedliche Erwartungshaltungen verursachen können und damit sowohl selbstwertdienliche, als auch selbstwertschädliche Effekte haben können. Während Schüler*innen mit hohen Fähigkeitsüberzeugungen Leistungssituationen einerseits besser bewältigen und kontrollieren können sollten, zeigen die Ergebnisse dieser Studie auf, dass für manche dieser Kinder solche auch mit negativen Effekten auf die Leistungsentwicklung verbunden sein können, indem bspw. PA den Abruf oder die Vorbereitung auf Klassenarbeiten behindern. Bei Schüler*innen mit geringen aSK besteht hingegen die Gefahr der Entstehung einer Negativspirale zwischen PA und Leistungsentwicklung. Idealerweise gelingt es der Grundschule, wichtige Weichenstellungen für die weitere Bildungslaufbahn von Schüler*innen auch dadurch zu setzen, dass sie Schüler*innen darin unterstützt, nicht durch ungünstige motivationale und emotionale Faktoren in ihrer Lernentwicklung behindert zu werden.

Anhang

Um möglichst gleiche Testbedingungen zu schaffen, wurden die Lernenden mit Hilfe verschiedener Maßnahmen durch den Fragebogen begleitet: Zu Beginn der Befragung wurde das Prozedere anhand eines Testitems erörtert. Im Folgenden wurden alle Items und Antwortmöglichkeiten (mit Ausnahme der Aufgaben im Leistungstest) durch die Testleiter vorgelesen. Um sicherzustellen, dass sich die Lernenden während der Testsituation auf der richtigen Seite des Fragebogens befinden, waren dessen Seiten mit Tierbildern gekennzeichnet und nach jeder Doppelseite mit einem Stoppschild versehen. So konnten die Testleiter immer angeben, auf welcher Seite sie sich befanden, wann gemeinsam umgeblättert wurde und es war gewährleistet, dass die Kinder keine Seiten überspringen.

Tab. 5 Items Skala akademisches Selbstkonzept (Mathematik)

„Wie denkst du über dich selbst?“

Wie gut bist du beim Rechnen?	1 = nicht so gut; 2 = gut; 3 = sehr gut
Fällt dir das Rechnen leicht oder schwer?	1 = schwerer; 2 = weder leichter noch schwerer; 3 = leichter
Wie viele Fehler machst du beim Rechnen?	1 = fast keine; 2 = ein paar; 3 = viele
Wie oft weißt du bei Rechenaufgaben die richtige Antwort?	1 = nicht so oft; 2 = oft; 3 = immer
Wie gut kannst du Zahlen zusammenzählen?	1 = nicht so gut; 2 = gut; 3 = sehr gut

Tab. 6 Items Skala Prüfungsangst

„Wenn wir in Mathematik eine Klassenarbeit schreiben ...“

... spüre ich ein komisches Gefühl im Magen
... mache ich mir Sorgen, ob ich auch alles schaffe
... frage ich mich, ob ich genug gelernt habe
... schlägt mein Herz mir bis zum Hals
... habe ich Angst

Tab. 7 Kennwerte Invarianzanalysen

	Chi ²	Df	CFI	RMSEA	SRMR
<i>Akademisches Selbstkonzept</i>					
Modell 0: ohne Restriktionen	96,056	47	0,988	0,041	0,027
Modell 1: Metrische Invarianz	98,909	52	0,989	0,038	0,030
Modell 2: Strenge Invarianz	110,859	57	0,987	0,039	0,030
<i>Prüfungsangst</i>					
Modell 0: ohne Restriktionen	111,684	29	0,970	0,068	0,028
Modell 1: Metrische Invarianz	114,155	33	0,971	0,063	0,030
Modell 2: Strenge Invarianz	123,196	37	0,969	0,061	0,032

Nach Chen (2007) gilt strenge Invarianz als gegeben, bei Veränderungen von 0,010 des CFI, 0,015 des RMSEA oder von 0,010 des SRMR zwischen den beiden Modellen

Danksagung Das Autorenteam bedankt sich herzlich bei Martin van Wickeren für wertvolle und konstruktive Rückmeldungen zu vorherigen Versionen des Manuskripts.

Funding Open Access funding enabled and organized by Projekt DEAL.

Open Access Dieser Artikel wird unter der Creative Commons Namensnennung 4.0 International Lizenz veröffentlicht, welche die Nutzung, Vervielfältigung, Bearbeitung, Verbreitung und Wiedergabe in jeglichem Medium und Format erlaubt, sofern Sie den/die ursprünglichen Autor(en) und die Quelle ordnungsgemäß nennen, einen Link zur Creative Commons Lizenz beifügen und angeben, ob Änderungen vorgenommen wurden.

Die in diesem Artikel enthaltenen Bilder und sonstiges Drittmaterial unterliegen ebenfalls der genannten Creative Commons Lizenz, sofern sich aus der Abbildungslegende nichts anderes ergibt. Sofern das betreffende Material nicht unter der genannten Creative Commons Lizenz steht und die betreffende Handlung nicht nach gesetzlichen Vorschriften erlaubt ist, ist für die oben aufgeführten Weiterverwendungen des Materials die Einwilligung des jeweiligen Rechteinhabers einzuholen.

Weitere Details zur Lizenz entnehmen Sie bitte der Lizenzinformation auf <http://creativecommons.org/licenses/by/4.0/deed.de>.

Literatur

- Asendorpf, J. (2005). Persönlichkeit: Stabilität und Veränderung. In H. Weber & T. Rammsayer (Hrsg.), *Handbuch der Persönlichkeitspsychologie und differentiellen Psychologie* (S. 15–27). Göttingen: Hogrefe.
- Bandalos, D.L., Yates, K., & Thorndike-Christ, T. (1995). Effects of math self-concept, perceived self-efficacy, and attributions for failure and success on test anxiety. *Journal of Educational Psychology*, 87(4), 611–623.
- Brinkman, T.M., Zhu, L., Zeltzer, L.K., Recklitis, C.J., Kimberg, C., Zhang, N., Muriel, A.C., Stovall, M., Srivastava, D.K., Robison, L.L., & Krull, K.R. (2013). Longitudinal patterns of psychological distress in adult survivors of childhood cancer. *British Journal of Cancer*, 109(5), 1373–1381.
- Carey, E., Hill, F., Devine, A., & Szűcs, D. (2016). The chicken or the egg? The direction of the relationship between mathematics anxiety and mathematics performance. *Frontiers in Psychology*, 6, 1–6.
- Chen, F.F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling*, 14(3), 464–504.
- Cortina, K. (2008). Leistungsängstlichkeit. In W. Schneider & M. Hasselhorn (Hrsg.), *Handbuch der Pädagogischen Psychologie* (S. 50–61). Göttingen: Hogrefe.
- DiStefano, C. (2012). Cluster analysis and latent class clustering techniques. In B. Laursen, T.D. Little & N.A. Card (Hrsg.), *Handbook of developmental research methods* (S. 645–666). New York: Guilford.

- Eden, C., Heine, A., & Jacobs, A. M. (2013). Mathematics anxiety and its development in the course of formal schooling—A review. *Psychology*, 4(6), 27–35.
- Ehm, J.-H., Hasselhorn, M., & Schmiedek, F. (2019). Analyzing the developmental relation of academic self-concept and achievement in elementary school children: Alternative models point to different results. *Developmental Psychology*, 55(11), 2336–2351.
- Faber, G. (2012). *Selbstkonzept, Kausalattributionen und Leistungsangst im Rechtschreiben. Schulfachspezifische Analysen zu den Selbsteinschätzungen von Grundschulkindern* (Bd. 2). Saarbrücken: AV Akademikerverlag.
- Ferguson, S. L., Moore, E. W. G., & Hull, D. M. (2019). Finding latent groups in observed data: a primer on latent profile analysis in Mplus for applied researchers. *International Journal of Behavioral Development*. <https://doi.org/10.1177/0165025419881721>.
- Geiser, C. (2011). *Datenanalyse mit Mplus. Eine anwendungsorientierte Einführung*. Wiesbaden: VS.
- Helmke, A. (1997). Individuelle Bedingungsfaktoren der Schulleistung: Ergebnisse aus dem SCHOLASTIK-Projekt. In F. E. Weinert & A. Helmke (Hrsg.), *Entwicklung im Grundschulalter* (S. 203–216). Weinheim: Psychologie Verlags Union.
- Helmke, A., & van Aken, M. A. G. (1995). The causal ordering of academic achievement and self-concept of ability during elementary school: A longitudinal study. *Journal of Educational Psychology*, 87(4), 624–637.
- Hembree, R. (1988). Correlates, causes, effects and treatment of test anxiety. *Review of Educational Research*, 58(1), 47–77.
- Hembree, R. (1990). The nature, effects, and relief of mathematics anxiety. *Journal for Research in Mathematics Education*, 21(2), 33–46.
- Jonberg, A., & Porsch, R. (2017). Leistungsangst in der Sekundarstufe I: Welchen Einfluss hat die soziale Herkunft? In H. Wendt & W. Bos (Hrsg.), *Auf dem Weg zum Ganztagsgymnasium. Erste Ergebnisse der wissenschaftlichen Begleitforschung zum Projekt GanzIn* (S. 474–492). Münster: Waxmann.
- Kastens, C. P., Gabriel, K., & Lipowsky, F. (2013). Selbstkonzept im Anfangsunterricht. In F. Lipowsky, G. Faust & C. Kastens (Hrsg.), *Persönlichkeits- und Lernentwicklung an staatlichen und privaten Grundschulen. Ergebnisse der PERLE-Studie zu den ersten beiden Schuljahren* (S. 99–128). Münster: Waxmann.
- Kastens, C. P., & van Wickeren, M. (in Vorbereitung). Gymnasium oder nicht Gymnasium? Eine auf Selbsteinschätzungen lernrelevanter Merkmale basierende Betrachtung von Übergangsempfehlungen. In C. Theurer, M. Hess, A.-K. Denn & F. Lipowsky (Hrsg.), *Determinanten und Effekte der Persönlichkeits- und Lernentwicklung in der Grundschule. Ergebnisse der PERLE Studie*. Wiesbaden: VS.
- Kaufmann, L., Nuerk, H.-C., Graf, M., Krinzinger, H., Delazer, M., & Willmes, K. (2009). *TEDI-MATH: Test zur Erfassung numerisch-rechnerischer Fertigkeiten für 4- bis 8-Jährige*. Bern: Huber.
- Krajewski, K., Küspert, P., & Schneider, W. (2002). *Deutscher Mathematiktest für zweite Klassen (DEMAT 2+)*. Göttingen: Hogrefe.
- Krinzinger, H., Kaufmann, L., & Willmes, K. (2009). Math anxiety and math ability in early primary school years. *Journal of Psychoeducational Assessment*, 27(3), 206–225.
- Lipowsky, F., Kastens, C. P., Lotz, M., & Faust, G. (2011). Aufgabenbezogene Differenzierung und Entwicklung des verbalen Selbstkonzepts im Anfangsunterricht. *Zeitschrift für Pädagogik*, 57(6), 868–884.
- Lüdtke, O., Robitzsch, A., Trautwein, U., & Köller, O. (2007). Umgang mit fehlenden Werten in der psychologischen Forschung. Probleme und Lösungen. *Psychologische Rundschau*, 58(2), 103–117.
- Mammarella, I. C., Donolato, E., Caviola, S., & Giofrè, D. (2018). Anxiety profiles and protective factors: A latent profile analysis in children. *Personality and Individual Differences*, 124, 201–208.
- Marsh, H. W. (1990). Causal ordering of academic self-concept: a multiwave longitudinal panel analysis. *Journal of Educational Psychology*, 82(4), 646–656.
- Marsh, H. W., & Craven, R. G. (2006). Reciprocal effects of self-concept and performance from a multidimensional perspective. Beyond seductive pleasure and unidimensional perspectives. *Perspectives on psychological science*, 1(2), 133–163.
- Marsh, H. W., Lüdtke, O., Trautwein, U., & Morin, A. J. (2009). Classical latent profile analysis of academic self-concept dimensions: Synergy of person- and variable-centered approaches to theoretical models of self-concept. *Structural Equation Modeling: A Multidisciplinary Journal*, 16(2), 191–225.
- Martschinke, S., & Kammermeyer, G. (2006). Selbstkonzept, Lernfreude und Leistungsangst und ihr Zusammenspiel im Anfangsunterricht. In A. Schröder-Lenzen (Hrsg.), *Risikofaktoren kindlicher Entwicklung. Migration, Leistungsangst und Schulübergang* (S. 125–139). Wiesbaden: VS.

- Moser, U., Berweger, S., & Lüchinger-Hutter, L. (2004). *Lest 4–7. Lern- und Entwicklungsstand bei 4- bis 7-jährigen* (Unveröffentlichter Test. Kompetenzzentrum für Bildungsevaluation und Leistungsmessung an der Universität Zürich).
- Muthén, L. K., & Muthén, B. O. (2012). *MPlus user's guide* (7. Aufl.). Los Angeles: Muthén & Muthén.
- Nylund-Gibson, K., & Choi, A. Y. (2018). Ten frequently asked questions about latent class analysis. *Translational Issues in Psychological Science*, 4(4), 440–461.
- Pekrun, R., & Götz, T. (2006). Emotionsregulation: Vom Umgang mit Prüfungsangst. In H. Mandl & H. F. Friedrich (Hrsg.), *Handbuch Lernstrategien* (S. 248–258). Göttingen: Hogrefe.
- Pekrun, R., Lichtenfeld, S., Marsh, H. W., Murayama, K., & Goetz, T. (2017). Achievement emotions and academic performance. Longitudinal models of reciprocal effects. *Child Development*, 88(5), 1653–1670.
- Pixner, S., & Kaufmann, L. (2013). Prüfungsangst, Schulleistung und Lebensqualität bei Schülern. *Lernen und Lernstörungen*, 2(2), 111–124.
- Praetorius, A.-K., Kastens, C. P., Hartig, J., & Lipowsky, F. (2016). Haben Schüler mit optimistischen Selbsteinschätzungen die Nase vorn? Zusammenhänge zwischen optimistischen, realistischen und pessimistischen Selbstkonzepten und der Leistungsentwicklung von Grundschulkindern. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 48(1), 14–26.
- Putwain, D. W., & Daniels, R. A. (2010). Is the relationship between competence beliefs and test anxiety influenced by goal orientation? *Learning and Individual Differences*, 20(1), 8–13.
- Ramirez, G., Gunderson, E. A., Levine, S. C., & Beilock, S. L. (2013). Math anxiety, working memory, and math achievement in early elementary school. *Journal of Cognition and Development*, 14(2), 187–202.
- Richter, D., Kuhl, P., & Pant, H. A. (2012). Soziale Disparitäten des Kompetenzerwerbs. In P. Stanat, H. A. Pant, K. Böhme & D. Richter (Hrsg.), *Kompetenzen von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe in den Fächern Deutsch und Mathematik. Ergebnisse des IQB-Ländervergleichs 2011* (S. 191–208). Münster: Waxmann.
- Schoreit, E. (2016). *Kompetent und trotzdem ängstlich? Profile über Kompetenzwahrnehmungen und Prüfungsängstlichkeit in der Grundschule und die Vorhersagbarkeit der Prüfungsängstlichkeit aufgrund elterlicher Merkmale*. Kassel: University Press.
- Schwarzer, R., & Jerusalem, M. (Hrsg.). (1999). *Skalen zur Erfassung von Lehrer- und Schülermerkmalen: Dokumentation der psychometrischen Verfahren im Rahmen der Wissenschaftlichen Begleitung des Modellversuchs Selbstwirksame Schulen*. Berlin: Freie Universität Berlin.
- Selig, J. P., & Little, T. D. (2012). Autoregressive and cross-lagged panel analysis for longitudinal data. Chapter 16. In B. P. Laursen, T. D. Little & N. A. Card (Hrsg.), *Handbook of developmental research methods* (S. 265–278). New York: Guilford.
- Shavelson, R. J., Hubner, J. J., & Stanton, G. C. (1976). Self-concept: Validation of construct interpretations. *Review of Educational Research*, 46(3), 407–444.
- Sorvo, R., Koponen, T., Viholainen, H., Aro, T., Räikkönen, E., Peura, P., Tolvanen, A., & Aro, M. (2019). Development of math anxiety and its longitudinal relationships with arithmetic achievement among primary school children. *Learning and Individual Differences*, 69, 173–181.
- Sparfeldt, J. R., Schilling, S. R., Rost, D. H., Stelzl, I., & Peipert, D. (2005). Leistungsängstlichkeit: Facetten, Fächer, Fachfacetten? Zur Trennbarkeit nach Angstfacette und Inhaltsbereich. *Zeitschrift für Pädagogische Psychologie*, 19(4), 225–236.
- Trautwein, U., & Möller, J. (2016). Self-concept: Determinants and consequences of academic self-concept in school contexts. In A. A. Lipnevich, F. Preckel & R. D. Roberts (Hrsg.), *Psychosocial skills and school systems in the 21st century. Theory, research, and practice* (S. 187–214). Basel: Springer.
- Valentine, J. C., DuBois, D. L., & Cooper, H. (2004). The relation between self-beliefs and academic achievement. A meta-analytic review. *Educational Psychologist*, 39(2), 111–133.
- Valtin, R., & Wagner, C. (2004). Der Übergang in die Sekundarstufe I: Psychische Kosten der externen Leistungsdifferenzierung. *Psychologie in Erziehung und Unterricht*, 51(1), 52–68.
- Wigfield, A., & Eccles, J. (2000). Expectancy–Value theory of achievement motivation. *Contemporary Educational Psychology*, 25(1), 68–81.
- Wu, S. S., Barth, M., Amin, H., Malcarne, V., & Menon, V. (2012). Math anxiety in second and third graders and its relation to mathematics achievement. *Frontiers in Psychology*, 3(162), 1–11.
- Zeidner, M. (1998). *Test anxiety: the state of the art*. New York: Plenum Press.