

Antoine Fischbach, Martin Brunner, Stefan Krauss & Jürgen Baumert

## **Die Bezugsnormorientierung von Mathematiklehrkräften am Ende der Sekundarstufe I: Konvergenz verschiedener Messverfahren und Wirkung auf motivational-affektive Aspekte des Mathematiklernens und Leistung**

### **Zusammenfassung**

*Lehrkräfte unterscheiden sich in den Normen, die sie zur Bewertung von Schülerleistungen heranziehen. Bisherige Studien zeigten, dass eine individuelle Bezugsnormorientierung der Lehrkraft positive Effekte auf die Entwicklung von Schülermotivation und -leistung hat. Wir untersuchten die Generalisierbarkeit früherer Befunde anhand von längsschnittlichen Daten (9. und 10. Klasse) der COACTIV-Studie ( $N_{\text{Schüler}} = 2465$ ;  $N_{\text{Lehrkräfte}} = 106$ ). Hierzu analysierten wir die Effekte von multimethodal und mehrperspektivisch erfasster Bezugsnormorientierung auf die Entwicklung zentraler Aspekte der Lernmotivation und der Leistung im Fach Mathematik getrennt für die Schulformen Gymnasium und Realschule. Korrelationsanalysen wiesen darauf hin, dass die verschiedenen Maße zur Erfassung von Bezugsnormorientierung höchstens moderat über Methoden oder Perspektiven hinweg konvergierten. Analysen hierarchisch linearer Modelle zeigten, dass insgesamt viele affektiv-motivationale Schülermerkmale nur geringe Variabilität auf Klassenebene aufwie-*

---

Dr. Antoine Fischbach (corresponding author), Luxembourg Centre for Educational Testing (LUCET), Fakultät für Sprachwissenschaften und Literatur, Geisteswissenschaften, Kunst und Erziehungswissenschaften, Universität Luxemburg, Porte des Sciences 11, 4366 Esch/Alzette, Luxemburg  
E-Mail: antoine.fischbach@uni.lu

Prof. Dr. Martin Brunner, Fachbereich Erziehungswissenschaft und Psychologie, Freie Universität Berlin & Institut für Schulqualität der Länder Berlin und Brandenburg e.V. (ISQ), Otto-von-Simson-Straße 15, 14195 Berlin, Deutschland  
E-Mail: martin.brunner@isq-bb.de

Prof. Dr. Stefan Krauss, Didaktik der Mathematik, Fakultät für Mathematik, Universität Regensburg, Universitätsstraße 31, 93053 Regensburg, Deutschland  
E-Mail: stefan1.krauss@mathematik.uni-regensburg.de

Prof. Dr. Drs. h.c. Jürgen Baumert, Direktor em., Max-Planck-Institut für Bildungsforschung Berlin, Lentzeallee 94, 14195 Berlin, Deutschland  
E-Mail: sekbaumert@mpib-berlin.mpg.de

sen. Effekte von Bezugsnormorientierungen auf die Entwicklung von Motivation und Leistung waren klein, meist statistisch nicht signifikant, sowie inkonsistent für verschiedene Erhebungsmethoden und Schulformen. Zukünftige Forschung sollte Probleme bei der Modellierung und Messung der Bezugsnormorientierung sowie die Analyse von Effekten der Lernumwelt bei nur geringer Variabilität auf der Klassenebene stärker in den Blick nehmen.

### **Schlagworte**

Bezugsnormorientierung; PISA; Mehrebenenmodelle; Schulische Leistungsbeurteilung

## **Mathematics teachers' frame of reference at the end of lower secondary school: Convergence of measures across methods and effects on motivational-affective aspects of students' mathematics learning and achievement**

### **Abstract**

*Teachers differ in the frame of reference they apply to evaluate students' performances. Previous research found positive effects of an ipsative frame of reference on students' achievement and achievement motivation. In the present study we investigated the generalizability of previous results with longitudinal data (9<sup>th</sup> and 10<sup>th</sup> grade) from the COACTIV study ( $N_{Students} = 2,465$ ;  $N_{Teachers} = 106$ ). The teacher frame of reference was assessed by means of a multi-method approach to study the differential effects on student achievement in mathematics and on various key aspects of their achievement motivation. Statistical analyses were conducted separately for the highest and intermediate academic track. Correlation analyses showed that measures of teachers' frame of reference converged only modestly across different methods of assessment. Results of hierarchical linear models showed that many motivational-affective student characteristics had only modest variability on the class level. The frame of reference effects were small and most of them were not statistically significant. Furthermore, results were inconsistent across frame of reference assessment methods and academic tracks. Future research should address modeling and measurement issues of teachers' frame of reference, and tackle how to analyze effects of learning environments when the variables of interest show only moderate variability on the class level.*

### **Keywords**

*Teacher frame of reference; PISA; Multilevel modeling; Educational assessment*

## 1. Einleitung

Lehrkräfte unterscheiden sich darin, welchen Normen sie bei ihrer Beurteilung von Schülerleistungen folgen (Heckhausen, 1974). Ziel unseres Artikels ist es zu untersuchen, wie sich diese *Bezugsnormorientierung* (BNO) auf die Entwicklung der Lernmotivation und der Leistung im Fach Mathematik in den Schulformen Gymnasium und Realschule auswirkt.

## 2. Theoretischer Rahmen

### 2.1 Definition und Erfassung der BNO

Unter Bezugsnormorientierung versteht man die Tendenz, einen bestimmten Vergleichsstandard zu bevorzugen, wenn man dazu den Bewertungsspielraum hat. Lehrkräfte unterscheiden sich in ihrer BNO (Rheinberg, 1980). Bei einer *sozialen Bezugsnormorientierung* (SBNO) setzt die Lehrkraft die Schülerleistung in Bezug zu den Leistungen anderer Schüler; der Fokus liegt auf der Position eines Schülers innerhalb der Leistungsverteilung einer sozialen Bezugsgruppe zu einem bestimmten Zeitpunkt (Heckhausen, 1974; Rheinberg, 1980). Bei einer *individuellen Bezugsnormorientierung* (IBNO) setzt die Lehrkraft die Schülerleistung in Bezug zu früheren Leistungen desselben Schülers; der Fokus liegt auf der Leistungsentwicklung eines bestimmten Schülers über die Zeit hinweg und zwar unabhängig davon wo er sich in der Leistungsverteilung der Bezugsgruppe befindet (Heckhausen, 1974; Rheinberg, 1980). Zur Erfassung der BNO von Lehrkräften wurde in bisherigen Studien meist jeweils eines von drei Instrumenten eingesetzt (Dickhäuser & Rheinberg, 2003, geben einen Überblick über die psychometrischen Eigenschaften dieser Instrumente): (a) Die *Schülerperzipierte Lehrerbezugsnorm* (SPLB) ist eine von Schwarzer, Lange und Jerusalem (1982) entwickelte Rating-Skala, welche die BNO mittels Schülerangaben bei 10 Items erfasst. (b) Der *Fragebogen zur Erfassung von Bezugsnormorientierung* (FEBO) ist ein von Rheinberg (1980) entwickelter Fragebogen für Lehrkräfte, der 39 Items mit Rating-Skalen als Antwortformat umfasst. (c) Rheinbergs (1980) *kleine Beurteilungsaufgabe* (KBA) versetzt Lehrkräfte in eine standardisierte Beurteilungssituation. Lehrkräfte beurteilen dabei für insgesamt neun fiktive Schüler jeweils die Leistungen in drei zeitlich aufeinanderfolgenden Tests. Die besagten Schülerleistungen variieren einerseits in ihrem absoluten Wert (unterdurchschnittlich, durchschnittlich oder überdurchschnittlich) und andererseits in der Tendenz der Leistungsentwicklung (abfallend, gleichbleibend oder ansteigend). Die Variation der Schülerleistung ist systematisch, was es erlaubt nachzuvollziehen, welche BNO die Lehrkraft bei ihrer Beurteilung der neun fiktiven Schülerleistungen tendenziell anwendet. Obwohl alle drei Instrumente genutzt werden können, um BNO zu messen, liegen interessanterweise bislang nur wenige Studien zur em-

pirischen Konvergenz von SPLB, FEBO und KBA vor: Zwischen dem FEBO (in Richtung IBNO gepolte Gesamtskala) und der KBA wurden für die IBNO ( $r = .33$  bei  $N = 77$  und  $r = .16$  bei  $N = 57$ ) und die SBNO ( $r = -.43$  bei  $N = 77$  und  $r = -.30$  bei  $N = 54$ ) jeweils moderate, konvergente Korrelationen beobachtet (Rheinberg, 1980). Die Zusammenhänge zwischen der SPLB einerseits und dem FEBO beziehungsweise der KBA andererseits wurden bislang noch nicht untersucht.

## **2.2 Effekte von BNO auf die Entwicklung von Lernmotivation und Leistung**

Es wird angenommen, dass sich die IBNO und SBNO von Lehrkräften unterschiedlich auf die Entwicklung von Lernmotivation und Leistung auswirken. Nach Heckhausen (1974) lenkt die IBNO die Aufmerksamkeit eines Schülers zum einen auf eine tatsächliche (intraindividuelle) Leistungsverbesserung und zum anderen auf die eigene Anstrengung und Ausdauer, die den Leistungsfortschritt bewirkten. Anstrengung und Ausdauer können willentlich gesteuert werden; dies impliziert positive Effekte auf Erfolgserwartungen und günstige Kausalattributionen des Lernerfolgs auf das eigene Handeln, die sich wiederum positiv auf die Leistungsentwicklung auswirken sollten (siehe auch Henry, Martinko & Pierce, 1993; Hosenfeld, 2002).

Erfolgserwartungen und günstige Kausalattributionen können sich aber nicht nur auf die Lernleistung selbst, sondern auch auf weitere motivational-affektive Aspekte des Lernens positiv auswirken. (a) Wie Rheinberg und Peter (1982) ausführen hängen Kausalattributionen insbesondere auch zusammen mit Angstentwicklung: Lehrkräfte mit IBNO halten Leistungserwartungen tendenziell eher in der Schwebe. Dadurch kann für den Schüler anregende Unsicherheit über den jeweiligen Effekt eigenen Bemühens erhalten bleiben. Durch die attributive Entkopplung von Begabung und Lernleistung müsste diese Unsicherheit als nicht bedrohlich empfunden werden, was somit Angst und Selbstzweifeln entgegenwirken sollte (Rheinberg & Peter, 1982). (b) Kausalattributionen und Erfolgserwartungen beeinflussen selbstbezogene Überzeugungen positiv (Stiensmeier-Pelster & Schwinger, 2008): Wer sich tendenziell mittelschwere, realistische Ziele setzt, erreicht diese auch zumeist; dies impliziert positive Effekte auf selbstbezogene Überzeugungen (Schwarzer & Jerusalem, 2002; Skaalvik, 1994; siehe auch Köller, 2004). Diese Prozesse werden zusätzlich verstärkt durch die positiven Wirkungen der Kausalattributionen auf die Leistung, da auch die Leistungsentwicklung einen positiven Einfluss auf die Entwicklung von selbstbezogenen Überzeugungen hat (Marsh & Craven, 2006; Marsh, Trautwein, Lüdtke, Köller & Baumert, 2005; siehe auch Helmke & van Aken, 1995).

Im Gegensatz zur IBNO lenkt nach Heckhausen (1974) die SBNO die Aufmerksamkeit eines Schülers auf seinen zeitlich weitestgehend stabilen, relativen Leistungsranplatz im Klassenverband, der sich auch trotz Anstrengung und Ausdauer kaum verändert (sofern die Mitschüler auch Lernfortschritte in ähnli-

chem Umfang machen). Der Fokus des Leistungsvergleichs liegt also hier auf interindividuellen Leistungsunterschieden im Klassenverband, wobei sich der Leistungsrangplatz dabei jedoch nur bedingt durch den Schüler beeinflussen lässt. Dies impliziert, dass die SBNO keine positiven Effekte auf Erfolgserwartungen, bzw. die kognitive oder motivationale Entwicklung der Schüler haben sollte, weil Schüler ihre Leistung auf zeitlich relativ stabile Faktoren wie die eigene Begabung attribuieren.

Die theoretischen Überlegungen zur Wirkung von IBNO und SBNO auf die Entwicklung von motivational-affektiven Aspekten des Lernens und auf die Lernleistung wurden durch zahlreiche Untersuchungen empirisch gestützt und fanden Eingang in die Lehrbücher (z.B. Rheinberg, 2006; 2008) und Lehrerausbildung. Die empirischen Befunde zur IBNO wurden in der Überblicksarbeit von Mischo und Rheinberg (1995, S. 141) zusammengestellt. Die Ergebnisse der darin zitierten Studien weisen darauf hin, dass sich eine stark ausgeprägte IBNO (erwartungskonform) positiv auf die Entwicklung der Lernmotivation und Leistung von Schülern auswirkt. So bewirkte eine stark ausgeprägte IBNO: (a) Günstigere Attributionsmuster und realistischere Zielsetzungen. (b) Eine Reduktion von Angst (vor Misserfolg, vor Prüfungen, sowie ganz allgemein). (c) Die Förderung von Selbstkonzepten (in den Originalarbeiten wird hier von selbstbezogenen Überzeugungen gesprochen; siehe auch Köller, 2004; Lüdtke, Köller, Marsh & Trautwein, 2005). (d) Eine Steigerung von Lern- und Schulleistungen.

Wichtig ist hierbei zu betonen, dass von einer *Interaktion* auszugehen ist: Effekte der IBNO sollten bei leistungsschwächeren Schülern stärker ausgeprägt sein als bei leistungsstarken Schülern. Die Effekte einer SBNO wurden weniger intensiv beforscht. Aufgezwungene soziale Leistungsvergleiche sollten Lernmotivation und Lernprozesse jedoch tendenziell negativ beeinflussen (Ames, 1992). Auch bei der SBNO ist wiederum von einem Interaktionseffekt auszugehen: Eine stark ausgeprägte SBNO dürfte sich vor allem bei leistungsschwächeren Schülern ungünstig auswirken (Jürgens & Sacher, 2008).

### **2.3 Forschungsanliegen dieser Arbeit**

Eine stark ausgeprägte IBNO sollte sich positiv auf leistungsbezogene Kausalattributionen und Erfolgserwartungen und darüber vermittelt auf die Entwicklung vieler zentraler Aspekte der Lernmotivation (z. B. Selbstwirksamkeit, Selbstkonzept, und Angst) und die Leistungsentwicklung auswirken. Des Weiteren sollten soziale Leistungsvergleiche, die durch eine SBNO angestoßen werden, die Entwicklung der Lernmotivation und Lernprozesse negativ beeinflussen. Die Wirkung von IBNO und SBNO sollte jeweils bei leistungsschwächeren Schülern stärker sein als bei leistungsstärkeren Schülern. Im vorliegenden Beitrag untersuchen wir diese Hypothesen für Jugendliche am Ende der Sekundarstufe I für das Fach Mathematik, um mehr über die Generalisierbarkeit bisheriger BNO-Forschung in Erfahrung zu bringen.

(1) Bislang gab es kaum Studien, in denen BNO mit mehreren Messinstrumenten erfasst wurde. Daher steht der empirische Nachweis zur konvergen-ten und diskriminanten Validität dieser Instrumente noch weitestgehend aus. So ist insbesondere auch nicht klar, ob die unterschiedlichen Instrumente tatsächlich das jeweils gleiche Konstrukt erfassen und inwieweit sich Befunde zur positiven Wirkung auf motivational-affektive Aspekte des Lernens und die Lernleistung über die verschiedenen Instrumente hinweg generalisieren lassen. Im vorliegenden Beitrag erfassen wir daher die BNO der Mathematiklehrkräfte mit der KBA, einer Kurzfassung des FEBOs sowie einer Kurzfassung der SPLB. Die Frage nach der Generalisierbarkeit ist insbesondere auch deshalb interessant, weil die aktuelle multimethodale Unterrichtsforschung (Clausen, 2002; Kunter & Baumert, 2006) zeigt, dass die Unterrichtswahrnehmung durch Schüler und Lehrer differentiell valide ausfällt.

(2) Schulformen stellen institutionell vorgeformte *differentielle Entwicklungsmilieus* dar (Baumert, Trautwein & Artelt, 2003), da sie sich u.a. systematisch in der Zusammensetzung der Schülerschaft, dem Lehrpersonal, didaktischen Traditionen, den Lehrplänen und auch in Dimensionen der Unterrichtsqualität unterscheiden. Es ist nicht auszuschließen, dass solche systematischen Unterschiede zwischen den Schulformen die Art und Weise, wie BNO wirkt, beeinflussen können. Daher untersuchen wir – im Gegensatz zu den meisten bisherigen Studien – die Frage, ob sich Ergebnisse über Schulformen hinweg generalisieren lassen.

(3) Die strukturelle Konzeption von BNO wurde noch nicht endgültig geklärt (siehe auch Lüdtke et al., 2005): Stellen IBNO und SBNO (a) jeweils eigenständige eindimensionale Konstrukte dar, handelt es sich (b) um ein eindimensionales bipolares Konstrukt mit den Polen IBNO vs. SBNO oder (c) um ein typologisches Konstrukt? In der bisherigen BNO-Forschung dominierte der Vergleich von Typen. Ein Großteil der Studien teilte die Lehrkräfte mittels Median-Split in zwei Gruppen. Wie MacCallum, Zhang, Preacher und Rucker (2002) ausführen, führt dies jedoch dazu, dass bivariate Zusammenhänge unterschätzt und multivariate Zusammenhänge „verschätzt“ werden. Einige Studien nutzten einen Extremgruppenvergleich; auch dies ist sehr kritisch, denn der selektive Ausschluss von Mischtypen verzerrt die Resultate insofern, als dass potentiell vorhandene Effekte künstlich verstärkt und infolgedessen überschätzt werden (Preacher, Rucker, MacCallum & Nicewander, 2005). In der vorliegenden Arbeit fassen wir daher IBNO und SBNO als jeweils eindimensionale Konstrukte auf, um interindividuelle Unterschiede zwischen Lehrkräften zu beschreiben und den Problemen effektiv zu begegnen, die sich durch einen Vergleich von Typen ergeben.

(4) Eine Kernannahme in der Unterrichtsforschung ist, dass kognitive und motivationale Schülermerkmale durch Merkmale des Unterrichts oder Merkmale einer Lehrkraft (z.B. die IBNO) beeinflusst werden. Zentral ist hierbei, dass Schüler bestimmten Klassen zugeordnet sind, also eine natürlich bedingte hierarchische Datenstruktur vorliegt. Dies impliziert die Anwendung von statistischen Mehrebenenmodellen (z.B. *Hierarchical Linear Models*, HLM), in denen Merkmale der Lehrkraft beziehungsweise des Unterrichts (Ebene 2) Merkmale

der Schüler (Ebene 1) beeinflussen. Wenn die hierarchische Datenstruktur nicht adäquat berücksichtigt wird, besteht die große Gefahr, dass Effekte der Lernumwelt mit Individualeffekten konfundiert werden und dass Standardfehler von Modellparametern „verschätzt“ werden, was zu verzerrten statistischen Schlussfolgerungen führen kann (z.B. Raudenbush & Bryk, 2002). Mit Ausnahme von Köller (2004) und Lüdtke und Kollegen (2005) wurden in bisherigen Studien zur Wirkung von BNO keine HLM-Modelle zur Analyse herangezogen. Beide Studien beschäftigten sich mit der Frage, ob eine IBNO den sogenannten *Big-fish-little-pond-effect* (BFLPE; Marsh, 1987), in dem sich negative Konsequenzen sozialer Leistungsvergleiche auf das fachspezifische Selbstkonzept abbilden, unterdrücken oder wenigstens teilweise kompensieren kann. Analog zu der vorliegenden Arbeit steht in beiden Studien jeweils das Fach Mathematik im Mittelpunkt des Interesses. Eine Teilstichprobe ( $n_{\text{Schüler}} = 4250$ ;  $n_{\text{Klassen}} = 236$ ) von Siebtklässlern aus der Studie *Bildungsverläufe und psychosoziale Entwicklung im Jugend- und jungen Erwachsenenalter* (BIJU) bildet die Datengrundlage der Arbeit von Köller (2004). Die Studie von Lüdtke und Kollegen (2005) basiert auf einer Teilstichprobe ( $n_{\text{Schüler}} = 2150$ ;  $n_{\text{Klassen}} = 112$ ) von Neuntklässlern aus der *Third International Mathematics and Science Study* (TIMSS). Sowohl Köller (2004) als auch Lüdtke und Kollegen (2005) konnten (auf Klassenebene) kleine positive IBNO-Effekte auf das mathematische Selbstkonzept nachweisen; eine IBNO hatte jedoch nur bei Köller (2004) einen kleinen positiven Einfluss auf die Größe des BFLPE.

(5) Die Mehrheit der Befunde zur Wirkung von BNO entstammt Studien Ende der Grundschule bzw. Beginn der Sekundarstufe I. Nur wenige Studien konnten bislang Effekte von BNO auf motivational-affektive Schülermerkmale oder die Lernleistung am Ende der Sekundarstufe I nachweisen (siehe Krampen, 1985; Lüdtke et al., 2005; Rheinberg & Hendricks, 1980; Rheinberg, Kühmel & Duscha, 1979). Inwieweit die Erkenntnisse auch für Jugendliche am Ende der Sekundarstufe I generalisierbar sind, scheint somit empirisch wenig gesichert. So könnten Effekte in jüngeren Jahrgangsstufen größer sein, da Schülerleistung (z.B. Hill, Bloom, Black & Lipsey, 2008), sowie die Persönlichkeit (z.B. Caspi & Roberts, 2001) hier noch größeren Veränderungen unterliegen.

### 3. Methodik

#### 3.1 Stichprobe

Die vorliegenden Analysen basieren auf den Daten der Studie *Professionswissen von Lehrkräften, kognitiv aktivierender Mathematikunterricht und die Entwicklung von mathematischer Kompetenz* (COACTIV; Kunter et al., 2011), welche in die nationale Ergänzung des *Programme for International Student Assessment* (PISA) 2003/2004 eingebunden war. Teilnehmende Schulen wurden, proportional nach Bundesland und Schulform, per Zufallsprinzip ausgewählt. Für jede

ausgewählte Schule wurden, wiederum per Zufallsprinzip, zwei vollständige 9. Klassen ausgewählt. Zusätzlich zur Schülerstichprobe stand eine Stichprobe der Mathematiklehrkräfte genau dieser Klassen zur Verfügung. Für beide Stichproben lagen zwei Messzeitpunkte vor (für Angaben zur Rücklaufquote siehe Baumert et al., 2006): der erste *gegen Ende der 9. Klasse* (K9; Erhebungszeitraum: April–Mai 2003) und der zweite *gegen Ende der 10. Klasse* (K10; Erhebungszeitraum: April–Mai 2004). Da Hauptschüler in der Regel nach der 9. Klasse ihren Schulabschluss machen und von der Schule abgehen, lagen für die COACTIV-Erhebung 2004 keine Daten von Hauptschülern bzw. Hauptschullehrkräften vor.

Ein Schwerpunkt der vorliegenden Studie ist die multimethodale Erfassung der BNO der Lehrkräfte sowie die Analyse der Wirkung dieser (unterschiedlich erfassten) BNO auf die Entwicklung von Schülermerkmalen. Zur Stichprobendefinition in dieser Studie haben wir daher zwei Kriterien kombiniert angewendet: Wir haben nur Lehrkräfte ausgewählt, die (a) dieselbe Klasse in der 9. und 10. Jahrgangsstufe unterrichteten und (b) für welche vollständige Daten für die KBA vorlagen, da dieses als einziges BNO-Messinstrument in der ungekürzten Originalfassung eingesetzt wurde. Somit basieren die Analysen der vorliegenden Arbeit auf Daten von 106 Mathematiklehrkräften (50 am Gymnasium und 56 an der Realschule; Alter:  $M = 47.4$  Jahre, Spannweite 28 bis 65 Jahre; Geschlecht: 45.7 % weiblich; Berufserfahrung:  $M = 20.8$  Jahre, Spannweite 4 bis 42 Jahre) aus 73 verschiedenen Schulen, die 106 Klassen mit insgesamt 2465 Schülern (1153 am Gymnasium und 1312 an der Realschule) unterrichteten. Für die gesamte Schülerstichprobe waren die durchschnittliche Klassengröße 23.9 Schüler (Spannweite 12 bis 30), das durchschnittliche Alter 15.1 Jahre (Spannweite 13 bis 18 Jahre) und der Anteil der Mädchen betrug 57.8 %.

## 3.2 Instrumente

### 3.2.1 Erfassung der BNO

In COACTIV wurden die BNO multimethodal erfasst (Baumert et al., 2006). Die KBA (Rheinberg, 1980) wurde in K10 erhoben und kam in der Originalfassung zum Einsatz. Aufgrund der Berechnungsvorschrift für die Skalenwerte der IBNO und SBNO gibt es unseres Wissens derzeit keine Möglichkeit, die Skalenreliabilitäten (im Sinne einer internen Konsistenz) zu berechnen. Die *BNO aus Sicht der Lehrkräfte* wurde in K9 anhand von zwei Skalen erfasst, welche sich jeweils aus Items des FEBO (Rheinberg, 1980) zusammensetzten und sich in dieser Form auch bei der nationalen Erweiterung der TIMS-Studie bewährten (siehe Clausen, 2002). Die IBNO-Skala umfasste fünf Items ( $\alpha = .69$ ; Itembeispiel: „Wenn ich von einer ‚schlechten Leistung‘ spreche, liegt sie unter der vorhergehenden Leistung des jeweiligen Schülers.“). Die SBNO-Skala umfasste zwei Items ( $\alpha = .74$ ; Beispielitem: „Ich orientiere meine Noten am Durchschnitt der Klasse.“). Die *BNO aus Sicht der Schüler* wurde in K9 anhand von zwei Skalen erfasst, welche jeweils Items der

SPLB (Schwarzer et al., 1982) nutzten und in dieser Form auch in den Studien von Köller (2004) und Lüdtke und Kollegen (2005) verwendet wurden. Die IBNO-Skala umfasste vier Items (Beispielitem: „Wenn sich ein schwacher Schüler verbessert, ist das für unseren Mathematiklehrer eine gute Leistung, auch wenn der Schüler immer noch unter dem Klassendurchschnitt liegt.“). Die SBNO-Skala umfasste zwei Items (Beispielitem: „Unser Mathematiklehrer lobt einen Schüler nur, wenn er sich im Vergleich zu seinen Klassenkameraden verbessert hat.“). Da die BNO aus Schülersicht auf Klassenebene modelliert wurde, sind für die Bewertung der Reliabilität der Skalen die Werte der ICC(2) maßgeblich (siehe Tabelle 2 in Kapitel 4.2). Zur besseren Interpretation der statistischen Koeffizienten wurden alle BNO-Skalen über die Gesamtstichprobe von 106 Lehrkräften z-standardisiert ( $M = 0$ ,  $SD = 1$ ).

### 3.2.2 Erfassung mathematikspezifischer Schülermerkmale

Zentraler Indikator der Schülerleistung war die *Mathematikkompetenz*, die durch die curricular-sensitive Skala des Mathematiktests der PISA-2004-Studie ( $\alpha_{K9} = .64$ ) gemessen wurde. Diese Kompetenzskala (Baumert et al., 2010) ist Rasch-homogen; wir verwendeten hier einen entsprechenden Personenparameter (WLE-Score). Alle weiteren Instrumente (inkl. ihrer psychometrischen Kennwerte) sind detailliert im PISA 2003 Skalenhandbuch (Ramm et al., 2006) beschrieben. Die *Mathematiknoten* sind durch Schülerangaben („Welche Note hattest du in deinem letzten Zeugnis in Mathematik?“) erhoben worden. Für die statistischen Analysen wurde die ursprüngliche Notenskala invertiert, so dass höhere Werte für bessere Leistungen stehen (1 = ungenügend bis 6 = sehr gut). Die *Mathematikangst* wurde mit neun Items erfasst ( $\alpha_{K9} = .90$ ; „Vor Mathematik habe ich Angst.“). Das *Selbstkonzept in Mathematik* wurde mit fünf Items erfasst ( $\alpha_{K9} = .92$ ; „Im Mathematikunterricht verstehe ich sogar die schwierigsten Aufgaben.“). Die *Selbstwirksamkeit in Mathematik* wurde mit acht Items erfasst ( $\alpha_{K9} = .79$ ; „Wie sicher glaubst Du, eine Mathematikaufgabe lösen zu können, in welcher es darum geht, anhand eines Zugfahrplans, die Dauer einer Fahrt zwischen zwei Ortschaften zu berechnen?“). Für die statistischen Analysen wurden alle Skalen (mit Ausnahme der Noten) über die Gesamtstichprobe unserer Studie z-standardisiert. Die Stabilität der mathematikspezifischen Schülermerkmale über die beiden Erhebungszeitpunkte K9 und K10 war mit einer mittleren Korrelation von  $r = .67$  (Spannweite  $r = .52$  bis  $r = .83$ ) als hoch bis sehr hoch zu bewerten (siehe auch Tabelle 2).

### 3.2.3 Auswahl und Erfassung der Kontrollvariablen

Um die Effekte der BNO auf die Entwicklung der mathematikspezifischen Schülermerkmale zu isolieren, wurden mehrere in K9 erhobene Kontrollvariablen auf Schülerenebene in das HLM-Modell aufgenommen. Grundsätzlich folgten wir bei der Auswahl der Kontrollvariablen einschlägigen methodischen Arbeiten (Antonakis & Dietz, 2011a; 2011b; Becker, 2005; Spector & Brannick, 2010) und haben Kontrollvariablen berücksichtigt für die theoretische Erklärungsmechanismen, bzw. empirische Evidenz vorliegt, dass sie das jeweilige mathematische Ergebniskriterium beeinflussen können. Zu den Kontrollvariablen gehörten (a) Alter, (b) Geschlecht, (c) die Lesekompetenz (PISA) und (d) die kognitiven Grundfähigkeiten (erfasst über Figurenalogien; Heller & Perleth, 2000). (e) Als Indikator des sozioökonomischen Familienhintergrunds wurde der höchste ISEI-Wert (Ganzeboom, De Graaf & Treiman, 1992) innerhalb einer Familie verwendet; zur Berechnung des ISEI wurden die Angaben der Schüler zum Beruf ihrer Eltern herangezogen. (f) Der Bildungshintergrund der Eltern wurde mit fünf Dummy-kodierten Qualifikationsindikatoren gemessen. (g) Der Migrationshintergrund wurde über das Geburtsland der Eltern definiert: Wenn zumindest ein Elternteil *nicht* in Deutschland geboren ist, so ging dieser Schüler als Schüler *mit* Migrationshintergrund in die Analysen ein.

Für die Leistungskriterien (Kompetenztest und Schulnoten) bildeten die Modelle schulischen Lernens (Baumert, Stanat & Demmrich, 2001; Carroll, 1963; 1989; Haertel, Walberg & Weinstein, 1983; Helmke & Schrader, 2010; Helmke & Weinert, 1997; Walberg, 1986; Wang, Haertel & Walberg, 1993) die theoretische Grundlage für die Auswahl unserer Kontrollvariablen. Diese Modelle synthetisieren eine Vielzahl von empirischen Studien in denen wiederholt gezeigt wurde, dass die schulische Leistung durch entwicklungsrelevante Faktoren, wie Alter, Geschlecht, sozioökonomischer Status, Bildungs- und/oder Migrationshintergrund, Schlussfolgerndes Denken und Vorwissen beeinflusst wird. Die Leseleistung ist insbesondere deshalb ausgewählt worden, da Lesen zentral ist, um eine mentale Problemrepräsentation bei mathematischen Textaufgaben aufzubauen, bzw. Lesen generell notwendig ist, um aus Mathematiklehrbüchern zu lernen (Carroll, 1996; Klieme, Neubrand & Lüdtke, 2001; Mayer & Hegarty, 1996). Die Leseleistung kann also über das mathematische Vorwissen in K9 hinaus, einen positiven Effekt auf die Mathematikleistung in K10 haben, der unabhängig von Effekten der weiteren Kontrollvariablen ist.

Auch für die Analyse der Entwicklung der Lernmotivation sollten diese Schülermerkmale als Kontrollvariablen aufgenommen werden. So stellt zum Beispiel Meece (2002) fest, dass die schulische Motivation durch den sozioökonomischen Status der Schüler beeinflusst wird: Kinder aus einem sozioökonomisch benachteiligten Umfeld sind schulisch weniger motiviert als Schüler aus einem sozioökonomisch privilegierten Umfeld (Borkowski & Thorpe, 1994). Weiterhin konnten beispielsweise Arbeiten von Brunner und Kollegen (2008; 2010) zeigen, dass das mathematische Selbstkonzept positiv mit dem Schlussfolgernden Denken (bei uns

durch die Figurenalogien des *Kognitiven Fähigkeitstests*, KFT; Heller & Perleth, 2000, erfasst) und negativ mit der Leseleistung kovariiert.

### 3.3 Statistische Analysen

Fehlende Daten stellen in längsschnittlich angelegten *large-scale* Untersuchungen eine unvermeidbare Realität dar.<sup>1</sup> In Anlehnung an Schafer und Graham (2002) erstellten wir daher 10 voneinander unabhängige Datensätze, in denen mittels eines multiplen Imputationsverfahrens fehlende Werte ersetzt wurden. Alle Analysen wurden dann für jeden der 10 Datensätze getrennt durchgeführt und anschließend mit den von Schafer und Graham (2002) beschriebenen Methoden zusammengefasst.

Die Auswirkungen der BNO einer Lehrkraft auf mathematikspezifische Schülermerkmale in K10 manifestieren sich als Effekt der Lernumwelt auf Klassenebene: *Ceteris paribus* sollte also zum Beispiel eine stärker ausgeprägte IBNO einer Lehrkraft mit einem niedriger ausgeprägten Klassenmittelwert der Mathematikangst einhergehen. Die gleichen Überlegungen gelten auch für die anderen Schülermerkmale: Hier sollte eine höher ausgeprägte IBNO mit höher ausgeprägten Klassenmittelwerten der mathematikspezifischen Schülermerkmale (z.B. Selbstkonzept) einhergehen. Es ist wichtig zu betonen, dass alle Analysen getrennt für die Schulformen Gymnasium und Realschule durchgeführt wurden. So kontrollierten wir (eventuelle) schulformspezifische Effekte der BNO auf die Entwicklung der mathematikspezifischen Schülermerkmale und berücksichtigten, dass Schulformen differenzielle Entwicklungsmilieus darstellen.

Zur Analyse der Effekte von Lernumwelten folgten wir den Empfehlungen von Lüdtke, Trautwein, Kunter und Baumert (2006). In einem ersten Schritt untersuchten wir die *Intraklassenkorrelationen* ICC(1) und ICC(2). Die Werte für die Intraklassenkorrelationen sind in Tabelle 2 dargestellt. Die ICC(1) gibt den prozentualen Anteil der Gesamtvarianz an, welcher auf mittlere Unterschiede zwischen Klassen zurückzuführen ist (Lüdtke et al., 2006). Die ICC(2) stellt einen Indikator der Reliabilität der Klassenmittelwerte dar. In die Berechnung der ICC(2) gehen die ICC(1), sowie die durchschnittliche Klassengröße (23.1 am Gymnasium und 23.4 an der Realschule) ein. Werte der ICC(2) können zwischen 0 und 1 variieren; höhere Werte indizieren eine höhere Reliabilität des jeweiligen Klassenmittelwerts. Komplexere HLM-Modelle sind vorzugsweise dann zu analysieren, wenn die Werte von ICC(1) substantiell sind und wenn für aggregierte Variablen zur Wahrnehmung der Lernumwelt (hier die schülerperzipierte BNO)

1 Bei der KBA gab es keine fehlenden Werte aufgrund der Stichprobendefinition. Beim FEBO fehlten 8 % der Werte. Bei der SPLB fehlten 47 % der Werte aufgrund des Multimatixdesigns von COACTIV bzw. PISA; hierbei wurde aber sichergestellt, dass in jeder Klasse mindestens 10 Jugendliche diese Items beantworteten. Bei den mathematikspezifischen Schülermerkmalen und den Kontrollvariablen fehlten – mit einer Ausnahme – jeweils maximal 5 % der Werte; aufgrund des Multimatixdesigns fehlten 48 % der Werte bei der Lesekompetenz.

auf Klassenebene die Werte der ICC(2) zumindest ausreichend sind. Darüber hinaus sind aber Mehrebenenanalysen auch dann anzuwenden, wenn substanztheoretische Überlegungen dafür sprechen (Snijders & Bosker, 2011, S. 106): BNO ist ein Merkmal der Lernumwelt und ihre Wirkung sollte sich auf Klassenebene entfalten. Aus diesem Grund verwendeten wir Mehrebenenmodelle, obwohl einige ICC-Werte relativ gering ausfielen (siehe Tabelle 2).

Im zweiten Analyseschritt orientierten wir uns bei der Spezifikation der HLM-Modelle eng an den Studien von Köller (2004) als auch Lüdtke und Kollegen (2005): Wir nutzten *random-intercept* HLM-Modelle, in denen ein bestimmtes mathematikspezifisches Schülermerkmal in K10 durch Variablen auf Schüler- und Klassenebene vorhergesagt wurde. Insbesondere wurde zusätzlich zu den Kontrollvariablen das jeweilige mathematikspezifische Schülermerkmal (erhoben in K9; zentriert am Gesamtmittelwert) auf Schülerebene in das HLM-Modell integriert: (Potentielle) BNO-Effekte können demnach im Sinne eines Effekts auf die Entwicklung der mathematikspezifischen Schülermerkmale von K9 nach K10 interpretiert werden. Auf Klassenebene wurde jeweils ein Indikator der IBNO *oder* der SBNO eingeführt.

In einem dritten Schritt untersuchten wir die Interaktionshypothese (stärkere BNO-Effekte bei leistungsschwächeren Schülern), indem wir Kompetenz und Noten als Indikatoren der Mathematikleistung jeweils getrennt analysierten. Zur Analyse der Interaktion erweiterten wir die *random-intercept*-Modelle aus Schritt 2 als *random-slope*-Modelle und prüften, ob die IBNO *oder* SBNO einer Lehrkraft den Zusammenhang (*slope*) beeinflusst, wie sich die Kompetenz in K9 (diesmal eingeführt als klassenzentrierte Variable; siehe hierzu die Empfehlung von Enders & Tofghi, 2007, S. 136) auf die Kompetenz in K10 auswirkt. Für die Analyse zur Interaktion von BNO und Mathematiknoten gingen wir analog vor.

## 4. Ergebnisse

### 4.1 Konvergenz der Erfassungsmethoden der BNO

Betrachten wir die Korrelationen der BNO, die über verschiedene Erhebungsmethoden erfasst wurde (Tabelle 1), so fällt auf, dass die verschiedenen BNO-Erhebungsmethoden nur mäßig konvergierten (siehe fett gedruckte Koeffizienten). Erwartungsgemäß hatten Gymnasiallehrkräfte mit einer stärker ausgeprägten IBNO (erfasst über die KBA) zwar tendenziell auch eine stärker ausgeprägte IBNO (erfasst über Lehrerselbstberichte), doch der Zusammenhang war mit  $r = .34$  nur moderat und nicht konsistent über Schulformen. Weiterhin konvergierten die Befunde der KBA und Lehrerselbstberichte nicht mit der Wahrnehmung der BNO durch die Schüler: Beispielsweise hatten schulformübergreifend Lehrkräfte mit einer stark ausgeprägten IBNO (erfasst über die KBA) tendenziell eine geringer ausgeprägte IBNO aus Sicht ihrer Schüler ( $r_{\text{Gymnasium}} = -.07$ ;  $r_{\text{Realschule}} = -.11$ ).

Tabelle 1: Interkorrelationen und deskriptive Statistiken der Bezugsnormorientierung in Abhängigkeit von Erhebungsmethode und Schulform

Bezugsnorm-orientierung der Lehrkraft nach Erhebungsmethode	KBA		Lehrerangabe		Schülerangabe	
	IBNO	SBNO	IBNO	SBNO	IBNO	SBNO
Gymnasium						
KBA						
IBNO	–					
SBNO	<b>-.50</b>	–				
Lehrerangabe						
IBNO	<b>.34</b>	<b>-.14</b>	–			
SBNO	<b>-.14</b>	<b>.26</b>	<b>.07</b>	–		
Schülerangabe						
IBNO	<b>-.07</b>	<b>-.24</b>	<b>.03</b>	<b>-.34</b>	–	
SBNO	<b>.10</b>	<b>.08</b>	<b>.09</b>	<b>-.11</b>	<b>-.21</b>	–
Mittelwert	<b>-.19</b>	<b>.09</b>	<b>-.14</b>	<b>-.11</b>	<b>-.22</b>	<b>-.16</b>
Standardabweichung	<b>.93</b>	<b>.94</b>	<b>.82</b>	<b>.89</b>	<b>.90</b>	<b>.92</b>
Realschule						
KBA						
IBNO	–					
SBNO	<b>-.56</b>	–				
Lehrerangabe						
IBNO	<b>.19</b>	<b>.13</b>	–			
SBNO	<b>-.04</b>	<b>.09</b>	<b>.25</b>	–		
Schülerangabe						
IBNO	<b>-.11</b>	<b>.16</b>	<b>.36</b>	<b>-.06</b>	–	
SBNO	<b>-.21</b>	<b>.11</b>	<b>-.16</b>	<b>.05</b>	<b>-.19</b>	–
Mittelwert	<b>.17</b>	<b>-.08</b>	<b>.13</b>	<b>.10</b>	<b>.20</b>	<b>.14</b>
Standardabweichung	<b>.87</b>	<b>.88</b>	<b>.97</b>	<b>.92</b>	<b>.88</b>	<b>.89</b>
Gymnasium vs. Realschule						
Mittelwertvergleich ( <i>d</i> )	<b>-.37</b>	<b>.18</b>	<b>-.27</b>	<b>-.21</b>	<b>-.43</b>	<b>-.30</b>

Allgemeine Anmerkungen. IBNO = Individuelle Bezugsnormorientierung; SBNO = Soziale Bezugsnormorientierung; KBA = kleine Beurteilungsaufgabe  
 Korrelationskoeffizienten der gleichen Bezugsnormorientierung, die mit unterschiedlichen Methoden erhoben wurden, sind fett gedruckt.

*d* gibt die Effektstärke des Mittelwertunterschieds zwischen Gymnasium und Realschule an. Positive Werte von *d* bedeuten, dass der Mittelwert am Gymnasium größer war als an der Realschule.

$N_{\text{Klassen}}(\text{Gymnasium}) = 50$ ;  $N_{\text{Klassen}}(\text{Realschule}) = 56$

Mit Blick auf Schulformunterschiede in der BNO (siehe Tabelle 1), stellen wir fest, dass sich die Schulformen hier teilweise deutlich unterschieden. Aus Sicht der Gymnasiasten war die IBNO ihrer Mathematiklehrkräfte zum Beispiel im Mittel deutlich geringer als die IBNO von Mathematiklehrkräften aus Sicht von Realschülern ( $d = -.43$ ). Weiterhin ist festzuhalten, dass (wie aufgrund der mäßigen Konsistenz der verschiedenen Erhebungsmethoden zu erwarten war) die Richtung der Schulformunterschiede nicht konsistent für die verschiedenen Erhebungsmethoden war.

## 4.2 Analyse der Effekte der BNO auf die Entwicklung von Lernmotivation und Leistung

Zur Analyse der Effekte der BNO auf die Entwicklung von Lernmotivation und Leistung untersuchten wir zunächst die Intraklassenkorrelationen ICC(1) und ICC(2) (Tabelle 2). Insgesamt ist festzustellen, dass die Werte der ICC(1) für alle untersuchten Variablen statistisch signifikant von Null verschieden waren, jedoch der Anteil der Varianz, der auf Unterschiede zwischen Klassen zurückgeht, jeweils relativ klein war. Weiterhin ist für die Schülerangaben zur IBNO und SBNO (siehe Tabelle 2) festzuhalten, dass in der IBNO schulformübergreifend deutliche Unterschiede zwischen Klassen, nicht aber in der SBNO bestanden. Da die ICC(1) in die Berechnung der ICC(2) mit einfließt, impliziert dies gleichzeitig eine relativ geringe Reliabilität der schülerberichteten SBNO auf Klassenebene.

Tabelle 2: Intraklassenkorrelationen der mathematikspezifischen Schülermerkmale und Schülerangaben zur Bezugsnormorientierung der Lehrkraft in Abhängigkeit der Schulform

Variable	Gymnasium						Realschule					
	$\tau^2$	$p$	$\sigma^2$	ICC(1)	ICC(2)	$r$	$\tau^2$	$p$	$\sigma^2$	ICC(1)	ICC(2)	$r$
Mathematikspezifisches Schülermerkmal in der 10. Klasse												
Angst	.04	.00	.92	.04	.47	.69	.02	.01	1.02	.02	.30	.70
Selbstkonzept	.01	.08	1.00	.01	.21	.83	.02	.01	.97	.02	.35	.78
Selbstwirk-samkeit	.04	.00	.81	.05	.53	.61	.04	.00	.99	.04	.50	.64
Kompetenz	.06	.00	.72	.08	.67	.52	.15	.00	.72	.18	.83	.56
Noten	.08	.00	.95	.07	.65	.64	.08	.00	.99	.07	.64	.62
Schülerangabe zur Bezugsnormorientierung der Lehrkraft												
IBNO	.12	.00	.86	.13	.77	–	.12	.00	.87	.12	.76	–
SBNO	.02	.01	.92	.02	.34	–	.01	.17	1.04	.01	.15	–

Allgemeine Anmerkungen. IBNO = Individuelle Bezugsnormorientierung; SBNO = Soziale Bezugsnormorientierung;  $\tau^2$  = Varianz zwischen Klassen;  $\sigma^2$  = Varianz innerhalb von Klassen

ICC(1) = Intraklassenkorrelation berechnet als  $\rho = \tau^2 / (\tau^2 + \sigma^2)$

ICC(2) = Reliabilität der Klassenmittelwerte berechnet als  $Rel(\rho) = (k * \rho) / [1 + (k - 1) * \rho]$

$r$  = Korrelation des gleichen Merkmals in der 9. und 10. Klasse

$N_{Schüler}(\text{Gymnasium}) = 1153$ ;  $N_{Klassen}(\text{Gymnasium}) = 50$ ;  $N_{Schüler}(\text{Realschule}) = 1312$ ;  $N_{Klassen}(\text{Realschule}) = 56$

$p = p$ -Wert von  $\tau^2$

Tabelle 3: Zusammenhang zwischen Bezugsnormorientierung und mathematikspezifischen Schülermerkmalen (*random-intercept*-Modell) in Abhängigkeit von Erhebungsmethode und Schulform

Mathematikspezifisches Schülermerkmal in der 10. Klasse	KBA			Lehrerangabe			Schülerangabe		
	$\gamma$	CI	<i>p</i>	$\gamma$	CI	<i>p</i>	$\gamma$	CI	<i>p</i>
Individuelle Bezugsnormorientierung									
Gymnasium									
Angst	.02	(-.03; .07)	.51	.03	(-.02; .08)	.33	-.03	(-.07; .01)	.19
Selbstkonzept	<b>.05</b>	(.00; .09)	.04	.02	(-.03; .07)	.42	-.03	(-.08; .01)	.19
Selbstwirksamkeit	.00	(-.07; .07)	.95	-.01	(-.09; .07)	.82	.05	(-.02; .11)	.18
Kompetenz	-.03	(-.09; .03)	.29	<b>-.06</b>	(-.12; .00)	.04	<b>.06</b>	(.01; .11)	.02
Noten	<b>.09</b>	(.01; .18)	.04	.04	(-.05; .13)	.38	-.04	(-.11; .03)	.24
Realschule									
Angst	.02	(-.04; .08)	.55	.00	(-.04; .05)	.84	-.02	(-.08; .04)	.59
Selbstkonzept	-.01	(-.06; .03)	.55	.00	(-.04; .05)	.86	.02	(-.03; .06)	.40
Selbstwirksamkeit	.03	(-.03; .09)	.29	.01	(-.05; .06)	.86	.00	(-.05; .05)	.97
Kompetenz	-.01	(-.08; .07)	.88	-.01	(-.06; .05)	.85	.02	(-.04; .09)	.50
Noten	.02	(-.05; .09)	.61	.01	(-.04; .07)	.64	.05	(-.02; .12)	.15
Soziale Bezugsnormorientierung									
Gymnasium									
Angst	.00	(-.04; .05)	.85	-.03	(-.07; .02)	.30	.00	(-.05; .05)	.96
Selbstkonzept	.00	(-.04; .04)	.90	-.01	(-.06; .03)	.59	.04	(-.01; .08)	.10
Selbstwirksamkeit	-.04	(-.09; .02)	.18	.01	(-.06; .08)	.71	-.01	(-.07; .05)	.78
Kompetenz	-.02	(-.07; .03)	.38	-.02	(-.08; .04)	.53	-.02	(-.08; .05)	.60
Noten	-.04	(-.11; .03)	.27	<b>-.11</b>	(-.20; -.01)	.03	.05	(-.05; .14)	.37
Realschule									
Angst	.01	(-.05; .07)	.76	.00	(-.06; .06)	.99	-.03	(-.10; .04)	.37
Selbstkonzept	-.01	(-.06; .03)	.53	.00	(-.04; .05)	.92	.02	(-.03; .07)	.49
Selbstwirksamkeit	.00	(-.06; .05)	.92	.01	(-.05; .06)	.76	-.04	(-.11; .02)	.21
Kompetenz	.02	(-.05; .09)	.64	.01	(-.05; .06)	.83	<b>-.07</b>	(-.13; -.01)	.04
Noten	-.01	(-.08; .06)	.73	-.03	(-.10; .04)	.35	.01	(-.07; .10)	.73

Allgemeine Anmerkungen. KBA = kleine Beurteilungsaufgabe;  $\gamma$  = HLM-Regressionskoeffizient; CI = 95 % Konfidenzintervall

Auf Schülerebene wurde in allen Modellen für Geschlecht, Alter, Migrationsstatus, sozioökonomischer Status (höchster ISEI), Lesekompetenz (PISA), kognitive Fähigkeiten (KFT), Ausbildung der Eltern und das jeweilige mathematische Schülermerkmal in der 9. Klasse kontrolliert.

$N_{\text{Schüler}}(\text{Gymnasium}) = 1153$ ;  $N_{\text{Klassen}}(\text{Gymnasium}) = 50$ ;  $N_{\text{Schüler}}(\text{Realschule}) = 1312$ ;  $N_{\text{Klassen}}(\text{Realschule}) = 56$   
HLM-Regressionskoeffizient mit *p*-Werten  $\leq .05$  sind fett gedruckt.

Wir gehen nun auf die Wirkung der BNO auf die Entwicklung der mathematikspezifischen Schülermerkmale ein. Tabelle 3 stellt die wichtigsten Befunde dar. Erstens, schulformübergreifend und methodenübergreifend waren die Effekte ( $\gamma$ ) der BNO auf die Entwicklung der mathematikspezifischen Schülermerkmale jeweils

meist klein und meist nicht statistisch signifikant von Null verschieden. Manche Zusammenhänge zwischen BNO und Schülermerkmalen waren zwar statistisch signifikant (siehe fett gedruckte Koeffizienten), jedoch nicht erwartungskonform: So hatte beispielsweise die lehrerberichtete IBNO am Gymnasium einen negativen Effekt von  $\gamma = -.06$  auf die Mathematikkompetenz in K10. Gymnasiasten einer Klasse, deren Lehrkraft eine IBNO von einer Standardabweichung *über* dem Mittelwert aufwies, hatten demnach im Mittel eine Mathematikkompetenz, die um  $.06$  Standardabweichungen *unter* dem Durchschnitt der Gesamtstichprobe lag. Zweitens, die Effekte von IBNO und SBNO auf die mathematikspezifischen Schülermerkmale waren nicht konsistent für die beiden Schulformen und die drei Erhebungsmethoden. Die mit den Daten der COACTIV-Studie erhaltenen Befunde zu Effekten der BNO lassen sich folglich weder über Erhebungsmethoden noch über Schulformen hinweg uneingeschränkt generalisieren.

Die Überprüfung der Interaktionshypothese (Tabelle 4) zeigte, dass die IBNO keinen stärkeren positiven, und die SBNO keinen stärkeren negativen Einfluss auf die Mathematikleistung bei leistungsschwächeren Schülern hatte; ausnahmslos alle Interaktionseffekte waren sowohl für Gymnasium, als auch für Realschule klein und statistisch nicht signifikant von Null verschieden.

Tabelle 4: Zusammenhang zwischen Bezugsnormorientierung und mathematikspezifischen Schülermerkmalen (*random-slope*-Modell) in Abhängigkeit von Erhebungsmethode und Schulform

Mathematikspezifisches Schülermerkmal in der 10. Klasse	KBA				Lehrerangabe				Schülerangabe			
	Haupteffekt		Interaktionseffekt		Haupteffekt		Interaktionseffekt		Haupteffekt		Interaktionseffekt	
	$\gamma$	<i>p</i>	$\gamma$	<i>p</i>	$\gamma$	<i>p</i>	$\gamma$	<i>p</i>	$\gamma$	<i>p</i>	$\gamma$	<i>p</i>
Individuelle Bezugsnormorientierung												
Gymnasium												
Kompetenz	-.07	.06	-.02	.43	-.08	.06	.00	.93	<b>.08</b>	.03	.00	.92
Noten	.08	.09	-.02	.39	.01	.86	.00	.90	.01	.90	.04	.17
Realschule												
Kompetenz	.04	.40	-.01	.65	-.03	.57	.02	.52	.01	.78	-.01	.79
Noten	.04	.38	.03	.39	.07	.11	.00	.95	.09	.08	-.03	.32
Soziale Bezugsnormorientierung												
Gymnasium												
Kompetenz	.00	.90	.00	.90	-.01	.79	.01	.66	-.03	.49	.03	.34
Noten	-.06	.22	.01	.61	-.07	.14	-.01	.62	.05	.43	.00	.99
Realschule												
Kompetenz	.01	.91	.02	.58	.00	.93	.02	.40	-.10	.07	.05	.16
Noten	.01	.82	-.02	.62	.00	.95	.01	.85	.02	.70	.00	.90

*Allgemeine Anmerkungen.* KBA = kleine Beurteilungsaufgabe;  $\gamma$  = HLM-Regressionskoeffizient. Auf Schülerebene wurde in allen Modellen für Geschlecht, Alter, Migrationsstatus, sozioökonomischer Status (höchster ISEI), Lesekompetenz (PISA), kognitive Fähigkeiten (KFT), Ausbildung der Eltern und das jeweilige mathematische Schülermerkmal in der 9. Klasse (gruppenzentriert) kontrolliert.

$N_{\text{Schüler}}(\text{Gymnasium}) = 1153$ ;  $N_{\text{Klassen}}(\text{Gymnasium}) = 50$ ;  $N_{\text{Schüler}}(\text{Realschule}) = 1312$ ;  $N_{\text{Klassen}}(\text{Realschule}) = 56$ . HLM-Regressionskoeffizient mit *p*-Werten  $\leq .05$  sind fett gedruckt.

## 5. Diskussion

Um mehr über die Generalisierbarkeit der Ergebnisse der bisherigen BNO-Forschung zu lernen, untersuchten wir im vorliegenden Beitrag (a) wie verschiedene Erhebungsmethoden von BNO konvergieren und (b) wie sich die IBNO und die SBNO von Mathematiklehrkräften am Ende der Sekundarstufe I auf die Entwicklung motivational-affektiver Aspekte der Lernmotivation und der Leistung auswirken. Die zentralen Ergebnisse lassen sich wie folgt zusammenfassen.

Erstens, die geringen Korrelationen zwischen den verschiedenen BNO-Erhebungsmethoden (siehe Tabelle 1) weisen im Einklang mit den wenigen einschlägigen Vorläuferstudien (Rheinberg, 1980; siehe auch Dickhäuser & Rheinberg, 2003) daraufhin, dass die verschiedenen Instrumente jeweils zu großen Teilen einen unique Aspekt der BNO abbilden, der teilweise instrumentenspezifisch ist und zum anderen eventuell durch unterschiedliche Erhebungszeitpunkte bzw. Berechnungsverfahren erklärt werden kann (FEBO und SPLB in K9, KBA in K10; siehe unten). Vor dem Hintergrund der aktuellen multimethodalen Unterrichtsforschung ist die Instrumentenspezifität jedoch nicht völlig überraschend (Clausen, 2002; Kunter & Baumert, 2006). Einige Studien zeigten hier, dass die Schülersicht, die traditionell in der Unterrichtsklimaforschung erfasst wird (siehe Gruehn, 2000), vor allem valide Indikatoren für Merkmale mit affektiver und sozialer Bedeutsamkeit liefert sowie für Merkmale, die Routinen des Unterrichts betreffen (Clausen, 2002; Gruehn, 2000; Kunter & Baumert, 2006). Die Perspektive der Lehrkräfte ist vor allem dann erforderlich, wenn die dem unterrichtlichen Handeln zugrunde liegenden Entscheidungskriterien und Leitvorstellungen erfasst werden sollen, da diese sich weitestgehend externer Beobachtung entziehen. Angesichts dieser Befunde stellt sich die Frage, ob die BNO *von Lehrkräften* durch *Schüler* überhaupt valide eingeschätzt werden kann. Eine geringe Konvergenz zwischen BNO aus Schüler- und Lehrersicht wäre daher zu erwarten. Insgesamt deuten die vorliegenden Ergebnisse daraufhin, dass angesichts der Spezifität der einschlägigen Erhebungsinstrumente der BNO, Effekte der IBNO, bzw. SBNO an bestimmte Erfassungsmethoden der IBNO, bzw. SBNO gebunden sind und Befunde früherer BNO-Forschung nicht ohne weiteres über verschiedene Instrumente hinweg generalisierbar sind.

Zweitens, die Effekte der IBNO wie auch der SBNO auf die Entwicklung der Lernmotivation und der Leistung waren meist klein und meist nicht statistisch signifikant von Null verschieden (siehe Tabelle 3). Dieses Ergebnis steht dabei im Einklang mit früheren Studien (Köller, 2004; Lüdtke et al., 2005), in denen ebenfalls Mehrebenenmodelle eingesetzt wurden, und in denen zwar statistisch signifikante aber dennoch nur kleine Effekte der IBNO auf die Entwicklung mathematischer Selbstkonzepte gefunden wurden. Weiterhin stützten unsere Ergebnisse nicht die Interaktionshypothesen: Die Effekte von BNO auf die mathematische Kompetenz und die Mathematiknoten waren bei leistungsschwächeren Schülern generell *nicht* stärker ausgeprägt, als bei leistungsstärkeren Schülern (siehe

Tabelle 4). Schließlich ließen sich Effekte der BNO kaum über Erhebungsmethoden oder über Schulformen hinweg generalisieren.

Eine substanztheoretische Erklärung für dieses Befundmuster liefert die Selbstkonzeptforschung (z.B. Möller, Pohlmann, Köller & Marsh, 2009). Im Rahmen dieser Forschung wurde wiederholt empirisch gezeigt, dass der Vergleich der eigenen Leistungen mit den Leistungen der Peergruppe eine grundlegende Basis für die Genese schulischer Selbstkonzepte darstellt. Dieser Befund hat wichtige Implikationen für die potentielle Wirkung von BNO: Auch wenn eine Lehrkraft ganz bewusst nach einer IBNO oder SBNO beurteilt und so möglicherweise die Aufmerksamkeit der Schüler auf Leistungsveränderungen (oder Leistungsstatus) lenkt, so ziehen Schüler unabhängig von der BNO ihrer Lehrkraft verschiedene Informationsquellen heran, um ihre Leistung zu bewerten (Skaalvik & Skaalvik, 2002). So zeigten auch die Ergebnisse von Lüdtke und Kollegen (2005), dass – entgegen den theoretischen Erwartungen – keine Interaktion zwischen der IBNO und sozialen Vergleichsprozessen bei der Genese von mathematischen Selbstkonzepten besteht. Eine stark ausgeprägte IBNO einer Lehrkraft kompensiert in dieser Studie nicht soziale Vergleichsprozesse seitens der Schüler. BNO-Effekte der Lehrkräfte können folglich durch soziale Vergleichsprozesse seitens der Schüler überlagert werden.

Zusammenfassend legen die Ergebnisse dieser Studie für das Fach Mathematik in der Sekundarstufe I nahe, dass die Befunde früher BNO-Studien nicht ohne weiteres generalisiert werden können. Trotz vieler methodischer Vorzüge hat die vorliegende Studie Grenzen, die gleichzeitig Impulse für die zukünftige Forschung zum Konzept der BNO geben können.

Erstens, wir fokussierten unsere Analysen auf Schülermerkmale für die auch in früheren BNO-Studien positive IBNO-Effekte gefunden wurden (siehe Mischo & Rheinberg, 1995). Wir nahmen an, dass die IBNO einen positiven Effekt auf Erfolgserwartungen und Kausalattributionen hat, die den Effekt der IBNO auf die Schülermerkmale mediieren. Angaben zu Erfolgserwartungen und Kausalattributionen wurden im Rahmen von PISA nicht erhoben, wodurch Effekte der BNO auf diese Variablen nicht untersucht werden konnten. Um ein vollständigeres Bild zur Wirkweise von BNO zu gewinnen, sollten daher auch die Mediationsvariablen erfasst werden.

Zweitens, die vorliegenden Ergebnisse basieren auf einer großen Längsschnittstichprobe, die die Entwicklung von Jugendlichen von der 9. auf die 10. Klassestufe nachzeichnet. Nur wenige Studien konnten bislang Effekte von BNO auf motivational-affektive Schülermerkmale oder die Lernleistung am Ende der Sekundarstufe I nachweisen (siehe Krampen, 1985; Lüdtke et al., 2005; Rheinberg & Hendricks, 1980; Rheinberg et al., 1979). Vermutlich sind Effekte in jüngeren Jahrgangsstufen größer, da schulische Kompetenzen hier noch größeren Veränderungen unterliegen (z.B. Hill et al., 2008). Ebenso ist davon auszugehen, dass in der 9. und 10. Klassenstufe motivationale Merkmale weniger veränderbar sind und weitestgehend stabil in die Persönlichkeitsstruktur integriert sind. Letztere ist in niedrigeren Klassenstufen noch deutlich stärker im Fluss (z. B. Caspi & Roberts, 2001). Ob

das Alter die Effekte von BNO moderiert, ist daher eine wichtige Frage für weitere Forschung zu diesem Konzept.

Drittens, die BNO von Lehrkräften dürfte die schulische Entwicklung dann besonders stark beeinflussen, wenn ein Wechsel der Lehrkraft stattfindet, dies eine positive Veränderung der BNO mit sich bringt und die Effekte auf die schulische Entwicklung kurz nach dem Wechsel erfasst werden. In der vorliegenden Studie unterrichtete dieselbe Lehrkraft in K9 und K10 und die Effekte wurden über ein Jahr erfasst. Somit informieren die vorliegenden Befunde vor allem über die langfristige Wirkung der BNO. Es ist dabei anzunehmen, dass die Effekte mit größeren Zeitintervallen geringer ausfallen als bei kürzeren Zeitintervallen. Eine wichtige Frage für die BNO-Forschung ist daher wie man die ggf. positive Wirkung von IBNO stabilisieren kann, damit sie nachhaltig die Entwicklung von Schülerinnen und Schülern positiv beeinflusst. So zeigten auch beispielsweise einige Interventionsstudien Anfangsschwierigkeiten bei einer Umstellung der BNO einer Lehrkraft, die nicht theoriekonform waren: Insbesondere Schüler am oberen und unteren Ende des Leistungsspektrums waren anfänglich irritiert beim Wechsel von SBNO zu IBNO (Krug & Leczybyl, 1999a, 1999b). Die BNO der Leistungsbewertung direkt zum Unterrichtsthema zu machen, wurde in diesem Zusammenhang als erfolgreiche Präventivmaßnahme diskutiert (siehe Krug & Leczybyl, 1999a). Eine regelmäßige offene Diskussion im Klassenverband über die BNO der Leistungsbewertung – sowie über Evaluation insgesamt – könnte demnach, ganz im Sinne einer Kultur authentischer Evaluation (siehe Kempfert & Rolff, 2005), zu einer langfristigen Stabilisierung von IBNO-Effekten führen.

Viertens, ein wichtiger Aspekt für zukünftige Forschung ist die Integration der BNO *von Schülern* (Skaalvik & Skaalvik, 2002). Schüler werden in verschiedenen Jahrgangsstufen und in unterschiedlichen Fächern von verschiedenen Lehrkräften (mit verschiedener BNO) unterrichtet. Entwicklungsförderlich für Schüler dürfte letztlich aber sein, ob *sie selbst* eine Verbesserung ihrer Ergebnisse als gute Leistung sehen. Dieser Aspekt sollte, zumindest als Mediatorvariable zwischen Lehrer-BNO und Schülerverhalten, zukünftig stärker berücksichtigt werden.

Fünftens, im Einklang mit den aktuellen Konzeptionen der Unterrichtsforschung betrachteten wir die BNO einer Lehrkraft als ein Merkmal der Lernumwelt. Daher sollte sich die Wirkung der BNO auf Klassenebene entfalten. Die Werte der Intraklassenkorrelationen zeigten, dass – wie auch in anderen repräsentativen Studien (Martin, Bobis, Anderson, Way & Vellar, 2011) – ein Großteil der Variabilität von motivational-affektiven Merkmalen auf Schülerebene angesiedelt ist und nicht etwa auf der Klassenebene. Aufgrund der hohen methodischen Qualität der Stichprobenziehung und Instrumentenentwicklung in PISA, bzw. COACTIV, betrachten wir die beobachtete Variation der motivational-affektiven und leistungsbezogenen Schülermerkmale auf Klassenebene nicht als einen Hinweis auf die mangelnde Qualität der Daten oder der psychometrischen Qualität der verwendeten Maße. So waren beispielsweise die internen Konsistenzen der verwendeten Maße (wie auch in früheren Studien zu BNO) als zufriedenstellend bis gut zu bewerten. Vor diesem Hintergrund ist die geringe klassenspezifische

sche Variabilität mathematikspezifischer Schülermerkmale ein wichtiger Befund den die zukünftige BNO-Forschung stärker in den Blick nehmen sollte, denn diese Variabilität markiert die Obergrenze für maximal aufzufindende BNO-Effekte im Schulkontext am Ende der Sekundarstufe I.

Sechstens, die strukturelle Konzeption von BNO wurde noch nicht endgültig geklärt (siehe Abschnitt 2.3, Punkt 3 sowie Lüdtke et al., 2005). In der vorliegenden Studie haben wir die IBNO und die SBNO getrennt voneinander untersucht. Zukünftige Studien sollten Interaktionen zwischen den beiden Dimensionen in den Blick nehmen, um so potentiell positive Effekte einer *Bezugsnormvielfalt* (Rheinberg, 1980) zu analysieren. Die Frage nach der strukturellen Konzeption von BNO ist auch zwangsläufig an die Frage nach der Erfassung von BNO gekoppelt. Nicht alle BNO-Instrumente sind gleichermaßen gut für die von uns gewählte, jeweils eindimensionale Modellierung von IBNO und SBNO geeignet. Während sowohl der FEBO, als auch die SPLB eine unabhängige Messung von IBNO und SBNO erlauben, ist dies bei der KBA nicht möglich (siehe Rheinberg, 1980). Insbesondere geht bei der KBA mit einer stärkeren Befürwortung der IBNO zwangsläufig auch eine geringere Befürwortung der SBNO einher und umgekehrt. Mangelnde Übereinstimmung zwischen, durch mit KBA ermittelter BNO einerseits, und durch Lehrkräfte und Schüler berichteter BNO andererseits, könnte auch teilweise durch diese Inkongruenz der Berechnungsverfahren erklärt werden.

Siebtens, die Erfassung der BNO erfolgte nicht für alle Maße zum gleichen Erhebungszeitpunkt. Im Gegensatz zu der BNO aus Sicht der Lehrkräfte und Schüler, welche beide in K9 erhoben wurden, kam die KBA erst in K10 zum Einsatz. Es stellt sich demnach die Frage, inwiefern die von uns diskutierte Instrumentenspezifität der unterschiedlichen BNO-Erfassungsmethoden (FEBO vs. KBA bzw. SPLB vs. KBA) nicht auch (zumindest teilweise) unterschiedlichen Erhebungszeitpunkten und damit verbundenen differenziellen Veränderungen der BNO der Lehrkräfte geschuldet war. Die in Tabelle 1 berichteten Korrelationen repräsentieren somit eher die Untergrenze der Zusammenhänge zwischen der BNO, die mit FEBO bzw. SPLB einerseits und KBA andererseits ermittelt wurde. Darüber hinaus ist allerdings zu bedenken, dass auch die Effekte der BNO erfasst über die KBA meist klein und meist nicht statistisch signifikant von Null verschieden waren, obwohl die KBA (im Vergleich zu FEBO bzw. SPLB) nahezu zeitgleich mit den mathematikspezifischen Ergebniskriterien in K10 erfasst wurde.

Achtens, die bisherige BNO-Forschung fokussierte auf die IBNO und SBNO, untersuchte jedoch kaum die kriteriale Bezugsnorm, bei der eine Leistung mit sachlichen Anforderungen verglichen wird (Heckhausen, 1974). Unseres Wissens liegen bislang keine empirisch belastbaren Befunde zur *kriterialen Bezugsnormorientierung* (KBNO) vor. Eine KBNO kommuniziert den Schülern, welche Standards zu erreichen sind und was eine gute Leistung auszeichnet. Da die Standards festgesetzt sind und somit Leistungsbeurteilungen auch gerechter und objektiver ausfallen als bei einer herkömmlichen Notengebung (Jürgens & Sacher, 2008), könnte dies ein Gefühl von Sicherheit auf Seiten der Schüler bewirken, das

sich wiederum positiv auf die Motivations- und Leistungsentwicklung auswirken könnte.

Angesichts der weitreichenden pädagogischen Implikationen (z.B. die Einbindung des BNO-Konzepts in der Lehrerbildung), die auf der BNO-Forschung basieren, hoffen wir, dass die vorliegenden Befunde zukünftige Forschung stimulieren kann, die die Grenzen unserer Studie thematisiert und dabei gleichzeitig die methodischen Stärken dieser Arbeit integriert.

## Danksagung

COACTIV wurde mit Mitteln der Deutschen Forschungsgesellschaft (DFG BA 1461/2-2) gefördert und ist Teil des DFG Schwerpunktprogramms Bildungsqualität von Schule (BIQUA). COACTIV ist ein Kooperationsprojekt des Max-Planck-Instituts für Bildungsforschung Berlin, der Universität Kassel und der Carl-von-Ossietzky-Universität Oldenburg. Die Autoren bedanken sich ganz herzlich bei Prof. Dr. Mareike Kunter (Goethe-Universität Frankfurt am Main) und Prof. Dr. Franzis Preckel (Universität Trier) für ihre wertvollen Hinweise zu einer früheren Version des Manuskripts.

## Literatur

- Ames, C. (1992). Classrooms: Goals, structures, and student motivation. *Journal of Educational Psychology*, 84(3), 261–271. <http://doi.org/10.1037/0022-0663.84.3.261>
- Antonakis, J. & Dietz, J. (2011a). Looking for validity or testing it? The perils of stepwise regression, extreme-scores analysis, heteroscedasticity, and measurement error. *Personality and Individual Differences*, 50(3), 409–415. <http://doi.org/10.1016/j.paid.2010.09.014>
- Antonakis, J. & Dietz, J. (2011b). More on testing for validity instead of looking for it. *Personality and Individual Differences*, 50(3), 418–421. <http://doi.org/10.1016/j.paid.2010.10.008>
- Baumert, J., Blum, W., Neubrand, M., Klusmann, U., Brunner, M., Jordan, A., Krauss, S., Kunter, M. & Löwen, K. (2006). *Professionswissen von Lehrkräften, kognitiv aktivierender Mathematikunterricht und die Entwicklung von mathematischer Kompetenz (COACTIV). Dokumentation der Erhebungsinstrumente*. Berlin: Max-Planck-Institut für Bildungsforschung.
- Baumert, J., Kunter, M., Blum, W., Brunner, M., Voss, T., Jordan, A., Klusmann, U., Krauss, S., Neubrand, M. & Tsai, Y.-M. (2010). Teachers' mathematical knowledge, cognitive activation in the classroom, and student progress. *American Educational Research Journal*, 47(1), 133–180. <http://doi.org/10.3102/0002831209345157>
- Baumert, J., Stanat, P. & Demmrich, A. (2001). PISA 2000: Untersuchungsgegenstand, theoretische Grundlagen und Durchführung der Studie. In J. Baumert, E. Klieme, M. Neubrand, U. Prenzel, U. Schiefele, W. Schneider, P. Stanat, K.-J. Tillmann, & M. Weiß (Hrsg.), *PISA 2000: Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich* (S. 15–68). Opladen: Leske & Budrich.
- Baumert, J., Trautwein, U. & Artelt, C. (2003). Schulumwelten. Institutionelle Bedingungen des Lehrens und Lernens. In Deutsches PISA-Konsortium (Hrsg.), *PISA 2000. Ein differenzierter Blick auf die Länder der Bundesrepublik Deutschland* (S. 261–331). Opladen: Leske & Budrich.

- Becker, T. E. (2005). Potential problems in the statistical control of variables in organizational research: A qualitative analysis with recommendations. *Organizational Research Methods*, 8(3), 274–289. <http://doi.org/10.1177/1094428105278021>
- Borkowski, J. G. & Thorpe, P. K. (1994). Self-regulation and motivation: A life-span perspective on underachievement. In D. H. Schunk & B. J. Zimmerman (Hrsg.), *Self-regulation of learning and performance: Issues and educational applications* (S. 45–73). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Brunner, M., Keller, U., Dierendonck, C., Reichert, M., Ugen, S., Fischbach, A. & Martin, R. (2010). The structure of academic self-concepts revisited: The nested Marsh/Shavelson model. *Journal of Educational Psychology*, 102(4), 964–981. <http://doi.org/10.1037/a0019644>
- Brunner, M., Lüdtke, O. & Trautwein, U. (2008). The internal/external frame of reference model revisited: Incorporating general cognitive ability and general academic self-concept. *Multivariate Behavioral Research*, 43(1), 137–172. <http://doi.org/10.1080/00273170701836737>
- Carroll, J. B. (1963). A model of school learning. *Teacher College Record*, 64(8), 723–733.
- Carroll, J. B. (1989). The Carroll model. A 25-year retrospective and prospective view. *Educational Researcher*, 18(1), 26–31.
- Carroll, J. B. (1996). Mathematical abilities: Some results from factor analysis. In R. J. Sternberg & T. Ben-Zeev (Hrsg.), *The nature of mathematical thinking* (S. 3–25). Mahwah, NJ: Lawrence Erlbaum Associates.
- Caspi, A. & Roberts, B. W. (2001). Personality development across the life course: The argument for change and continuity. *Psychological Inquiry*, 12(2), 49–66.
- Clausen, M. (2002). *Qualität von Unterricht – Eine Frage der Perspektive?* Münster: Waxmann.
- Dickhäuser, O. & Rheinberg, F. (2003). Bezugsnormorientierung: Erfassung, Probleme, Perspektiven. In J. Stiensmeier-Pelster & F. Rheinberg (Hrsg.), *Diagnostik von Motivation und Selbstkonzept* (S. 41–55). Göttingen: Hogrefe.
- Enders, C. K. & Tofighi, D. (2007). Centering predictor variables in cross-sectional multilevel models: A new look at an old issue. *Psychological Methods*, 12(2), 121–138. <http://doi.org/10.1037/1082-989X.12.2.121>
- Ganzeboom, H. B. G., De Graaf, P. M. & Treiman, D. J. (1992). A standard international socio-economic index of occupational status. *Social Science Research*, 21(1), 1–56.
- Gruehn, S. (2000). *Unterricht und schulisches Lernen. Schüler als Quellen der Unterrichtsbeschreibung*. Münster: Waxmann.
- Haertel, G. D., Walberg, H. J. & Weinstein, T. (1983). Psychological models of educational performance: A theoretical synthesis of constructs. *Review of Educational Research*, 53(1), 75–91.
- Heckhausen, H. (1974). *Leistung und Chancengleichheit*. Göttingen: Hogrefe.
- Heller, K. & Perleth, C. (2000). *Kognitiver Fähigkeitstest für 4. bis 12. Klassen, Revision (KFT 4–12 + R)*. Göttingen: Hogrefe.
- Helmke, A. & Schrader, F.-W. (2010). Determinanten der Schulleistung. In D. H. Rost (Hrsg.), *Handwörterbuch Pädagogische Psychologie* (4. Aufl., S. 90–102). Weinheim: Beltz.
- 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. <http://doi.org/10.1037/0022-0663.87.4.624>
- Helmke, A. & Weinert, F. E. (1997). Bedingungsfaktoren schulischer Leistungen. In F. E. Weinert (Hrsg.), *Psychologie der Schule und des Unterrichts* (Enzyklopädie der Psychologie, Serie Pädagogische Psychologie, Bd. 3, S. 71–176). Göttingen: Hogrefe.

- Henry, J. W., Martinko, M. J. & Pierce, M. A. (1993). Attributional style as a predictor of success in a first computer science course. *Computers in Human Behavior*, 9(4), 341–352. [http://doi.org/10.1016/0747-5632\(93\)90027-P](http://doi.org/10.1016/0747-5632(93)90027-P)
- Hill, C. J., Bloom, H. S., Black, A. & Lipsey, M. W. (2008). Empirical benchmarks for interpreting effect sizes in research. *Child Development Perspectives*, 2(3), 172–177.
- Hosenfeld, I. (2002). *Kausalitätsüberzeugungen und Schulleistungen*. Münster: Waxmann.
- Jürgens, E. & Sacher, W. (2008). *Leistungserziehung und Pädagogische Diagnostik in der Schule*. Stuttgart: Kohlhammer.
- Kempfert, G. & Rolff, H.-G. (2005). *Qualität und Evaluation. Ein Leitfaden für Pädagogisches Qualitätsmanagement*. Weinheim: Beltz.
- Klieme, E., Neubrand, M. & Lüdtke, O. (2001). Mathematische Grundbildung: Testkonzeption und Ergebnisse. In J. Baumert, E. Klieme, M. Neubrand, U. Prenzel, U. Schiefele, W. Schneider, P. Stanat, K.-J. Tillmann & M. Weiß (Hrsg.), *PISA 2000: Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich* (S. 139–190). Opladen: Leske & Budrich.
- Köller, O. (2004). *Konsequenzen von Leistungsgruppierungen*. Münster: Waxmann.
- Krampen, G. (1985). Differentielle Effekte von Lehrerkommentaren zu Noten bei Schülern. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 17(2), 99–123.
- Krug, S. & Lecybyl, R. (1999a). Die Veränderung von Einstellung, Mitarbeit und Lernleistung im Verlauf einer Bezugsnormspezifischen Motivationsintervention. In F. Rheinberg & S. Krug (Hrsg.), *Motivationsförderung im Schulalltag* (2. Aufl., S. 95–114). Göttingen: Hogrefe.
- Krug, S. & Lecybyl, R. (1999b). Die Wirkung experimentell variierten Lehrerverhaltens auf Unterrichts-wahrnehmung, Lernbereitschaft und Leistung von Schülern. In F. Rheinberg & S. Krug (Hrsg.), *Motivationsförderung im Schulalltag* (2. Aufl., S. 81–94). Göttingen: Hogrefe.
- Kunter, M. & Baumert, J. (2006). Who is the expert? Construct and criteria validity of student and teacher ratings of instruction. *Learning Environments Research*, 9(3), 231–251. <http://doi.org/10.1007/s10984-006-9015-7>
- Kunter, M., Baumert, J., Blum, W., Klusmann, U., Krauss, S. & Neubrand, M. (Hrsg.). (2011). *Professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV*. Münster: Waxmann.
- Lüdtke, O., Köller, O., Marsh, H. W. & Trautwein, U. (2005). Teacher frame of reference and the big-fish–little-pond effect. *Contemporary Educational Psychology*, 30(3), 263–285. <http://doi.org/10.1016/j.cedpsych.2004.10.002>
- Lüdtke, O., Trautwein, U., Kunter, M. & Baumert, J. (2006). Analyse von Lernumwelten. Ansätze zur Bestimmung der Reliabilität und Übereinstimmung von Schülerwahrnehmungen. *Zeitschrift für Pädagogische Psychologie*, 20(1), 85–96.
- MacCallum, R. C., Zhang, S., Preacher, K. J. & Rucker, D. D. (2002). On the practice of dichotomization of quantitative variables. *Psychological Methods*, 7(1), 19–40. <http://doi.org/10.1037//1082-989X.7.1.19>
- Marsh, H. W. (1987). The big-fish–little-pond effect on academic self-concept. *Journal of Educational Psychology*, 79(3), 280–295. <http://doi.org/10.1037/0022-0663.79.3.280>
- 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. <http://doi.org/10.1111/j.1745-6916.2006.00010.x>

- Marsh, H. W., Trautwein, U., Lüdtke, O., Köller, O. & Baumert, J. (2005). Academic self-concept, interest, grades, and standardized test scores: Reciprocal effects models of causal ordering. *Child Development*, 76(2), 397–416. <http://doi.org/10.1111/j.1467-8624.2005.00853.x>
- Martin, A. J., Bobis, J., Anderson, J., Way, J. & Vellar, R. (2011). Patterns of multi-level variance in psycho-educational phenomena: Comparing motivation, engagement, climate, teaching, and achievement factors. *Zeitschrift Für Pädagogische Psychologie*, 25(1), 49–61. <http://doi.org/10.1024/1010-0652/a000029>
- Mayer, R. E. & Hegarty, M. (1996). The process of understanding mathematical problems. In R. J. Sternberg & T. Ben-Zeev (Hrsg.), *The nature of mathematical thinking* (S. 29–53). Mahwah, NJ: Lawrence Erlbaum Associates.
- Meece, J. L. (2002). *Child and adolescent development for educators* (2. Aufl.). New York, NY: McGraw Hill.
- Mischo, C. & Rheinberg, F. (1995). Erziehungsziele von Lehrern und individuelle Bezugsnormen der Leistungsbewertung. *Zeitschrift für Pädagogische Psychologie*, 9(3), 139–151.
- Möller, J., Pohlmann, B., Köller, O. & Marsh, H. W. (2009). A meta-analytic path analysis of the internal/external frame of reference model of academic achievement and academic self-concept. *Review of Educational Research*, 79(3), 1129–1167. <http://doi.org/10.3102/0034654309337522>
- Preacher, K. J., Rucker, D. D., MacCallum, R. C. & Nicewander, W. A. (2005). Use of the extreme groups approach: A critical reexamination and new recommendations. *Psychological Methods*, 10(2), 178–192. <http://doi.org/10.1037/1082-989X.10.2.178>
- Ramm, G., Prenzel, M., Baumert, J., Blum, W., Lehmann, R., Leutner, D., Neubrand, M., Pekrun, R., Rolff, H.-G., Rost, J. & Schiefele, U. (Hrsg.). (2006). *PISA 2003. Dokumentation der Erhebungsinstrumente*. Münster: Waxmann.
- Raudenbush, S. W. & Bryk, A. S. (2002). *Hierarchical linear models. Applications and data analysis methods* (2. Aufl.). Thousand Oaks, CA: Sage.
- Rheinberg, F. (1980). *Leistungsbewertung und Lernmotivation*. Göttingen: Hogrefe.
- Rheinberg, F. (2006). Bezugsnormorientierung. In D. H. Rost (Hrsg.), *Handwörterbuch Pädagogische Psychologie* (3. Aufl., S. 55–62). Weinheim: Beltz.
- Rheinberg, F. (2008). Bezugsnormen und die Beurteilung von Lernleistung. In W. Schneider & M. Hasselhorn (Hrsg.), *Handbuch der Pädagogischen Psychologie* (S. 178–186). Göttingen: Hogrefe.
- Rheinberg, F. & Hendricks, U. (1980). Verbesserte Wahrnehmung eigenen Lernzuwachses. Anlage und Ergebnisse eines Unterrichtsexperiments. *Zeitschrift für Pädagogik*, 26(6), 895–906.
- Rheinberg, F., Kühmel, B. & Duscha, R. (1979). Experimentell variierte Schulleistungsbewertung und ihre motivationalen Folgen. *Zeitschrift für Empirische Pädagogik*, 3, 1–12.
- Rheinberg, F. & Peter, R. (1982). Selbstkonzept, Ängstlichkeit und Schulunlust von Schülern: Eine Längsschnittstudie zum Einfluß des Klassenlehrers. In F. Rheinberg (Hrsg.), *Jahrbuch für Empirische Erziehungswissenschaft. Schulleistungsbewertung* (S. 143–159). Düsseldorf: Schwann.
- Schafer, J. L. & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods*, 7(2), 147–177. <http://doi.org/10.1037//1082-989X.7.2.147>
- Schwarzer, R. & Jerusalem, M. (2002). Das Konzept der Selbstwirksamkeit. In M. Jerusalem & D. Hopf (Hrsg.), *Selbstwirksamkeit und Motivationsprozesse in Bildungsinstitutionen* (S. 28–53). Weinheim: Beltz.
- Schwarzer, R., Lange, B. & Jerusalem, M. (1982). Die Bezugsnorm des Lehrers aus der Sicht des Schülers. In F. Rheinberg (Hrsg.), *Jahrbuch für Empirische Erziehungswissenschaft. Schulleistungsbewertung* (S. 161–172). Düsseldorf: Schwann.

- Skaalvik, E. M. (1994). Attribution of perceived achievement in school in general and in maths and verbal areas: Relations with academic self-concept and self-esteem. *British Journal of Educational Psychology*, 64(1), 133–143. <http://doi.org/10.1111/j.2044-8279.1994.tb01090.x>
- Skaalvik, E. M. & Skaalvik, S. (2002). Internal and external frames of reference for academic self-concept. *Educational Psychologist*, 37(4), 233–244.
- Snijders, T. A. B. & Bosker, R. J. (2011). *Multilevel analysis. An introduction to basic and advanced multilevel modeling* (2. Aufl.). London, United Kingdom: Sage.
- Spector, P. E. & Brannick, M. T. (2010). Methodological urban legends: The misuse of statistical control variables. *Organizational Research Methods*, 14(2), 287–305. <http://doi.org/10.1177/1094428110369842>
- Stiensmeier-Pelster, J. & Schwinger, M. (2008). Kausalattribution. In W. Schneider & M. Hasselhorn (Hrsg.), *Handbuch der Pädagogischen Psychologie* (S. 74–83). Göttingen: Hogrefe.
- Walberg, H. J. (1986). Syntheses of research on teaching. In M. C. Wittrock (Hrsg.), *Handbook of research on teaching* (3. Aufl., S. 214–229). New York, NY: Macmillan.
- Wang, M. C., Haertel, G. D. & Walberg, H. J. (1993). Toward a knowledge base for school learning. *Review of Educational Research*, 63(3), 249–294.

## UNSERE BUCHEMPFEHLUNG



Wilfried Bos, Ramona Lorenz,  
Manuela Endberg, Heike Schaumburg,  
Renate Schulz-Zander, Martin Senkbeil  
(Hrsg.)

### Schule digital – der Länderindikator 2015

Vertiefende Analysen zur  
schulischen Nutzung  
digitaler Medien im  
Bundesländervergleich

2015, 238 Seiten, br., 29,90 €,  
ISBN 978-3-8309-3400-4

E-Book: 26,99 €,  
ISBN 978-3-8309-8400-9

Mit dem Länderindikator 2015 wird erstmalig eine bundeslandspezifische Betrachtung der schulischen Nutzung digitaler Medien in Deutschland realisiert. Auf Grundlage einer repräsentativen Befragung von Lehrkräften werden Analysen zum kompetenten Umgang mit digitalen Medien vorgelegt, die aktuelle Befunde und Handlungsbedarfe aufgreifen.

In diesem Band werden die Aspekte der schulischen Ausstattung mit digitalen Medien, ihrer Nutzung in Lehr- und Lernkontexten sowie der IT-bezogenen Einstellungen der Lehrkräfte und Fähigkeiten der Schülerinnen und Schüler behandelt. Die Ergebnisse zeigen, dass die Integration digitaler Medien in der Schule in einigen Bundesländern recht weit vorangeschritten ist, in anderen Ländern aber noch Nachholbedarf besteht.

