

Ehm, Jan-Henning; Hasselhorn, Marcus; Schmiedek, Florian
**Der wechselseitige Einfluss von Selbstkonzept und Leistung bei
Grundschulkindern im Lichte verschiedener längsschnittlicher
Analysemethoden**

Zeitschrift für pädagogische Psychologie 36 (2021) 4, S. 279-288



Quellenangabe/ Reference:

Ehm, Jan-Henning; Hasselhorn, Marcus; Schmiedek, Florian: Der wechselseitige Einfluss von Selbstkonzept und Leistung bei Grundschulkindern im Lichte verschiedener längsschnittlicher Analysemethoden - In: Zeitschrift für pädagogische Psychologie 36 (2021) 4, S. 279-288 - URN: urn:nbn:de:01111-pedocs-252358 - DOI: 10.25656/01:25235

<https://nbn-resolving.org/urn:nbn:de:01111-pedocs-252358>

<https://doi.org/10.25656/01:25235>

Nutzungsbedingungen

Dieses Dokument steht unter folgender Creative Commons-Lizenz: <http://creativecommons.org/licenses/by/4.0/deed.de> - Sie dürfen das Werk bzw. den Inhalt vervielfältigen, verbreiten und öffentlich zugänglich machen sowie Abwandlungen und Bearbeitungen des Werkes bzw. Inhaltes anfertigen, solange Sie den Namen des Autors/Rechteinhabers in der von ihm festgelegten Weise nennen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use

This document is published under following Creative Commons-License: <http://creativecommons.org/licenses/by/4.0/deed.en> - You may copy, distribute and render this document accessible, make adaptations of this work or its contents accessible to the public as long as you attribute the work in the manner specified by the author or licensor.

By using this particular document, you accept the above-stated conditions of use.



Kontakt / Contact:

peDOCS
DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation
Informationszentrum (IZ) Bildung
E-Mail: pedocs@dipf.de
Internet: www.pedocs.de

Mitglied der


Leibniz-Gemeinschaft



Der wechselseitige Einfluss von Selbstkonzept und Leistung bei Grundschulkindern im Lichte verschiedener längsschnittlicher Analysemethoden

Jan-Henning Ehm , Marcus Hasselhorn und Florian Schmiedek

IdEA – Center for Individual Development and Adaptive Education, DIPP | Leibniz-Institut für Bildungsforschung und Bildungsinformation, Frankfurt am Main, Deutschland

Zusammenfassung: Der Zusammenhang zwischen dem akademischen Selbstkonzept und der Leistung wird als reziprok angesehen. Während eine große Anzahl von Studienergebnissen im Sinne einer bidirektionalen Beziehung interpretiert wurden, basieren bisherige Analysen zumeist auf Variationen des klassischen Cross-Lagged-Panel-Modells und beziehen oft nur einen Leistungsindikator in die Modelle mit ein. Ergebnisse basierend auf neueren Modellen, wie beispielsweise dem Random-Intercept Cross-Lagged Panel Model liegen bisher kaum vor. Das Ziel der vorliegenden Studie bestand darin, die längsschnittliche Beziehung zwischen Selbstkonzept und Leistung mit unterschiedlichen Modellen zu analysieren, um herauszufinden, ob die Modelle zu vergleichbaren Ergebnissen hinsichtlich der wechselseitigen Effekte kommen. Basierend auf einer Stichprobe von 1952 Grundschulkindern von Klasse eins bis drei, ergaben sich deutliche Unterschiede. Während Effekte von der Leistung auf das Selbstkonzept in allen Modellen nachgewiesen werden konnten, ist die Evidenz für umgekehrte Effekte eher schwach. Die Ergebnisse werden vor dem Hintergrund theoretischer Annahmen und der Angemessenheit methodischer Verfahren zur Analyse von längsschnittlichen Daten diskutiert.

Schlüsselwörter: Akademisches Selbstkonzept, längsschnittliche Analysen, Random-Intercept Cross-Lagged Panel Model, Grundschulkindern

The Developmental Relation of Academic Self-Concept and Achievement in Elementary School Children in the Light of Alternative Models

Abstract: The association between academic self-concept and achievement is assumed to be reciprocal. Although a large amount of research has been interpreted as demonstrating a bidirectional association, these results are typically based on variants of the classical cross-lagged panel model and often include only one achievement indicator in the models. Results with more recently developed methodological approaches, for example, the random-intercept cross-lagged panel model, are hardly found so far. The present study aimed to test the association between math self-concept and math achievement with different models to evaluate whether positive cross effects can be demonstrated with the alternative models. Drawing on a sample of 1952 elementary students from Grade 1 to Grade 3, results of structural equation modeling yielded noticeable differences. While effects of achievement on self-concept can be demonstrated in all models, the evidence for the reverse effects is rather weak. The results are discussed with regard to the theoretical assumptions and appropriate methodological approaches.

Keywords: Academic self-concept, longitudinal association, random-intercept cross-lagged panel model, elementary school children

Dem Selbstkonzept, insbesondere in der akademischen Domäne, wird seit einigen Jahrzehnten ein reges Forschungsinteresse zuteil. Im Zentrum der Forschung stehen die Struktur und der Zusammenhang einzelner Domänen, die Entwicklung über die akademische Laufbahn hinweg, geschlechtsspezifische Unterschiede sowie der Zusammenhang zu anderen Konstrukten. Dabei bezieht sich eine zentrale Forschungsfrage, deren bisherige Ergebnisse gleichzeitig das rege Forschungsinteresse be-

gründen, auf die möglichen wechselseitigen Einflüsse von Selbstkonzept und Leistung. Drei theoretische Modelle werden dabei angenommen (Calsyn & Kenny, 1977; Valentine, DuBois & Cooper, 2004). Nach dem *Self-Enhancement-Ansatz* stellt das Selbstkonzept eine Determinante der schulischen Leistung dar, während der *Skill-Development-Ansatz* davon ausgeht, dass die schulische Leistung vor allem das Selbstkonzept beeinflusst. Die beiden Ansätze sind nicht nur von theoretischer, sondern

auch von praktischer Relevanz. Folgt man dem *Self-Enhancement-Ansatz*, schlägt sich eine Verbesserung des Selbstkonzepts direkt in einer Verbesserung des individuellen Lernzuwachses nieder, während nach dem *Skill-Development-Ansatz* eine Verbesserung des Selbstkonzepts keinen Einfluss auf die Leistungsentwicklung hat. Empirische Evidenz findet sich für beide Ansätze (*Self-Enhancement*, z.B. durch Marsh 1990, der *Skill-Development-Ansatz*, z.B. durch Helmke & van Aken, 1995; Retelsdorf, Köller & Möller, 2014; Skaalvik & Valås, 1999). Die meisten empirischen Befunde zeigen jedoch wechselseitige Zusammenhänge auf und unterstützen damit den dritten Ansatz, das sogenannte *reciprocal effect model* (REM, z.B. Guay, Ratelle, Roy & Litalien, 2010; Marsh & Craven, 2006; Marsh & Martin, 2011). Entsprechend wird derzeit das REM als der Ansatz angesehen, welcher den Zusammenhang am besten erklärt (siehe Huang, 2011 für eine Übersicht).

Ein Blick auf den Altersbereich der jeweiligen Studien zeigt jedoch auch, dass die Befundlage im Grundschulbereich eher inkonsistent ist (Weidinger, Steinmayr & Spinath, 2017). Weidinger und Kollegen (2017) zeigten zudem auf, dass es im Verlauf der Grundschule zu einem Wechsel in der Dominanz der Kreuzeffekte kommen kann. Neben diesen entwicklungsbedingten Veränderungen kann auch die Operationalisierung der Konstrukte einen Einfluss auf die Dominanz der Kreuzeffekte haben. So zeigte sich beispielsweise in den Daten von Helmke und van Aken (1995), dass durch die gleichzeitige Modellierung von verschiedenen Leistungsindikatoren (Noten und Leistungstests) in einem Modell der Effekt vom Selbstkonzept auf die Leistung gegen Null ging (vgl. Marsh et al., 2018; Pinxten, De Fraine, Van Damme & D'Haenens, 2010). Hinzu kommt, dass bisherige Studien zur Analyse des Zusammenhangs von Selbstkonzept und Schulleistung sich nur auf das klassische Cross-Lagged Panel Model (CLPM) und das sog. Full-Forward Cross-Lagged Panel Modell (FF-CLPM) stützen. Eine Überprüfung mit neueren Modellen, z.B. dem Random-Intercept Cross-Lagged Panel Model (RI-CLPM), bei denen stabile interindividuelle Unterschiede kontrolliert werden, wurde bisher nur in der Studie von Ehm, Hasselhorn und Schmiedek (2019) im Bereich des Lesens durchgeführt. Dabei konnte aufgezeigt werden, dass die Modelle zu unterschiedlichen Ergebnissen und damit auch zu unterschiedlichen Schlussfolgerungen hinsichtlich der kausalen Wirkrichtungen kommen können (siehe auch Usami, Murayama & Hamaker, 2019, für eine Übersicht).

Im folgenden Abschnitt wollen wir kurz auf die zentralen Annahmen des klassischen CLPM und des RI-CLPM eingehen (siehe Abbildung 1). Das CLPM, auch bekannt unter der Bezeichnung *Autoregressives Cross-Lagged Modell* (ARCL), wird zur Analyse des Zusammenhangs von zwei

oder mehreren längsschnittlich erfassten Konstrukten herangezogen. Die Stabilität interindividueller Unterschiede von einem Messzeitpunkt (MZP) zum nächsten wird durch Autoregressionen modelliert, d.h. durch Regressionspfade von einem Konstrukt auf das gleiche Konstrukt zum nächsten MZP. Der Effekt von einem Konstrukt auf ein anderes Konstrukt wird durch längsschnittliche Kreuzeffekte erfasst. Kernelement des CLPM ist demnach die Schätzung von Kreuzeffekten unter der Kontrolle vorangehender individueller Unterschiede des als abhängig betrachteten Konstrukts. Die Stärke beider Effekte ist dabei u.a. von der Stabilität interindividueller Unterschiede in den Konstrukten abhängig. So führen z.B. deutliche Veränderungen in der Rangreihe in einem Konstrukt zu geringeren Autoregressionen. Kreuzeffekte sind hingegen dann zu erwarten, wenn die individuelle Position in einem Konstrukt zu einem früheren Zeitpunkt im Zusammenhang mit der Position in einem anderen Konstrukt zu einem späteren Zeitpunkt steht. Je höher die Stabilität in einem Konstrukt, desto geringer ist die Wahrscheinlichkeit für Kreuzeffekte. Beispielsweise ließen sich in der Studie von Helmke und van Aken (1995), in der die Leistung hohe Autoregressionen aufwies, keine reziproke Kreuzeffekte nachweisen (vgl. Guay, Marsh & Boivin, 2003).

Das RI-CLPM stellt eine Erweiterung des klassischen CLPM dar (Hamaker, Kuiper & Grasman, 2015). Wie die Abbildung 1B aufzeigt, werden darin Random Intercepts, also übergeordnete latente Variablen, deren Faktorladungen auf 1 fixiert werden, in das Modell mitaufgenommen. Die Random Intercepts kontrollieren für über alle MZP überdauernde Stabilitäten, und nicht nur für die Stabilität von jeweils einem MZP zum nächsten. Ausgangspunkt für die Erweiterung des CLPM um die Random Intercepts ist die Annahme, dass zeitstabile interindividuelle Unterschiede durch die Autoregressionen im klassischen Modell nicht adäquat berücksichtigt werden. Nach Hamaker und Kollegen (2015) stellen die Kreuzeffekte im klassischen Modell daher nicht die intraindividuelle Beziehung zwischen zwei Konstrukten dar, was ihrer Auffassung nach zu fehlerhaften Schlussfolgerungen bezüglich des Vorhandenseins, der Dominanz und des Vorzeichens kausaler Effekte führen kann. Neben dem RI-CLPM werden auch im STARTS-Modell (Kenny & Zautra, 2001) stabile interindividuelle Unterschiede durch übergeordnete latente Variablen kontrolliert. In beiden Modellen beeinflussen die Stable Trait bzw. Random Intercepts direkt die einzelnen Indikatoren. Anders als im RI-CLPM werden beim STARTS-Modell jedoch Residualvarianzen modelliert (Hamaker et al., 2015).

Insgesamt scheint die Frage danach, ob stabile interindividuelle Unterschiede kontrolliert werden und ob unterschiedliche Leistungsindikatoren gleichzeitig in einem Modell aufgenommen werden, entscheidend für den

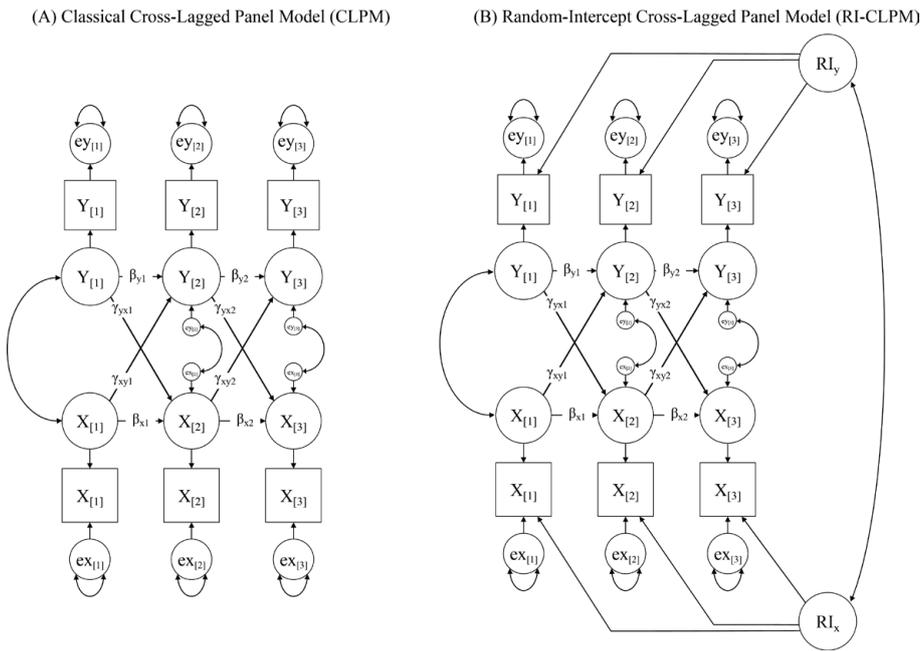


Abbildung 1. Darstellung der unterschiedlichen Modelle zur Analyse des längsschnittlichen Zusammenhangs von Selbstkonzept (Y) und Leistung (X). Manifeste Variablen sind als Quadrate dargestellt, latente Variablen als Kreise, Regressionen durch gerichtete Pfeile, Varianzen und Kovarianzen durch Doppelpfeile dargestellt. Für eine bessere Übersicht wurde auf die Darstellung des Selbstkonzept-Messmodells (3 Indikatoren und itemspezifische Faktoren) verzichtet. RI-CLPM: RI_y/RI_x stellen die Random Intercepts dar.

Nachweis von Kreuzeffekten zu sein. Das Ziel des vorliegenden Beitrags besteht daher darin, zu überprüfen, ob (a) Modelle mit getrennter vs. gemeinsamer Modellierung der Leistungsindikatoren und (b) Modelle, bei denen für zeitlich überdauernde interindividuelle Stabilitäten kontrolliert wird (RI-CLPM) vs. nicht (CLPM), zu vergleichbaren Ergebnissen kommen, um so unser Verständnis über die Dynamik der Beziehung zwischen Selbstkonzept und Leistung zu verbessern.

Basierend auf den bisherigen Studien erwarten wir bedeutsame Kreuzeffekte im CLPM mit getrennter Modellierung der Leistungsindikatoren, bei gemeinsamer Modellierung lediglich Effekte von der Leistung auf das Selbstkonzept. Entsprechend der Analysen von Hamaker und Kollegen (2015) und den Ergebnissen von Ehm und Kollegen (2019) gehen wir davon aus, dass sich im RI-CLPM weder mit getrennter noch mit gemeinsamer Modellierung bedeutsame Kreuzeffekte nachweisen lassen.

Methode

Stichprobe und Durchführung

Die Erhebungen wurden im Rahmen des Projektes „Schulreifes Kind“ durchgeführt (Hasselhorn et al, 2012). Für die vorliegende Studie wurden die Daten von drei MZP herangezogen (Ende der 1. Klasse bis Ende der 3. Klasse). In diesem Zeitraum umfasste die Stichprobe 1952 Schülerinnen und Schüler (51.2% männlich), verteilt auf 90 Klassen (Alter am Ende der 1. Klasse $M = 7.30$ Jahre, $SD = 0.47$). An

den Erhebungen nahmen im ersten Schuljahr $n = 1591$, im zweiten $n = 1705$ und im dritten $n = 1690$ teil. Von $n = 1350$ Kindern lagen die Daten zu allen MZP vor.

Das mathematische Selbstkonzept und die mathematischen Leistungen wurden zu allen Messzeitpunkten zeitgleich (innerhalb einer Woche) erhoben. Im ersten Schuljahr wurde das Selbstkonzept der Schülerinnen und Schüler einzeln, im zweiten und dritten Schuljahr im Klassenkontext erfasst. Um einen Einfluss der Lesefähigkeiten auf das Antwortverhalten der Schülerinnen und Schüler zu kontrollieren, wurden die Items im ersten und zweiten Schuljahr durch geschulte Testleiterinnen und Testleiter vorgelesen. Die Erfassung der Leistung erfolgte in allen drei Klassenstufen im Klassenverband.

Instrumente

Selbstkonzept

Zur Erfassung des mathematischen Selbstkonzepts wurde ein Verfahren verwendet, das konzeptuell auf dem Selbstkonzeptverfahren der LOGIK-Studie basiert (Helmke, 1999; siehe Ehm, Lindberg & Hasselhorn, 2014). Anhand einer siebenstufigen Antwortskala (1 bis 7; niedrige Werte stehen für ein geringes Selbstkonzept) wurde das mathematische Selbstkonzept anhand der drei Items (1) „Im Rechnen bin ich am besten/am schlechtesten“ (2) „Beim Rechnen mache ich sehr viele Fehler/gar keine Fehler“ und (3) „In Mathe bin ich gar nicht gut/sehr gut“ erfasst. Die einzelnen Stufen waren durch Strichmännchen illustriert. Die Schüler sollten das Strichmännchen ankreuzen, das am ehesten ihren eigenen Fähigkeiten entspricht. Die

Tabelle 1. Mittelwert, Standardabweichung, Min, Max und latente Korrelationen der Leistungseinschätzung, der Testleistung und des Selbstkonzepts von Klasse 1 bis 3

	1	2	3	4	5	6	7	8	9
(1) Leistungseinschätzung 1. Klasse	–								
(2) Leistungseinschätzung 2. Klasse	.71	–							
(3) Leistungseinschätzung 3. Klasse	.64	.72	–						
(4) Testleistung 1. Klasse	.65	.64	.62	–					
(5) Testleistung 2. Klasse	.64	.71	.71	.64	–				
(6) Testleistung 3. Klasse	.55	.60	.67	.50	.62	–			
(7) Selbstkonzept 1. Klasse	.30	.23	.22	.27	.21	.18	–		
(8) Selbstkonzept 2. Klasse	.37	.48	.42	.38	.46	.35	.46	–	
(9) Selbstkonzept 3. Klasse	.40	.46	.51	.39	.51	.46	.44	.62	–
<i>M</i>	2.52	2.51	2.40	23.58	18.79	18.28	5.56	5.15	5.08
<i>SD</i>	.99	1.01	1.05	7.45	9.10	5.94	1.21	1.31	1.30
Minimum	0	0	0	2	0	1	1	1	1
Maximum	4	4	4	36	36	31	7	7	7

Anmerkung: Alle Korrelationen $p < .01$.

interne Konsistenz war zufriedenstellend ($t1: \alpha = .77$; $t2: \alpha = .85$; $t3: \alpha = .89$).

Leistung

Als Maße für die Mathematikleistung wurden die Ergebnisse von standardisierten Schulleistungstests und die Bewertungen der Lehrkräfte herangezogen. Die Lehrkräfte schätzen die mathematische Leistung der Schülerinnen und Schüler auf einer fünfstufigen Skala (0 = deutlich unterdurchschnittlich bis 4 = sehr gut) ein. Zur Erfassung der Mathematikleistung wurde die DEMAT Reihe eingesetzt (DEMAT 1+, Krajewski, Küspert & Schneider, 2002; DEMAT 2+, Krajewski, Liehm & Schneider, 2004; DEMAT 3+, Roick, Göllitz & Hasselhorn, 2004).

Statistische Analysen

Die Analysen wurden mit Mplus 8 durchgeführt (Muthén & Muthén, 1998–2017). In den alternativen Modellen wurden alle Konstrukte als latente Variablen modelliert. Um itemspezifische Effekte über die MZP zu kontrollieren, wurden für die Selbstkonzeptindikatoren itemspezifische Faktoren gebildet (Eid, 2000). Dem Messmodell wurde starke faktorielle Invarianz unterstellt (Widaman & Thompson, 2003).¹ Da Leistungseinschätzungen der Lehrkräfte nicht über die Klassen hinweg vergleichbar sind,

wurden diese gruppenzentriert (Goldstein, 2003; Raudenbush & Bryk, 2002). Die Werte der Leistungstests wurden am Gesamtmittelwert standardisiert.

Für die Modellschätzung wurde der Full Information Maximum Likelihood Robust Schätzer (MLR) eingesetzt. Um dem Mehrebenencharakter der Daten Rechnung zu tragen und den korrekten Standardfehler zu erhalten, wurden alle Analysen mit der Option «type is complex» durchgeführt (Klassenzugehörigkeit als Clustervariable). Berichtet werden die gängigen Modellfitstatistiken, die Signifikanz der Regressionsgewichte auf Basis der von Mplus ausgegebenen *z*-Werte. Darüber hinaus wurde die Bedeutsamkeit der Autoregressions- und Kreuzeffekte geprüft, indem die entsprechenden Regressionspfade (von $t1$ zu $t2$ und von $t2$ zu $t3$) simultan auf Null gesetzt wurden. Unterschiede im Modellfit der Modelle mit freier Schätzung vs. Restriktionen wurden mit dem Satorra–Bentler Scaled Chi-Square Difference (TRd) Test geprüft (Satorra & Bentler, 2001).

Ergebnisse

Zunächst werden die deskriptiven Statistiken und Korrelationen für das Selbstkonzept, die Leistungseinschätzung und Leistungstests berichtet. Es folgt die Überprüfung der Messinvarianz des Selbstkonzeptfragebogens über die Zeit. Der Hauptfokus der Ergebnisse liegt auf der

¹ D.h., anders als im RI-CLPM nach Hamaker und Kollegen (2015) wurden Residualvarianzen modelliert um Messfehler zu kontrollieren, wodurch die hier berichteten Modelle mit Random Intercept für das Selbstkonzept dem STARTS-Modell entsprechen. Um Konvergenzprobleme zu vermeiden, wurden die Fehlervarianzen der Leistungsindikatoren hingegen auf Null gesetzt. Da auch das STARTS-Modell ein Random Intercept beinhaltet, wurde die Bezeichnung RI-CLPM für das Modell beibehalten.

Darstellung des längsschnittlichen Zusammenhangs zwischen Selbstkonzept und Leistung anhand der unterschiedlichen Modelle.

Deskriptive Statistiken

Die deskriptiven Statistiken und Korrelationen finden sich in Tabelle 1 (Rohwerte, alle Korrelationen $p < .01$). Der Zusammenhang zwischen den Leistungseinschätzungen von einem MZP zum nächsten erwies sich als relativ hoch (zwi-

schen $r = .71$ und $.72$). Ein vergleichbares Bild zeigte sich bei der Testleistung (zwischen $r = .62$ und $.64$). Der Zusammenhang zwischen dem Selbstkonzept in Klasse 1 und 2 fiel mit $r = .46$ geringer aus als der Zusammenhang von $r = .62$ zwischen Klasse 2 und 3. Die messzeitspezifische Korrelation zwischen der Leistungseinschätzung und der Testleistung lag zwischen $r = .65$ und $r = .71$. Der Zusammenhang zwischen dem Selbstkonzept und der Leistungseinschätzung lag zwischen $r = .30$ (1. Klasse) und $r = .51$ (3. Klasse), der Zusammenhang zwischen Selbstkonzept und Testleistung zwischen $r = .27$ (1. Klasse) und $r = .46$ (3. Klasse).

Tabelle 2. Fitindices der unterschiedlichen Modelle und Vergleich der genesteten Modelle (2 und 3)

	$\chi^2(df)$	CFI	TLI	RMSEA	SRMR	AIC	TRd (<i>df</i>)
CLPM – getrennte Modellierung der Leistungsindikatoren							
Selbstkonzept und Leistungseinschätzung							
(1) Freie Schätzung	165.59 (53)	.982	.977	.033	.037	52157.24	
(2) Leistung auf Selbstkonzept $\gamma_{yx} = 0$	296.06 (55)	.961	.954	.047	.075	52300.31	138.50 (2)
(3) Selbstkonzept auf Leistung $\gamma_{xy} = 0$	181.35 (55)	.980	.976	.034	.043	52172.71	16.39 (2)
Selbstkonzept und Leistungstest							
(1) Freie Schätzung	150.40 (53)	.984	.980	.031	.037	52458.30	
(2) Leistung auf Selbstkonzept $\gamma_{yx} = 0$	305.71 (55)	.959	.950	.049	.075	52616.30	125.51 (2)
(3) Selbstkonzept auf Leistung $\gamma_{xy} = 0$	161.54 (55)	.982	.979	.032	.043	52465.20	11.61 (2)
CLPM – gemeinsame Modellierung der Leistungsindikatoren							
(1) Freie Schätzung	172.51 (89)	.990	.989	.022	.031	61819.47	
(2) Leistung auf Selbstkonzept $\gamma_{yx} = 0$	356.83 (91)	.969	.965	.039	.087	62016.23	150.10 (2)
(3) Selbstkonzept auf Leistung $\gamma_{xy} = 0$	186.86 (91)	.989	.987	.023	.031	61831.78	12.72 (2)
RI-CLPM							
Selbstkonzept und Leistungseinschätzung							
(1) Freie Schätzung	92.18 (50)	.993	.991	.021	.026	52077.16	
(2) Leistung auf Selbstkonzept $\gamma_{yx} = 0$	109.93 (52)	.991	.988	.024	.033	52091.52	22.45 (2)
(3) Selbstkonzept auf Leistung $\gamma_{xy} = 0$	106.47 (52)	.991	.989	.023	.031	52089.33	13.43 (2)
Selbstkonzept und Leistungstest							
(1) Freie Schätzung	85.96 (50)	.994	.992	.019	.022	52397.09	
(2) Leistung auf Selbstkonzept $\gamma_{yx} = 0$	132.90 (52)	.987	.983	.029	.033	52439.99	51.48 (2)
(3) Selbstkonzept auf Leistung $\gamma_{xy} = 0$	92.53 (52)	.993	.992	.020	.024	52399.73	6.55 (2)
RI-CLPM – gemeinsame Modellierung der Leistungsindikatoren							
(1) Freie Schätzung	157.69 (87)	.992	.990	.020	.024	61806.50	
(2) Leistung auf Selbstkonzept $\gamma_{yx} = 0$	198.49 (89)	.987	.985	.025	.045	62016.23	35.19 (2)
(3) Selbstkonzept auf Leistung $\gamma_{xy} = 0$	169.91 (89)	.991	.989	.022	.024	61815.10	13.37 (2)

Anmerkungen: Die Modelle 2 und 3 wurden als Restriktionen des Modells mit freier Schätzung (1) spezifiziert. Eine sig. Veränderung des Fits ist durch den Satorra-Bentler Scaled Chi-Square Difference (TRd) angegeben, alle Veränderungen $p < .05$. Um Schätzprobleme zu vermeiden, wurden beim RI-CLPM mit gemeinsamer Modellierung der Leistungsindikatoren die Random Intercepts über die latenten Leistungsfaktoren operationalisiert.

https://econtent.hogrefe.com/doi/pdf/10.1024/1010-0652/a000303 - Tuesday, August 23, 2022 1:36:04 AM - Deutsches Institut für Internationale Pädagogische Forschung DIPP IP Address: 195.37.189.82

Messinvarianz

Eine Voraussetzung für die Analyse des Zusammenhangs zwischen Selbstkonzept und Leistung besteht in der Bestätigung von Messinvarianz des Selbstkonzeptfragebogens über die Zeit. Um herauszufinden, ob die Items über die drei MZP hinweg das Selbstkonzept in vergleichbarer Weise messen, wurde eine CFA mit starker faktorieller Invarianz (Gleichheit der Ladungen und Intercepts) durchgeführt. Dieses restriktivere Modell wies ein vergleichbaren Modellfit wie das Modell ohne Restriktionen auf (freies Modell $\chi^2(22) = 21.903, p = 0.466, CFI = 1.00, TLI = 1.00, RMSEA = .000, SRMR = .011$, Invarianz Modell $\chi^2(30) = 47.949, p < .05, CFI = .996, TLI = .995, RMSEA = .036, SRMR = .027, RDR = .033$, siehe Browne & Du Toit, 1992). Folglich kann starke Messinvarianz angenommen werden (Meade, Johnson & Braddy, 2008; Widaman & Thompson, 2003).

Längsschnittlicher Zusammenhang zwischen Selbstkonzept und Leistungseinschätzungen / Leistung

CLPM – getrennte Modellierung der Leistungsindikatoren. Der Fit für beide Modelle, d.h. für Selbstkonzept und Leistungseinschätzung $\chi^2(53) = 165.59, p < .001, CFI = .982, TLI = .977, RMSEA = .033, SRMR = .037$, und Selbstkonzept und Testleistung $\chi^2(53) = 150.40, p < .001, CFI = .984, TLI = .980, RMSEA = .031, SRMR = .037$ erwiesen sich als gut. Signifikante Kreuzeffekte von der Leistung auf das Selbstkonzept konnten im Modell mit Leistungseinschätzungen zu beiden Messzeitpunkten ($\beta = .26/.21$), vom Selbstkonzept auf die Leistung jedoch nur von Messzeitpunkt zwei zu drei nachgewiesen werden ($\beta = .10$). Ein vergleichbares Bild ergibt sich auch im Modell mit Leistungstests (siehe Tabelle 3). Eine Fixierung der Kreuzeffekte auf Null führte zu einer signifikanten Verschlechterung des Modellfits (siehe Tabelle 2).

CLPM – gemeinsame Modellierung der Leistungsindikatoren. Auch das Modell, bei dem Leistungseinschätzungen und Leistungstest als Indikatoren einer latenten Variable modelliert wurden, zeigte eine sehr gute Modellanpassung $\chi^2(80) = 172.51, p < .001, CFI = .990, TLI = .989, RMSEA = .022, SRMR = .031$. Signifikant positive Kreuzeffekte konnten hier jedoch nur von der Leistung auf das Selbstkonzept ($\beta = .36/.29$), nicht aber in umgekehrter Richtung nachgewiesen werden.

Random-Intercept CLPM (RI-CLPM). Wie aus Tabelle 2 hervorgeht, zeigt auch das RI-CLPM eine gute Modellanpassung (Modell Leistungseinschätzung $\chi^2(50) = 92.18, p < .001, CFI = .993, TLI = .991, RMSEA = .021, SRMR = .026$; Modell Test $\chi^2(50) = 85.96, p < .001, CFI = .994, TLI = .992, RMSEA = .019, SRMR = .022$). Die Random Intercepts der Leistung und des Selbstkonzepts korrelierten moderat miteinander (Modell Leistungseinschätzung:

$r = .63$; Modell Test: $r = .49$). Signifikante Kreuzeffekte von der Leistung auf das Selbstkonzept konnten zu allen Messzeitpunkten (Modell Leistungseinschätzungen: $\beta = .21/.18$; Modell Leistungstest: $\beta = .23/.31$) und vom Selbstkonzept auf die Leistung von Messzeitpunkt zwei zu drei nachgewiesen werden (siehe Tabelle 3).

RI-CLPM – gemeinsame Modellierung der Leistungsindikatoren. Das Modell zeigte eine sehr gute Modellanpassung $\chi^2(87) = 157.690, p < .001, CFI = .992, TLI = .990, RMSEA = .020, SRMR = .024$. Signifikant positive Kreuzeffekte konnten auch hier nur von der Leistung auf das Selbstkonzept ($\beta = .22/.23$), nicht aber in umgekehrter Richtung nachgewiesen werden. Die Random Intercepts der Leistung und des Selbstkonzepts korrelierten zu $r = .58$ miteinander.

Diskussion

Das Ziel der vorliegenden Studie bestand darin, die längsschnittliche Beziehung zwischen Selbstkonzept und Leistung mit unterschiedlichen Modellen zu analysieren. Die vorliegenden Ergebnisse zeigen ein unterschiedliches Befundmuster auf. Entsprechend der Hypothesen konnten bedeutsame Kreuzeffekte im CLPM mit getrennter Modellierung der Leistungsindikatoren nachgewiesen werden, vom Selbstkonzept auf die Leistung jedoch nur von MZP 2 auf MZP 3. Bei gemeinsamer Modellierung ließen sich lediglich Effekte von der Leistung auf das Selbstkonzept nachweisen. Die Modellierung des Zusammenhangs anhand des RI-CLPM bestätigen die hier formulierten Annahmen nur teilweise. Analog zu den Befunden des CLPM lassen sich im RI-CLPM mit getrennter, jedoch nicht mit gemeinsamer Modellierung der Leistungsindikatoren bedeutsame Kreuzeffekte nachweisen. Im folgendem Abschnitt werden die vorliegenden Befunde in die bestehende Literatur eingebettet, und es wird auf die Frage eingegangen, welches Modell die Daten am besten erklärt.

Die anhand des CLPM mit getrennter Modellierung der Leistungsindikatoren berichteten Ergebnisse bestätigen frühere Studien im Grundschulbereich, die reziproke Effekte berichten und auf eine klare Dominanz der Effekte von der Leistung auf das Selbstkonzept hinweisen (Helmke & van Aken, 1995; Praetorius, Kastens, Hartig & Lipowsky, 2016; Skaalvik & Valås, 1999). Die Dominanz des Skill-Development-Ansatzes wird nicht nur dadurch erkenntlich, dass Effekte von der Leistung auf das Selbstkonzept zu zwei Messzeitpunkten, ein umgekehrter Effekt jedoch nur zu einem Messzeitpunkt zu beobachten ist, sondern auch durch die Höhe der Regressionsgewichte (siehe Tabelle 3). Entsprechend führt die Fixierung der Regressionen von der Leistung auf das Selbstkonzept auf Null zu einer deutlicheren Verschlechterung des Fits als

Tabelle 3. Standardisierte Autoregression und Kreuzeffekte für das CLPM und RI-CLPM

CLPM – getrennte Modellierung auf der Leistungsindikatoren		Regressionen
Selbstkonzept und Leistungseinschätzung	Autoregression – Leistung	.70* / .67*
	Autoregression – Selbstkonzept	.44* / .50*
	Kreuzeffekt – Leistung auf Selbstkonzept	.26* / .21*
	Kreuzeffekt – Selbstkonzept auf Leistung	.03 / .10*
Selbstkonzept und Leistungstest	Autoregression – Leistung	.63* / .60*
	Autoregression – Selbstkonzept	.47* / .50*
	Kreuzeffekt – Leistung auf Selbstkonzept	.25* / .24*
	Kreuzeffekt – Selbstkonzept auf Leistung	.04 / .07*
CLPM – gemeinsame Modellierung der Leistungsindikatoren		
Selbstkonzept und Leistung	Autoregression – Leistung	.98* / .98*
	Autoregression – Selbstkonzept	.39* / .45*
	Kreuzeffekt – Leistung auf Selbstkonzept	.36* / .29*
	Kreuzeffekt – Selbstkonzept auf Leistung	-.08* / -.06*
RI-CLPM		
Selbstkonzept und Leistungseinschätzung	Autoregression – Leistung	.34* / .34*
	Autoregression – Selbstkonzept	.21* / .39*
	Kreuzeffekt – Leistung auf Selbstkonzept	.21* / .18*
	Kreuzeffekt – Selbstkonzept auf Leistung	-.04 / .18*
Selbstkonzept und Leistungstest	Autoregression – Leistung	.37* / .31*
	Autoregression – Selbstkonzept	.19* / .34*
	Kreuzeffekt – Leistung auf Selbstkonzept	.23* / .31*
	Kreuzeffekt – Selbstkonzept auf Leistung	-.02 / .12*
RI-CLPM – gemeinsame Modellierung der Leistungsindikatoren		
Selbstkonzept und Leistung	Autoregression – Leistung	.15* / .07*
	Autoregression – Selbstkonzept	.18* / .33*
	Kreuzeffekt – Leistung auf Selbstkonzept	.22* / .23*
	Kreuzeffekt – Selbstkonzept auf Leistung	-.08* / .05

Anmerkungen: vor dem Spiegelstrich: Regressionen auf AV der 2. Klasse, nach dem Spiegelstrich: Regression auf AV der 3. Klasse. * $p < .05$.

die Fixierung der umgekehrten Regressionen (Vergleiche Modell 2 und 3 in Tabelle 2).

Auf den ersten Blick können die hier aufgezeigten Ergebnisse auch als eine Bestätigung der Annahme von Weidinger und Kollegen (2017) aufgefasst werden, die zeigen, dass zunächst Effekte von der Leistung auf das Selbstkonzept vorhanden sind und es erst im Verlauf der Grundschule zu wechselseitigen Kreuzeffekten kommt. Allerdings setzt die hier durchgeführte Studie zeitlich früher an und umfasst deutlich weniger Messzeitpunkte, so dass über den weiteren Verlauf keine Aussagen getroffen werden können. Auch ist der Abstand zwischen den Messzeitpunkten in der Studie von Weidinger und Kollegen (2017) deutlich kürzer, was erheblichen Einfluss auf die Höhe der Autoregressionen haben kann. Dies zeigen in der vorliegenden Studie auch die Befunde zum CLPM mit gemeinsamer Modellierung der Leistungsindikatoren. In Übereinstimmung mit den Befunden von Helmke und van Aken (1995) führt die gemeinsame Modellierung zu deutlich höheren Autoregressionen, so dass keine positiven Effekte vom Selbstkonzept auf die Leistung nachgewiesen werden können. In der vorliegenden Studie wird die Leistung zu einem Messzeitpunkt fast komplett durch die Leistung zum vorherigen Messzeitpunkt erklärt ($\beta = .98 / .98$) – was für eine erhebliche Stabilität der Leistung spricht – und hat darüber hinaus einen erheblichen Einfluss auf die Entwicklung des Selbstkonzepts ($\beta = .36 / .29$). Die gemeinsame Modellierung der Leistungsindikatoren kann somit als eine validere Schätzung der tatsächlichen Leistung aufgefasst werden, was mit Veränderungen im Zusammenspiel zwischen Leistung und Selbstkonzept verbunden ist.

Die Frage der Stabilität interindividueller Unterschiede und ihrer Kontrolle steht im Fokus des RI-CLPM. Gegenüber dem CLPM können vor allem substantielle Unterschiede hinsichtlich des Fits beobachtet werden. Hinsichtlich der Zusammenhänge zeigen sich Unterschiede in den Autoregressionen, nicht aber bezüglich der Kreuzeffekte. Genau wie beim CLPM Modell lassen sich Kreuzeffekte bei getrennter, nicht jedoch bei gemeinsamer Modellierung nachweisen. Deutlich wird auch, dass die Korrelationen zwischen den Random Intercepts über einen Zeitraum von drei Jahren auf einen höheren Zusammenhang zwischen Selbstkonzept und Leistung hinweisen als bei einer querschnittlichen Betrachtung (Random Intercept getrennt $r = .62 / .49$, gemeinsam $r = .58$ vs. querschnittliche Betrachtung $r = .27 - .51$). Die Ergebnisse des RI-CLPM stehen im Einklang mit den hier berichteten Ergebnissen des CLPM, jedoch in Kontrast zu den Ergebnissen von Ehm und Kollegen (2019), bei denen für Testleistungen und -einschätzungen im Bereich Lesen keine Kreuzeffekte im RI-CLPM nachgewiesen werden konnten. Die unterschiedlichen Befunde könnten zum einen auf die Betrachtung

tung unterschiedlicher Domänen, zum anderen auf die hier geringere Anzahl von Messzeitpunkten zurück zu führen sein. Je mehr Messzeitpunkte zur Verfügung stehen, desto besser können stabile interindividuelle Unterschiede kontrolliert werden (Bainter & Howard, 2016).

Insgesamt sprechen die hier aufgezeigten Befunde deutlich für eine Dominanz des Skill-Development-Ansatzes. Auch in der Literatur besteht weitgehende Einigkeit darin, dass das Selbstkonzept durch die Erfahrungen im Schul- und Leistungskontext, insbesondere durch die Rückmeldung signifikanter Anderer (z.B. Lehrkräfte), beeinflusst wird (Harter, 1999). Gerade der zunehmende Zusammenhang zwischen Selbstkonzept und Leistung in der Grundschule wird als ein Indiz dafür betrachtet, dass sich das Selbstkonzept langsam den tatsächlichen Leistungen anpasst. Der Prozess der Anpassung wird zudem durch die Vergabe von Noten beschleunigt (Harter, 1999). Bei den vorliegenden Leistungseinschätzungen durch die Lehrkräfte handelt es sich jedoch nicht um Noten im klassischen Sinne. Hier werden typischerweise höhere Zusammenhänge als zwischen Selbstkonzepten und Leistungstests erzielt (Marsh et al., 2018).

Während die interindividuellen Differenzen hinsichtlich der Leistung bereits in der Grundschule relativ stabil sind, ist das Selbstkonzept deutlicheren Schwankungen unterworfen. Dies führt zu der Frage, ob es aus theoretischer Sicht überhaupt sinnvoll ist, dem Selbstkonzept bereits in der Grundschule einen Trait-Charakter zu unterstellen. Unterstellt man jedoch dem Selbstkonzept eine gewisse Stabilität, so gilt es, diese richtig zu berücksichtigen. Wie Hamaker und Kollegen (2015) aufzeigen, werden im CLPM intra- und interindividuelle Prozesse nicht getrennt, was zu fehlerhaften Schlussfolgerungen bezüglich des Vorhandenseins, der Dominanz und des Vorzeichens von Kreuzeffekten führen kann. Wie Usami, Hayes und McArdle (2015) anmerken, ist jedoch die Wahrscheinlichkeit, mögliche Kreuzeffekte zwischen zwei Konstrukten nachzuweisen, in Modellen, bei denen die Stabilität durch übergeordnete Faktoren zusätzlich berücksichtigt wird, deutlich geringer. Das RI-CLPM kann demnach als eine strengere Prüfung möglicher Kreuzeffekte aufgefasst werden. Neben einer Bevorzugung des Modells aus theoretischen Gründen sprechen auch die hier durchgeführten Analysen für die Angemessenheit des Modells und der dahinter liegenden Annahmen. Der Vergleich der Fitindizes zeigt deutlich auf, dass das RI-CLPM die vorliegenden Daten besser erklärt als das CLPM.

Limitationen und Schlussfolgerung

Eine Limitation, die möglicherweise direkt zu weiterer Forschung führt, ergibt sich aus der Tatsache, dass die vor-

liegende Studie auf die Grundschule beschränkt ist. Daher sollten die vorliegenden Befunde mit den hier verwendeten Methoden auch für die Sekundarstufe überprüft werden. Möglicherweise ergeben sich dort, auch unter Kontrolle stabiler interindividueller Unterschiede, nicht nur statistisch, sondern auch praktisch bedeutsame Effekte vom Selbstkonzept auf die Leistung.

Insgesamt können die hier durchgeführten Analysen bisherige Befunde weitestgehend bestätigen, die auf eine klare Dominanz der Effekte von der Leistung auf das Selbstkonzept im Grundschulalter hinweisen. Zwar lassen sich in einigen Modellen auch Effekte vom Selbstkonzept auf die Leistung nachweisen, deren praktische Relevanz kann jedoch in Frage gestellt werden. Dies führt zu der Frage, ob auf Maßnahmen, die auf eine Verbesserung des Selbstkonzepts abzielen, im pädagogischen Kontext verzichtet werden sollte (vgl. O'Mara, Marsh, Craven & Debus, 2006). Diese Schlussfolgerung wäre nicht nur voreilig, sondern auch falsch. Denn unabhängig davon, ob das Selbstkonzept die Leistung beeinflusst, stellt die Erlangung eines positiven Selbstkonzepts selbst ein wichtiges pädagogisches Ziel dar. Zudem geht ein positives Selbstkonzept mit anderen leistungsförderlichen Konstrukten, wie beispielsweise einer hohen Selbstwirksamkeit oder einer hohen Lern- und Leistungsmotivation einher (Bong & Skaalvik 2003; Marsh, Trautwein, Lüdtke, Köller & Baumert, 2006). Auch kann nicht ausgeschlossen werden, dass das Befundmuster in der Sekundarschule ein anderes ist und dass der Einfluss des Selbstkonzepts auf die Leistung in experimentellen Studien oder Längsschnittstudien mit kürzeren Zeitabständen nachgewiesen werden kann (vgl. Dormann & Griffin, 2015). Allein basierend auf den vorliegenden Ergebnissen können daher keine abschließenden Schlussfolgerungen gezogen werden.

Literatur

- Bainter, S. A., & Howard, A. L. (2016). Comparing within-person effects from multivariate longitudinal models. *Developmental Psychology*, 52, 1955 – 1968. <http://doi.org/10.1037/dev0000215>
- Bong, M. & Skaalvik, E. M. (2003). Academic self-concept and self-efficacy: How different are they really? *Educational Psychology*, 15, 1 – 40.
- Browne, M. W. & Du Toit, S. H. (1992). Automated fitting of nonstandard models. *Multivariate Behavioral Research*, 27, 269 – 300.
- Calsyn, R. J., & Kenny, D. A. (1977). Self-concept of ability and perceived evaluation of others: Cause or effect of academic achievement? *Journal of Educational Psychology*, 69, 136 – 145. <https://dx.doi.org/10.1037/0022-0663.69.2.136>
- Dormann, C. & Griffin, M. A. (2015). Optimal time lags in panel studies. *Psychological Methods*, 20, 489 – 505.
- Ehm, J.-H., Lindberg, S., & Hasselhorn, M. (2014). Reading, writing, and math self-concept in elementary school children: influence of dimensional comparison processes. *European Journal of*

- Psychology of Education*, 29, 277 – 294. <https://doi.org/10.1007/s10212-013-0198-x>
- 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, 2336 – 2351. <https://doi.org/10.1037/dev0000796>
- Eid, M. (2000). A multitrait-multimethod model with minimal assumptions. *Psychometrika*, 65, 241 – 261. <https://dx.doi.org/10.1007/BF02294377>
- Goldstein, H. (2003). *Multilevel statistical models* (3rd ed.). London: Hodder Arnold.
- Guay, F., Marsh, H. W. & Boivin, M. (2003). Academic self-concept and academic achievement: Developmental perspectives on their causal ordering. *Journal of Educational Psychology*, 95, 124 – 136.
- Guay, F., Ratelle, C. F., Roy, A. & Litalien, D. (2010). Academic self-concept, autonomous academic motivation, and academic achievement: Mediating and additive effects. *Learning and Individual Differences*, 20, 644 – 653. <https://doi.org/10.1016/j.lindif.2010.08.001>
- Hamaker, E. L., Kuiper, R. M. & Grasman, R. P. P. (2015). A Critique of the Cross-Lagged Panel Model. *Psychological Methods*, 20, 102 – 116.
- Harter, S. (1999). *The construction of the self: A developmental perspective*. New York: Guilford.
- Hasselhorn, M., Schöler, H., Schneider, W., Ehm, J.-H., Johnson, M., Keppler, I., et al. (2012). Gezielte Zusatzförderung im Modellprojekt "Schulreifes Kind". *Frühe Bildung*, 1, 3 – 10. <https://dx.doi.org/10.1026/2191-9186/a000019>
- Helmke, A. (1999). From optimism to realism? Development of children's academic self-concept from kindergarten to grade six. In W. Schneider & F. E. Weinert (Eds.), *Individual development from 3 to 12. Findings from the Munich Longitudinal Study* (pp. 198 – 221). Cambridge: Cambridge University Press.
- 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, 624 – 637.
- Huang, C. (2011). Self-concept and academic achievement: A meta-analysis of longitudinal relations. *Journal of School Psychology*, 49, 505 – 528.
- Kenny, D. A., & Zautra, A. (2001). Trait–state models for longitudinal data. In L. M. Collins & A. G. Sayer (Eds.), *Decade of behavior. New methods for the analysis of change* (pp. 243 – 263). Washington, DC: American Psychological Association.
- Krajewski, K., Liehm, S., & Schneider, W. (2004). *DEMAT 2+: Deutscher Mathematiktest für zweite Klassen*. Beltz Test.
- Krajewski, K., Küspert, P. & Schneider, W. (2002). *Deutscher Mathematiktest für erste Klassen: DEMAT 1+*. Beltz Test.
- Marsh, H. W. (1990). The structure of academic self-concept: The Marsh & Shavelson model. *Journal of Educational Psychology*, 82, 623 – 636.
- 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, 133 – 163. <https://dx.doi.org/10.1111/j.1745-6916.2006.00010.x>
- Marsh, H. W., & Martin, A. J. (2011). Academic self-concept and academic achievement: Relations and causal ordering. *British Journal of Educational Psychology*, 81, 59 – 77. <https://dx.doi.org/10.1348/000709910X503501>
- Marsh, H. W., Pekrun, R., Murayama, K., Arens, A. K., Parker, P. D., Guo, J., & Dicke, T. (2018). An integrated model of academic self-concept development: Academic self-concept, grades, test scores, and tracking over 6 years. *Developmental Psychology*, 54, 263 – 280. <https://doi.org/10.1037/dev0000393>
- Marsh, H. W., Trautwein, U., Lüdtke, O., Köller, O. & Baumert, J. (2006). Integration of multidimensional self-concept and core personality constructs: construct validation and relations to well-being and achievement. *Journal of Personality*, 74, 403 – 456.
- Meade, A. W., Johnson, E. C. & Braddy, P. W. (2008). Power and sensitivity of alternative fit indices in tests of measurement invariance. *Journal of Applied Psychology*, 93, 568 – 592.
- Muthén, L. K., & Muthén, B. O. (1998 – 2017). *Mplus user's guide* (7th ed.). Los Angeles: Muthén & Muthén.
- O'Mara, A. J., Marsh, H. W., Craven, R. G. & Debus, R. L. (2006). Do self-concept interventions make a difference? A synergistic blend of construct validation and meta-analysis. *Educational Psychologist*, 41, 181 – 206.
- Pinxten, M., De Fraine, B., Van Damme, J., & D'Haenens, E. (2010). Causal ordering of academic self-concept and achievement: Effects of type of achievement measure. *British Journal of Educational Psychology*, 80, 689 – 709.
- Praetorius, A. K., Kastens, C., Hartig, J., & Lipowsky, F. (2016). Haben Schüler mit optimistischen Selbsteinschätzungen die Nase vorn? *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 48, 14 – 26. <https://dx.doi.org/10.1026/0049-8637/a000140>
- Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousand Oaks, CA: Sage.
- Retelsdorf, J., Köller, O., & Möller, J. (2014). Reading achievement and reading self-concept – Testing the reciprocal effects model. *Learning and Instruction*, 29, 21 – 30. <https://dx.doi.org/10.1016/j.learninstruc.2013.07.004>
- Roick, T., Göllitz, D. & Hasselhorn, M. (2004). *DEMAT 3+: Deutscher Mathematiktest für dritte Klassen*. Beltz Test.
- Sattora, A. & Bentler, P.M. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika*, 66, 507 – 514.
- Skaalvik, E. M. & Valås, H. (1999). Relations among achievement, self-concept, and motivation in mathematics and language arts: A longitudinal study. *The Journal of Experimental Education*, 67, 135 – 139.
- Usami, S., Hayes, T. & McArdle, J. J. (2015). Inferring Longitudinal Relationships Between Variables: Model Selection Between the Latent Change Score and Autoregressive Cross-Lagged Factor Models. *Structural Equation Modeling*, 50, 1 – 12.
- Usami, S., Murayama, K. & Hamaker, E. L. (2019). A Unified Framework of Longitudinal Models to Examine Reciprocal Relations. *Psychological Methods*. <https://dx.doi.org/10.1037/met0000210>
- Valentine, J., DuBois, D. & Cooper, H. (2004). The Relation Between Self-Beliefs and Academic Achievement: A Meta-Analytic Review. *Educational Psychologist*, 39, 111 – 133.
- Weidinger, A. F., Steinmayr, R. & Spinath, B. (2017). Changes in the relation between competence beliefs and achievement in math across elementary school years. *Child Development*, 89, 138 – 156. <https://doi.org/10.1111/cdev.12806>
- Widaman, K. F. & Thompson, J. S. (2003). On specifying the null model for incremental fit indices in structural equation modeling. *Psychological Methods*, 8, 16 – 37.

Historie

Manuskript eingereicht: 11.09.2020

Manuskript nach Revision angenommen: 07.01.2021

Onlineveröffentlichung: 11.02.2021

Open Data

The data that support the findings of this study are openly available in the OSF at: <https://osf.io/g5t86/>

Förderung

Gefördert durch das Ministerium für Kultus, Jugend und Sport Baden-Württemberg.
Open Access-Veröffentlichung ermöglicht durch das DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation.

ORCID

Jan-Henning Ehm

 <https://orcid.org/0000-0001-8901-0309>

Dr. Jan-Henning Ehm

DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation
IDeA – Center for Individual Development and Adaptive Education
Frankfurt, Germany
Rostocker Straße 6
60323 Frankfurt am Main
Deutschland
ehm@dipf.de