

Hövel, Dennis Christian; Hennemann, Thomas; Urban, Mareike; Osipov, Igor
**"Die Mischung macht den Unterschied". Selektive Wirksamkeitsstudie zur
Förderung emotionaler und sozialer Kompetenzen mit dem Ben & Lee
Programm**

*Emotionale und soziale Entwicklung in der Pädagogik der Erziehungshilfe und bei Verhaltensstörungen
: ESE 2 (2020) 2, S. 30-44*



Quellenangabe/ Reference:

Hövel, Dennis Christian; Hennemann, Thomas; Urban, Mareike; Osipov, Igor: "Die Mischung macht den Unterschied". Selektive Wirksamkeitsstudie zur Förderung emotionaler und sozialer Kompetenzen mit dem Ben & Lee Programm - In: Emotionale und soziale Entwicklung in der Pädagogik der Erziehungshilfe und bei Verhaltensstörungen : ESE 2 (2020) 2, S. 30-44 - URN: urn:nbn:de:0111-pedocs-251969 - DOI: 10.25656/01:25196

<https://nbn-resolving.org/urn:nbn:de:0111-pedocs-251969>

<https://doi.org/10.25656/01:25196>

in Kooperation mit / in cooperation with:



<http://www.klinkhardt.de>

Nutzungsbedingungen

Dieses Dokument steht unter folgender Creative Commons-Lizenz: <http://creativecommons.org/licenses/by-nc-nd/4.0/deed.de> - Sie dürfen das Werk bzw. den Inhalt unter folgenden Bedingungen vervielfältigen, verbreiten und öffentlich zugänglich machen: Sie müssen den Namen des Autors/Rechteinhabers in der von ihm festgelegten Weise nennen. Dieses Werk bzw. dieser Inhalt darf nicht für kommerzielle Zwecke verwendet werden und es darf nicht bearbeitet, abgewandelt oder in anderer Weise verändert werden.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use

This document is published under following Creative Commons-Licence: <http://creativecommons.org/licenses/by-nc-nd/4.0/deed.en> - You may copy, distribute and transmit, adapt or exhibit the work in the public as long as you attribute the work in the manner specified by the author or licensor. You are not allowed to make commercial use of the work or its contents. You are not allowed to alter, transform, or change this work in any other way.

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

**„Die Mischung macht den Unterschied“ –
Selektive Wirksamkeitsstudie zur Förderung
emotionaler und sozialer Kompetenzen mit dem
Ben & Lee Programm**

*Dennis Christian Hövel, Thomas Hennemann,
Mareike Urban und Igor Osipov*

Abstract

Hintergrund. Bei der Umsetzung präventiver Maßnahmen in der Schule bestehen einige Herausforderungen. Viele Grundschulen haben u.a. Probleme, die zeitlichen Ressourcen für Förderprogramme aufzubringen und die Maßnahmen mit dem Lehrplan in Einklang zu bringen. Ein Lösungsansatz für diese Herausforderung könnte die Integration Sozial-Emotionalen-Lernens (SEL) in das reguläre Curriculum sein, so wie es von Ben & Lee umgesetzt wird. Die Wirkung von SEL-Programmen fällt in der Primarstufe stark heterogen aus. Die Zielstellung der vorliegenden Studie ist daher einerseits, die Wirksamkeit von Ben & Lee (Urban, Hövel & Hennemann, 2018) zu überprüfen und andererseits, mögliche Einflussfaktoren auf die Effekte zu identifizieren.

Methodik. Ben & Lee wurde mittels quasi-experimentellem Kontrollgruppen-Design in 36 Klassen aus 3. und 4. Jahrgangsstufen umgesetzt. Das sozial-emotionale Wissen, das Verhalten und das Erleben von 405 Kindern mit auffälligem Verhalten im Lehrerurteil wurde vor, direkt nach und drei Monaten nach Beendigung der Förderung mittels standardisierter Verfahren (Test, Fremd- und Selbstbeurteilung) erhoben. Die Datenanalyse erfolgte mit Hilfe linearer Mehrebenenwachstumsmodelle.

Ergebnisse & Schlussfolgerungen. Kinder mit Schwierigkeiten im Verhalten zeigten nach dem Einsatz von Ben & Lee im Vergleich zur Kontrollgruppe ein signifikant erhöhtes sozial-emotionales Wissen ($d = 0.5$). Dieses führt jedoch nicht unmittelbar zu Veränderungen im Verhalten und im Erleben der Schülerinnen und Schüler. Für eine Verbesserungen in diesen Bereichen kommt es auf die Klassenzusammensetzung sowie das Feedbackverhalten der Lehrperson an. Der Beitrag diskutiert Handlungsschritte sowohl auf Ebene der Lehrpersonen als auch auf schulorganisatorischer Ebene.

Keywords

Prävention, Sozial-Emotionales-Lernen, Grundschule

1 Einleitung

Für eine präventive Förderung emotionaler und sozialer Kompetenzen in der Grundschule lässt sich, insbesondere für Kinder unter erhöhten psychosozialen Risiken, entlang internationaler (Wilson & Lipsey, 2007) sowie nationaler Metaanalysen (Beelmann, Pfof & Schmitt, 2014; Hövel, Hennemann & Rietz, 2019) eine gute Wirksamkeit nachweisen. In ihrer Langzeitstudie mit randomisiertem Experimental-Kontrollgruppendesign ($N = 1.634$) untersuchten McCormick et al. (2019) die Auswirkungen Sozial-Emotionalen-Lernen (SEL) über einen Zeitraum von fünf Jahren. Die Autorengruppe belegt einen signifikanten Effekt von SEL auf die Notwendigkeit einer Zuweisung sonderpädagogischen Förderbedarfs, d.h. die Schülerinnen und Schüler, die am SEL teilgenommen haben, zeigten eine geringere Quote an sonderpädagogischem Förderbedarf. Die aktuelle KiGGS-Studie (Klipker, Baumgarten, Göbel, Lampert & Hölling, 2018) diskutiert präventive Maßnahmen zudem als möglichen Einflussfaktor für den signifikanten Rückgang der Gesamtprävalenz von Verhaltensstörungen bei Kindern- und Jugendlichen von rund 20% auf ca. 17%.

Bei der Umsetzung präventiver Maßnahmen in der Schule gibt es jedoch noch einige Herausforderungen. Karing, Beelmann und Haase (2015) ermitteln, dass viele Grundschulen u.a. Probleme haben, die zeitlichen Ressourcen für Förderprogramme aufzubringen und die Zulassung solcher Programme aufgrund des Lehrplans schwierig sei. Ein Lösungsansatz für diese Herausforderung könnte die Integration SELs in das reguläre Curriculum sein (Reicher, 2010). Ein solches Programm, welches SEL mit den Inhalten des Faches Deutsch und Sachunterrichts verzahnt, ist Ben & Lee (Urban, Hövel & Hennemann, 2018).

2 Curricularer Aufbau von Ben & Lee

Das Unterrichtswerk ist entlang einer Rahmengeschichte konzipiert. Die beiden Kinder Ben und Lee geraten durch einen geheimnisvollen Kompass auf eine Zeitreise, auf der sie sich zuerst bei Indianern, dann bei Piraten und schließlich in der Zukunft wiederfinden. Auf ihren Reisen stoßen sie immer wieder auf Rätsel und Probleme im Umgang mit Gefühlen und in der sozialen Interaktion.

Das Programm ist in drei Bausteine gegliedert und umfasst insgesamt 24 curriculare Doppelstunden zu je 90 Minuten. Zusätzlich gibt es drei Bausteinabschlüsse, meist in Zusammenarbeit mit den Eltern, vier Vertiefungsstunden und zahlreiche Vertiefungsideen. Jeder Baustein verfolgt neben einem Entwicklungsanliegen auch konkrete Fachanliegen aus den Fächern Deutsch und Sachunterricht. Hierbei werden die curricularen Themenschwerpunkte ‚Lebensweisen und Gebräuche indigener Kulturen‘, ‚planen, schreiben und überarbeiten von Geschichten‘, ‚Merkmale einer gesunden Lebensweise‘, ‚Experimente‘ sowie ‚verantwortungsvoller Umgang mit Ressourcen/Energie‘ behandelt. Alle Anliegen entsprechen den Anforderungen der Lehrpläne für den Deutsch- und Sachunterricht 3. und 4. Jahrgangsstufen (u.a. des Lehrplans für Grundschulen in NRW).

Im Bereich der Entwicklungsanliegen sequenziert das Modell der sozial-kognitiven Informationsverarbeitung (SKI; Crick & Dodge, 1994; Lemerise & Arsenio, 2000) das SEL. Der erste Baustein adressiert die Fremd- und Selbstwahrnehmung sowie die Analyse komplexer sozialer Situationen. Im zweiten Baustein wird am Strategierepertoire zur Emotionsregulation gearbeitet und der dritte Baustein thematisiert das komplexe Feld des sozialen Problemlösens mit den Teilaspekten Zielantizipation sowie Generierung und Überprüfung von Hand-

lungsmöglichkeiten. Entlang des SKI-Modells werden alle SEL-Inhalte in eine strukturierte und visualisierte Problemlöseformel überführt, die den Transfer der erlernten Inhalte in den Alltag erleichtern soll.

3 Mögliche Einflussfaktoren auf die Wirksamkeit

Die Wirksamkeit von SEL-Programmen fällt in der Primarstufe stark heterogen aus (Hövel et al., 2019). Hinsichtlich der Zielgruppe lässt sich festhalten, dass indizierte und selektive Gruppen ($d = 0.39$ & 0.29) stärker profitieren als universelle ($d = 0.11$). Neben Aspekten der Umsetzungsmodalitäten (z.B. strukturierter Transfer der Inhalte in den Alltag, $r = .31$ oder Anzahl der Einheiten, $r = .33$) beeinflussen unter anderem auch die Erhebungsmethoden ($r = .22$) die Evaluationsergebnisse (Hövel et al., 2019, S. 50), was auch internationale Untersuchungen (u.a. Wilson & Lipsey, 2007) belegen. Die mittels Tests erhobenen Befunde fallen größer aus als die Ergebnisse aus Fremd- und Selbsteinschätzungen. Während die Tests das im Training vermittelte sozial-emotionale Wissen erfassen, bilden die standardisierten Einschätzungen das von den Schülerinnen und Schülern selbst und/oder das von ihren Bezugspersonen wahrgenommene Verhalten ab. Hier scheint es eine Diskrepanz zu geben. Mit einer durchschnittlichen Effektstärke von $d = 0.22$ erzielen SEL-Programme eine praktische bedeutsame Wirkung im Bereich der Wissensvermittlung. Die Überführung dieses Wissens in beobachtbares und erlebbares Verhalten fällt an vielen Stellen jedoch noch schwer, was die im Vergleich zum Wissen kleineren gemittelten Effektstärken von $d = 0.13 - 0.11$ für das Fremd- und Selbsturteil belegen (Hövel et al., 2019).

Bei kognitionspsychologisch ausgerichteten SEL-Maßnahmen handelt es sich um komplexe Interventionen für den Unterricht. Faktoren der Unterrichtsführung spielen hierbei eine wichtige Rolle. Die Manuale der Programme greifen in der Regel zwar explizit Kriterien einer effektiven Klassenführung wie Regeln, Konsequenzen, Umgang mit Störungen, kooperative Lernformen usw. auf, diese sind jedoch nicht isoliert zu betrachten und zu bearbeiten, sondern fügen sich in einem größeren Gesamtzusammenhang erfolgreicher Unterrichtsgestaltung in Abhängigkeit zu den Bedarfen der jeweiligen Lerngruppe ein. Entlang des Angebot-Nutzungsmodells von Helmke (2009) wird die Klassenführung einerseits durch den Klassenkontext und deren Zusammensetzung, sowie andererseits durch Merkmale der Lehrperson beeinflusst. Zu beiden Aspekten liegen insbesondere für den Verhaltensbereich Theorien vor, die diese Einflüsse beschreiben.

3.1 Klassenkomposition

In ihrem Leitartikel des Themenhefts zu den Zusammenhängen der Klassenkomposition und der sozial-emotionalen Entwicklung geben Müller und Zurbriggen (2016), ausgehend von der Theorie der differentiellen Kontakte von Sutherland, einen Überblick zur aktuellen Befundlage. Die vereinfacht dargestellte Grundthese der Theorie ist, dass die Wahrscheinlichkeit für ein Individuum, delinquentes Verhalten zu zeigen, steigt, je mehr Kontakt zu anderen delinquenten Personen besteht. Wichtige Prozesse sind hierbei die von Gleichaltrigen gebotenen Möglichkeiten, soziale Modellierungen, Verstärkungen und Vergleiche sowie emotionale Ansteckung (Müller & Zurbriggen, 2016). Müller und Zurbriggen referieren zudem Studien, die belegen, dass die Gruppe der Gleichaltrigen nicht nur Einfluss auf die Externalisierung, sondern auch auf die Internalisierung und das sozialkompetente Verhalten von Kindern und Jugendlichen hat (Barry & Wentzel, 2006; Prinstein, 2007).

Integriert man diese Befunde in die dargestellten Effekte von SEL-Programmen, so liegt die These nah, dass präventive Maßnahmen zwar das sozial-emotionale Wissen verbessern, sich dieses Wissen aber erst bei günstiger Klassenkompositionen im gezeigten Verhalten der Schülerinnen und Schüler manifestiert. D.h. Trainings können wahrscheinlich das Wissen für alle sechs Phasen der SKI vermitteln, um dieses Wissen zu habituierten werden aber eine entsprechende Bezugsnormorientierung in der Gruppe sowie positive Vorbildern benötigt.

3.2 Lehrkraftfeedback

Ein weiterer Mechanismus, der das gezeigte Verhalten innerhalb einer Klasse beeinflussen könnte, ist das Rückmeldeverhalten seitens der Lehrperson. Die Theorie der sozialen Referenzierung (Feinman, 1992) geht davon aus, dass sich Kinder für ihr Verhalten in unbekanntem Situationen an dem von Bezugspersonen orientieren. Beim Übertrag dieser Theorie auf die Schule wird davon ausgegangen (Huber, 2019), dass die Kinder einer Klasse das öffentliche Lehrkraftfeedback als sozialen Referenzrahmen für ihr Verhalten untereinander nutzen. In experimentellen Studien konnte, sowohl für die Grundschule allgemein (Huber, Gebhardt & Schwab, 2015), als auch für Kinder mit Behinderung (Huber, Gerullis, Gebhardt & Schwab, 2018), ein signifikanter Zusammenhang zwischen dem Lehrkraftfeedback und der sozialen Akzeptanz der Schülerinnen und Schüler ermittelt werden. Kinder, die von ihrer Lehrperson häufiger öffentlich eine positive Rückmeldung erhielten, waren besser sozial in die Klasse integriert. Demgegenüber waren Schülerinnen und Schüler, die öfter im Plenum ein negatives Feedback von der Lehrkraft erhielten, deutlich schlechter in die Gruppe der Gleichaltrigen eingegliedert.

Eine Implikation dieses Befunds auf die Wirksamkeit von SEL könnte sein, dass die Überführung des im Programm erworbenen Wissens auf das im Klassenzimmer gezeigte Verhalten auch durch das Feedbackverhalten der Lehrperson beeinflusst wird.

4 Fragestellung

Die Fragestellung der vorliegenden Studie ist daher, ob (1) das Ben & Lee Programm wirksam das sozial-emotionale Wissen, Verhalten und Erleben von Schülerinnen und Schülern mit Verhaltensproblemen der Primarstufe verbessern kann und ob (2) die Effekte zudem durch die Klassenkomposition und das Lehrkraftfeedback beeinflusst werden.

5 Methode

5.1 Studiendesign

Die Wirksamkeitsüberprüfung wurde in der vorliegenden Studie mittels quasi-experimentelles Design mit zwei Gruppen und drei Messzeitpunkten umgesetzt. Hierzu wurde eine Experimentalgruppe (EG) mit einer nicht mittels SEL geförderten Kontrollgruppe (KG) jeweils zwei Wochen vor (Prä-Test), zwei Wochen nach (Post-Test) und etwa drei Monate nach Durchführung des Programms (Follow-up) miteinander verglichen. An der Studie beteiligten sich insgesamt 35 Klassen an 26 Grundschulen mit zusammen 778 Schülerinnen und Schülern. Dabei fungierten 17 Klassen als Kontroll- und 19 Klassen als Versuchsgruppe, in der das Ben & Lee Programm jeweils mit der gesamten Klasse durchgeführt wurde. Bei der Stichprobenakquise wurden die Schulen zufällig auf beide Bedingungen verteilt. Für die Veränderungsmessung wurden alle Kinder mit erhöhten sozial-emotionalen Risiken herangezogen.

5.2 Stichprobe

Mittels standardisierter Beurteilung durch die Lehrkräfte wurden die Schülerinnen und Schüler für die Veränderungsmessung ermittelt. Hierbei wurde ein niedrighschwelliges Einschlusskriterium festgelegt. Ein grenzwertiger Normwert in mindestens einer der insgesamt fünf Problemskalen des SDQ qualifizierte für die Datenerhebung. Hierbei wurden 233 Schülerinnen und Schüler für die Experimentalgruppe und 172 für die Kontrollgruppe identifiziert. Die Gesamtproblembelastung unterschied sich zum Prä-Test mit Mittelwerten von 12.29 ($SD = 6.28$) in der Fördergruppe und 12.87 ($SD = 6.01$) in der Kontrollgruppe nicht signifikant ($t(390) = -0.05$, n.s.; vgl. Tab. 1). Auch hinsichtlich der Einschätzungen des Sozialklimas, des prosozialen Verhaltens und des sozial-emotionales Wissens gab es zwischen der Kontroll- und Experimentalgruppe keine Prä-Test-Unterschiede. Das Durchschnittsalter der Stichprobe lag bei 106.5 Monaten in der EG ($SD=7.5$) und 108.1 Monaten in der KG ($SD=8.5$) und die Geschlechterverteilung bei 37.8% Mädchen bzw. 62.2% Jungen in der EG und 38.4% Mädchen bzw. 61.6% Jungen in der KG. Auch hierbei zeigten sich keine signifikanten Unterschiede zwischen den Gruppen (Alter: $t(381)=1.91$, n.s.; Geschlecht: $\chi^2(1)=0.02$, n.s.).

Tab. 1: Deskriptive Statistik

Konstrukt	MZP	N _{KG}	N _{EG}	M _{KG}	SD _{KG}	M _{EG}	SD _{EG}	t(df)
Sozialklima								
FEESS SK	0	161	231	44.20	(10.41)	44.26	(10.78)	-0.05(390)
FEESS SK	1	170	218	43.70	(11.28)	44.08	(12.66)	
FEESS SK	2	171	214	44.22	(11.27)	45.79	(11.97)	
Prosoziales Verhalten								
SDQ	0	171	233	6.46	(2.36)	6.30	(2.42)	0.66(402)
SDQ	1	170	218	6.85	(2.42)	7.10	(2.41)	
SDQ	2	171	217	7.28	(2.40)	6.89	(2.41)	
Verhaltens-auffälligkeiten								
SDQ Problemwert	0	171	233	12.87	(6.01)	12.39	(6.28)	0.77(402)
SDQ Problemwert	1	170	218	11.96	(6.61)	10.47	(5.77)	
SDQ Problemwert	2	172	218	10.51	(6.73)	11.17	(7.04)	
Emotional-soziales Wissen								
IDS	0	158	214	47.61	(7.42)	46.86	(7.73)	0.93(370)
IDS	1	161	194	51.96	(7.88)	55.80	(6.26)	
IDS	2	160	187	54.41	(7.63)	55.99	(6.77)	

Anmerkungen: MZP=Messzeitpunkt, KG=Kontrollgruppe, EG=Experimentalgruppe; M=Mittelwert, SD=Standardabweichung; t=t-Wert; df=Freiheitsgrade.

5.3 Erhebungsinstrumente

Strengths and Difficulties Questionnaire (SDQ). Der SDQ ist ein Screening zur Erfassung von Verhaltensauffälligkeiten sowie prosozialem Verhalten bei Kindern und Jugendlichen im Alter von 4 bis 16 Jahren (Goodman, 1997). Das Instrument umfasst vier einzelne Problemskalen, einen Gesamtproblemwert und die Skala prosoziales Verhalten. Alle Bereiche sind durch jeweils fünf Items repräsentiert.

Die Lehrkraftversion zeigt mit $\alpha = .70$ bis $.88$ (Goodman, 2001) eine akzeptable bis gute interne Konsistenz. Die Test-Retest-Reliabilität weist nach vier bis sechs Monaten eine Stabilität von $r_{tt} = .65$ bis $.80$ auf (Goodman, 2001). Die Validität der deutschsprachigen Version des SDQs wurde u.a. von Saile (2007) in einer Studie mit 642 Kindern bestätigt.

Im Rahmen dieser Studie wurde das jeweilige Klassenmittel des Gesamtproblemwerts des SDQs als Merkmal der Klassenkomposition berücksichtigt.

Intelligence and Development Scales (IDS). Die Erfassung des emotional-sozialen Wissens erfolgte mittels vier Skalen der IDS (Grob, Meyer & Hagmann von Arx, 2009). Diese beinhalten die Funktionsbereiche Emotionen erkennen, Emotionen regulieren, soziale Situationen verstehen und sozial kompetent handeln. Die vierfaktorielle Struktur sowie der Hauptfaktor konnten konfirmatorisch bestätigt werden (Meyer, Hagmann-von Arx & Grob, 2009). Die Autorgruppe ermittelt für diesen Hauptfaktor eine sehr gute Reliabilität von $\alpha = .87$.

Fragebogen zur Erfassung der emotionalen und sozialen Schulerfahrungen (FEESS).

Die emotional-sozialen Schulerfahrungen wurden mit Hilfe des FEESS (Rauer & Schuck, 2004) erfasst. Der Fragebogen gliedert sich in zwei Teilfragebögen, welche im ersten Teil die Dimensionen Sozialklima sowie Fähigkeitskonzept erfassen und im zweiten Teil den schulischen Erfahrungsraum des Kindes bezogen auf das eigene Lernen, die Lehrkräfte und die Schule als Ganzes abfragen. Den drei Hauptfaktoren sind insgesamt 90 Items in sieben Subskalen zugeordnet. In der vorliegenden Studie kamen nur die zwei Subskalen zur Dimension Sozialklima (soziale Integration, Klassenklima) als abhängige Variable sowie die Subskala Gefühl des Angenommenseins als mögliche Einflussvariable zum Einsatz. Das Sozialklima umfasst insgesamt 22 Items, wie z.B. „in der Klasse halten wir alle zusammen“ oder „meine Mitschüler trösten mich, wenn ich traurig bin“. Das durch die Schülerinnen und Schüler wahrgenommene Lehrkraftfeedback wurde durch die 13 Items der Skala ‚Gefühl des Angenommenseins‘ mit Aussagen wie z.B. „meine Lehrer schimpfen zu viel mit mir“ oder „meine Lehrer werden ärgerlich mit mir, wenn ich etwas falsch mache“ erfasst.

Hinsichtlich der Reliabilität konnten Untersuchungen (Rauer & Schuck, 2004) zeigen, dass mit $\alpha = .77$ bis $.89$ die überwiegende Mehrheit der Kennwerte der internen Konsistenz mindestens als befriedigend, häufig als gut bzw. auch sehr gut angesehen werden können.

5.4 Datenanalyse

Die Auswertung erfolgte anhand diskreter Multilevel-Modelle für jede Outcome-Variable (u.a. Hox, 2010). Mehrebenenmodelle sind flexible statistische Instrumente, welche die gleichzeitige Berücksichtigung fixer und zufälliger Effekte und somit z.B. unterschiedlicher Charakteristika einer komplexen bzw. hierarchischen Datenstruktur erlauben. Im vorliegenden Fall werden auf Ebene eins wiederholte Messungen, die auf zweiter Ebene in einzelnen Individuen genestet sind, modelliert. Auf dritter Ebene werden wiederum Schulklasseneffekte berücksichtigt. Da in diesem Fall Entwicklungsverläufe von in Klassen zusammengefassten Schülerinnen und Schüler modelliert werden, handelt es sich um einen Spezialfall der Mehrebenenmodelle, nämlich um Mehrebenenwachstumsmodelle (Finch, Bolin & Kelley,

2014). Dabei steht im Fokus die Frage, ob in der Interventions- und Kontrollgruppe unterschiedliche Wachstumsverläufe vorliegen.

Die Vorgehensweise bei der Auswertung bestand aus folgenden Schritten: Zunächst wurde ein random-intercept-Modell mit zwei Ebenen geschätzt (Modell 1). In diesem Fall handelt es sich um ein Wachstumsmodell mit wiederholten Messungen der Outcome-Variable auf Ebene eins, die in Schülerinnen und Schülern auf Ebene zwei genestet sind. Das Modell erlaubt Schätzung der zufälligen Variation der Intercepte (bzw. Prätest-Messungen) zwischen den Probanden. Der Slope-Parameter (die Steigung bzw. der lineare Wachstumstrend) wird für alle Schülerinnen und Schüler im Gegenteil gleichgesetzt, so dass keine interindividuelle Variation der Steigung erlaubt ist und somit nur ein Durchschnittstrend für alle Probanden modelliert wird. Als zweites Modell wurde ein random-slope-Modell mit zwei Ebenen geschätzt (Modell 2). Hier werden individuelle Steigungsparameter (Slope für die Variable Messzeitpunkt) erlaubt, so dass interindividuelle Unterschiede in Wachstumstrends über Zeit als zufällige Effekte auf der Ebene zwei modelliert werden können. Im dritten Schritt wurde das Modell 1 oder 2, je nachdem, welches davon über einen besseren Fit verfügt, um eine dritte Ebene erweitert, und zwar wurde die zufällige Variation der Intercepte auf der Ebene der Schulklassen hinzugefügt. Die Klassenzugehörigkeit fungiert hierbei sowohl als Clustervariable auf Ebene drei als auch für die Gruppenzugehörigkeit (EG/KG). Das vierte Modell stellt schließlich eine Erweiterung des dritten Modells um einen zufälligen Steigungseffekt auf der dritten Ebene dar, so dass Variationen schulklassenbezogener Wachstumstrends modelliert werden können. Im Modell 2 bzw. 4 wird darüber hinaus die Kovarianz zwischen den beiden zufälligen Effekten auf der jeweiligen Ebene geschätzt. Diese steht für den Zusammenhang zwischen der Prätestmessung und dem linearen Wachstumstrend.

Die endgültige Modellauswahl erfolgte auf der Grundlage der Modellvergleiche. Hierzu wurden die relative Modellfitstatistik (AIC/BIC) und der Likelihood-Quotienten-Test herangezogen. Im Mittelpunkt der Analyse steht der Cross-Level-Interaktionsterm zwischen dem Messzeitpunkt und der Zugehörigkeit zur Kontroll- bzw. Experimentalgruppe. Der Interaktionsterm gibt den Unterschied in linearen Wachstumstrends zwischen Kontroll- und Interventionsgruppe wieder. Die Kontrollgruppe fungierte dabei als Referenzkategorie, so dass ein positiver statistisch signifikanter Interaktionsparameter für einen steileren linearen Entwicklungstrend in der Experimentalgruppe spricht.

Zur Berechnung der Effektstärken wurde die Mittelwertdifferenz zwischen EG und KG auf der Grundlage der durch das Modell vorhergesagten Individualwerte zum jeweiligen Messzeitpunkt herangezogen.

Die Auswertung erfolgte mit der Software R, Pakete lme4 (Version v1.1-21) und lmerTest (Version v3.1-0). Das Schätzverfahren war Maximum Likelihood.

6 Ergebnisse

6.1 Emotional-soziales Wissen

Hinsichtlich des sozial-emotionalen Wissens hat die Mehrebenenanalyse ergeben, dass wie im Falle des Verhaltens beim Modell mit schulklassenspezifischer Variation beider zufälligen Effekte die beste Datenanpassung vorliegt (Modell 4, vgl. Tab. 3). Somit hängt die individuelle Entwicklung des Wissens über Emotionen und soziale Situationen bei verhaltensauffälligen Schülern und Schülerinnen vom spezifischen Klassenkontext (Klassenzusammensetzung und Lehrerverhalten) ab. Konkreter steht die Entwicklung des emotional-sozialen Wissens

beispielsweise in einem Zusammenhang mit dem Erleben des Lehrerfeedbacks ($\gamma_{GdA} = 0.06$, $p=0.06$; standard. $\gamma_{GdA} = 0.01$). Das Ben & Lee Curriculum trägt ebenfalls positiv zur Verbesserung des emotional-sozialen Wissens bei: Die Cross-Level-Interaktion zeigt, dass der Steigungsparameter in der Experimentalgruppe steiler ausfällt als in der Kontrollgruppe ($\gamma_{MZP*KG/EG} = 1.59$, $p=0.06$; standard. $\gamma_{MZP*KG/EG} = 0.19$). Der Interaktionsterm verpasst zwar knapp das konventionelle α -Level von 0.05, ist aber nichtdestotrotz als statistisch bedeutsam einzuschätzen (vgl. Diskussion zu Limitationen der Studie). Die negative Korrelation zwischen dem Intercept und Slope des schulklassenbezogenen Wachstumstrends ($r_{int,slope} = -0.41$) zeigt, dass Klassen mit schwächerer mittlerer Ausprägung des emotional-sozialen Wissens eher dazu neigen, sich über die Zeit zu verbessern als Klassen mit höherem Wissensanteil.

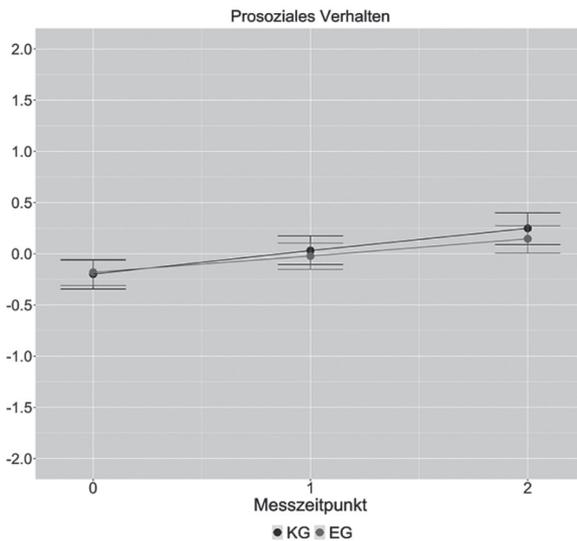


Abb. 1: Durch die Mehrebenenwachstumsmodelle vorhergesagte standardisierte Werte der vier Outcome-Variablen mit Konfidenzintervallen

6.2 Sozialklima

Bezüglich des Sozialklimas zeigte sich, dass das Modell mit interindividueller Variabilität linearer Wachstumstrends und schulklassenspezifischer Variabilität der Intercepts (M3, vgl. Tabelle 3) über eine bessere Datenpassung verfügt als die restlichen drei Modelle, d.h. dass u.a. die Annahme einer statistisch bedeutsamen schulklassenspezifischen Variation sozialintegrativer Entwicklungsprozesse sich nicht bestätigen ließ. Die statistisch signifikante Cross-Level-Interaktion zwischen dem Messzeitpunkt und der Teilnahme am Ben & Lee Curriculum deutet darauf hin, dass die Steigung in der Experimentalgruppe um einen Skaleneinheit steiler ausfällt (vgl. Tab. 2, $\gamma_{MZP*KG/EG} = 1.02$, $p<0.05$; standard. $\gamma_{MZP*KG/EG} = 0.09$). Die Abbildung 1 zeigt durch das Modell 3 vorhergesagte standardisierte Maße des Sozialklimas zu drei Zeitpunkten und veranschaulicht unterschiedliche Entwicklungen in der KG bzw. EG: Beide Gruppen schneiden im Prä-Test vergleichsmäßig ähnlich ab, zum Post-Test und Follow-up verbessert sich die EG immer stärker. Die Differenz zwischen EG und KG zum Post-Test beträgt ca. 0.2 Standardabweichungen, zum Follow-Up ca. 0.5 Standardabweichungen (vgl.

Abbildung 1). Die positive mittelstarke Korrelation zwischen Intercept und Slope auf der Ebene der Schüler und Schülerinnen ($r_{\text{intslope}} = .55$) deutet darauf hin, dass Schülerinnen und Schüler, die zu Beginn der Studie das Sozialklima in der Klasse besser einschätzten, auch einen steileren linearen Wachstumstrend aufweisen als die Kinder mit einer schlechteren Einschätzung des Sozialklimas. Unabhängig vom Treatment steht die Verbesserung des Sozialklimas in einem positiven Zusammenhang mit dem individuell erlebten Lehrkraftfeedback ($G_{\text{dA}} = 0.67, p < 0.001$; standard. $G_{\text{dA}} = 0.06$).

Tab. 2: Schätzer der fixen und zufälligen Effekte, Korrelation der zufälligen Effekte und Modellfitstatistik aus der endgültigen Auswahl der Mehrebenenwachstumsmodellen

Fixe Effekte	Sozialklima		Prosoz. Verhalten		SDQ Gesamt		IDS	
	Schätzer (SE)	t	Schätzer (SE)	t	Schätzer (SE)	t	Schätzer (SE)	t
Intercept	27.99*** (3.32)	8.44	6.60*** (0.89)	7.44	10.82*** (1.62)	6.67	46.13*** (2.68)	17.20
MZP	-1.24 (0.80)	-1.54	0.47 (0.32)	1.45	-1.23 (0.79)	-1.57	1.68 (1.35)	1.24
KG/EG	-0.20 (1.37)	-0.15	0.01 (0.38)	0.03	-0.87 (0.72)	-1.21	-0.36 (1.15)	-0.31
MZP*KG/EG	1.02* (0.49)	2.09	-0.08 (0.20)	-0.39	0.24 (0.48)	0.51	1.59# (0.84)	1.90
FEES GdA	0.67*** (0.04)	17.41	0.04*** (0.01)	3.86	-0.12*** (0.02)	-5.21	0.06# (0.03)	1.92
SDQ Klassenmittelwert	-0.41 (0.27)	-1.50	-0.14 (0.08)	-1.85	0.76*** (0.11)	7.11	0.07 (0.21)	0.35
Zufällige Effekte	SD		SD		SD		SD	
Residuum	6.39		1.28		3.63		5.00	
RI (Schüler)	5.76		1.67		4.66		4.46	
RS MZP (Schüler)	1.14							
RI (Klassen)	3.17		0.92		1.26		2.72	
RS MZP (Klassen)			0.51		1.18		2.19	
Intercept-Slope Korrelation	r		r		r		r	
Schüler	0.55							
Klassen			-0.25		-0.96		-0.41	
Modellfit								
LL	-4111.3		-2341.8		-3517.9		-3488.2	
df	1152		1149		1150		1052	
AIC	8244.6		4705.5		7057.7		6998.5	
BIC	8300.3		4761.1		7113.3		7053.1	

Anmerkungen: MZP=Messzeitpunkt, KG=Kontrollgruppe, EG=Experimentalgruppe; FEES GdA=Gefühl des An-genommenseins; RI=Random intercept; RS MZP=Random slope für die Variable Messzeitpunkt; SE=Standardfehler; SD=Standardabweichung; t=t-Wert; LL= LogLikelihood; df=Freiheitsgrade; AIC=Akaike Information Criterion; BIC=Bayesian Information Criterion; * $p < .05$, ** $p < .01$, *** $p < .001$, # $p = .06$.

6.3 Prosoziales Verhalten und Verhaltensprobleme

Die beiden Modelle aus dem Verhaltensbereich zeigten dann die beste Datenanpassung, wenn schulklassenbezogene Variationen von Prä-Testwerten und linearen Wachstumstrends (Modell 4, vgl. Tab. 3) berücksichtigt wurden. Dabei zeigte sich, dass interindividuelle Unterschiede in der Verhaltensentwicklung über die drei Zeitpunkte statistisch vernachlässigt werden können. Die Bedeutsamkeit klassenbezogener Varianz der Verhaltensentwicklung kann damit erklärt werden, dass schülerbezogene Verhaltenseinschätzungen mittels der Lehrerversion des SDQ-Fragebogens erhoben wurden und somit „innerhalb“ der jeweiligen Lehrkraft korreliert sind. Ferner zeigte die Mehrebenenanalyse, dass sowohl die Entwicklung des prosozialen Verhaltens, als auch die Entwicklung von Verhaltensauffälligkeiten in der Kontrollgruppe und in der Gruppe mit Ben & Lee Curriculum parallel verläuft (vgl. Tab. 2 und Abb.1) und somit keine bedeutsame Interaktion zwischen Messzeitpunkt und Teilnahme an der Ben & Lee Förderung vorliegt. Ferner lässt sich aus den beiden Modellen schlussfolgern, dass unabhängig vom Treatment die Verbesserung des prosozialen Verhaltens vom Erleben des Lehrerfeedbacks ($\gamma_{GdA} = 0.04$, $p < 0.001$; standard. $\gamma_{GdA} = 0.01$) und der Rückgang von Verhaltensauffälligkeiten sowohl vom Erleben des Lehrerfeedbacks ($\gamma_{GdA} = -0.12$, $p < 0.001$; standard. $\gamma_{GdA} = -0.02$) als auch von der Klassenkomposition abhängt ($\gamma = 0.76$, $p < 0.001$; standard. $\gamma = 0.12$). Die negative Korrelation zwischen dem Intercept und Slope des Wachstumstrends für prosoziales Verhalten auf Schulklassenebene ($r_{int,slope} = -.25$) deutet darauf hin, dass Schulklassen mit besseren mittleren Sozialverhaltenseinschätzungen stärker dazu tendieren, sich über die drei Messzeitpunkte zu verschlechtern, als Klassen mit ungünstigeren Einschätzungen des Verhaltens. Die starke negative Korrelation zwischen dem Intercept und Slope des Wachstumstrends für Verhaltensauffälligkeiten auf Schulklassenebene ($r_{int,slope} = -.96$) spricht im Gegenteil dafür, dass Klassen mit ungünstigen Verhaltenseinschätzungen sich mit der Zeit eher verbessern. Hierzu muss man zusätzlich berücksichtigen, dass der Slope des linearen Wachstumstrends für Verhaltensauffälligkeiten insgesamt sehr flach ausfällt.

Tab. 3: Modellfitstatistik und Ergebnisse der Likelihood-Ratio-Tests für alle geschätzten Mehrebenenmodelle

Modell	LL	df	AIC	BIC	Test	$\Delta\chi^2(\Delta df)$	p
Soziale Integration							
M1 RI (Schüler)	-4131.4	1155	8278.9	8319.3	-	-	-
M2 RI + RS (Schüler)	-4127.9	1153	8275.8	8326.3	M1	7.11(2)	.03
M3 RI + RS (Schüler) + RI (Klassen)	-4111.3	1152	8244.6	8300.3	M2	33.15(1)	.00
M4 RI + RS (Schüler) + RI + RS (Klassen)	-4111.2	1150	8248.4	8314.2	M3	0.20(2)	.90
Prosoziales Verhalten							
M1 RI (Schüler)	-2391.6	1152	4799.3	4839.7	-	-	-
M2 RI + RS (Schüler)	-2389.7	1150	4799.5	4850.0	M1	3.82(2)	.15
M3 RI (Schüler) + RI (Klassen)	-2371.3	1151	4760.6	4806.1	M1	40.67(1)	.00
M4 RI (Schüler) + RI + RS (Klassen)	-2341.8	1149	4705.5	4761.1	M3	59.07(2)	.00
SDQ Problemwert							

Modell	LL	df	AIC	BIC	Test	$\Delta\chi^2(\Delta df)$	p
M1 RI (Schüler)	-3536.0	1153	7088.1	7128.6	-	-	-
M2 RI + RS (Schüler)	-3534.0	1151	7088.4	7139.0	M1	3.69(2)	.16
M3 RI (Schüler) + RI (Klassen)	-3536.0	1152	7090.0	7135.5	M1	0.09(1)	.76
M4 RI (Schüler) + RI + RS (Klassen)	-3517.9	1150	7057.7	7113.3	M1	36.40(3)	.00
IDS							
M1 RI (Schüler)	-3533.3	1055	7086.7	7126.4	-	-	-
M2 RI + RS (Schüler)	-3534.8	1053	7089.6	7139.3	M1	1.10(2)	.58
M3 RI (Schüler) + RI (Klassen)	-3519.0	1054	7056.0	7100.7	M1	32.67(1)	.00
M4 RI (Schüler) + RI + RS (Klassen)	-3488.2	1052	6998.5	7053.1	M3	61.50(2)	.00

Anmerkung: M1–M4=Modelle 1 bis 4; RI=Random Intercept; RS=Random Slope (für die Variable Messzeitpunkt); LL=LogLikelihood; df =Freiheitsgrade; AIC=Akaike Information Criterion; BIC=Bayesian Information Criterion; Test=gegen welches Modell wird getestet; $\Delta\chi^2$ =Chi-Quadrat-Wert aus dem Likelihood-Quotienten-Test; Δdf =Differenz der Freiheitsgrade zwischen zwei verglichenen Modellen; die endgültig ausgewählten Modelle sind grau unterlegt.

7 Diskussion

International wie national liegen die Effektstärken von SEL-Maßnahmen beim Wissensaufbau in Bezug auf den Präventionsgegenstand bei $d = 0.32$ bis 0.70 (Beelmann et al., 2014; Durlak, Weissberg, Dymnicki, Taylor & Schellinger, 2011; Hövele et al., 2019). Mit einer Effektstärke von $d = 0.5$ bei der Erweiterung des sozial-emotionalen Wissens der Schülerinnen und Schüler und von $d = 0.2$ hinsichtlich der Verbesserung des Sozialklimas zeigt das Ben & Lee Programm im Vergleich zu weiteren Fördermaßnahmen eine probate Wirksamkeit.

Das emotional-soziale Wissen sowie das Sozialklima wird zudem von dem von den Kindern wahrgenommenen Lehrkraftfeedback beeinflusst, und dies unabhängig vom Ben & Lee Curriculum. In Weiterentwicklungen des Programms sollte daher ein Schwerpunkt daraufgesetzt werden, Lehrpersonen in ihrem Rückmeldeverhalten positiv zu unterstützen.

Für das von den Lernpersonen wahrgenommene Verhalten der Schülerinnen und Schüler ergeben sich keine signifikanten Programmeffekte. Die Klassenkompositionen und das Feedbackverhalten nehmen jedoch einen bedeutsamen Einfluss und sollten bei zukünftigen Umsetzungen der Maßnahme Berücksichtigung finden.

7.1 Limitationen

Zu den zentralen Limitationen der Studie zählt die vergleichsweise geringe Stichprobe ($N_{\text{Schüler}} = 405$, $N_{\text{Klassen}} = 36$). Diese ermöglicht auf der einen Seite zwar Schätzung komplexer Mehrebenenwachstumsmodelle mit relativ stabilen Maximum-Likelihood-Schätzern von bis zu fünf zufälligen Effekten. Die statistische Teststärke solcher komplexer Modelle bei einer geringen Stichprobe ist auf der anderen Seite deutlich eingeschränkt und führt somit zu höheren Standardfehlern. Im Rahmen vorliegender Datenanalysen wurde eine der Komplexität der Daten am besten entsprechende Modellierungsmöglichkeit gewählt und in Kauf genommen, dass

dies bei kleiner Stichprobe hohe Standardfehler nach sich ziehen würde. Deshalb wurden bei der Interpretation der Ergebnisse bezüglich Treatmenteffektivität eher Effektstärken als das konventionelle Signifikanzniveau von 0.05 fokussiert. Fokussierung von p -Werten bei Entscheidungen über Treatmenteffektivität wird in der Forschung aktuell verstärkt kritisiert (vgl. z.B. Goodman, 2008).

Eine weitere Limitation besteht hinsichtlich der Operationalisierungen der Klassenkomposition und des Lehrkraftfeedbacks. Mit Items wie ‚meine Lehrer schimpfen zu viel mit mir‘ wird in der Skala GdA des FEES das Empfinden der Schülerinnen und Schüler erfasst, jedoch nicht, ob die Lehrperson tatsächlich häufig öffentliches Feedback gibt. Um dies in zukünftigen Studien zu erfassen, wären Unterrichtsbeobachtungen oder Videografien hilfreich.

7.2 Praktische Implikationen

Die Integration von SEL in das reguläre Curriculum des Deutsch- und Sachunterrichts, so wie sie im Ben & Lee Programm vorgenommen wird, scheint ein vielversprechender Zugang zur schulischen Verhaltensförderung zu sein. Trotz der gleichzeitigen Fokussierung von Fach- und Entwicklungszielen konnten Schülerinnen und Schüler mit risikobehafteten Verhaltensweisen ihr sozial-emotionales Wissen bedeutsam verbessern. Vor diesem Hintergrund kann der Ansatz als Option zur Lösung des von Karing, Beermann und Haase (2015) beschriebenen Problems von Schulen verstanden werden, die zeitlichen Ressourcen für präventive Maßnahmen aufzubringen und die Programme in das Curriculum zu integrieren bzw. die Zulassung der Programme für den Lehrplan zu erhalten.

Für die nachhaltige Veränderung des Verhaltens ist neben dem Einsatz von SEL-Programmen zudem aber auch eine bessere Planung der Klassenzusammensetzung erforderlich. Müller und Zurbriggen (2016) stellten fest, dass Schulklassen im Unterschied zu Cliques und außerschulischen Freundschaftsgruppen i.d.R. als unfreiwillige Gleichaltrigengruppen angesehen werden können. D.h., dass die Kinder und Jugendlichen im Normalfall nicht selbst entscheiden können, in welche Klasse sie gehen. Die besuchte Klasse und die dort entstandene Klassenkomposition hat jedoch maßgeblichen Einfluss auf die Entwicklung des gezeigten Verhaltens. Um einerseits Kinder und Jugendliche mit herausforderndem Verhalten nicht zu benachteiligen und andererseits die Wirksamkeit von SEL zu erhöhen, sollte in der Praxis auf eine ausgewogene Klassenzusammensetzung geachtet werden. Von Überlegungen der Praxis, mit der Intention (sonder-)pädagogische Ressourcen zu bündeln, „Schwerpunktklassen“ mit einer Vielzahl an Schülerinnen und Schülern mit Schwierigkeiten im Verhalten zu bilden, ist auf Basis der hier vorliegenden Befunde klar abzuraten. Anstelle solcher Spezialklassen sollten Schulen Kinder mit Verhaltensauffälligkeiten schulweit aufteilen und SEL-Programme in allen Klassen durch die Klassenleitung umsetzen lassen. Durlak et al. (2011) ermitteln in diesem Zusammenhang, dass sich die Wirksamkeit von universellen SEL-Programmen insbesondere bei einer Umsetzung durch die Klassenlehrerinnen und -lehrer einstellt.

Ebenso wichtig wie die Klassenkomposition ist das Feedbackverhalten der Lehrperson. Huber (2019) führt an, dass Lehrkräfte durch eine Fokussierung auf ihr proaktives Classroom-Management Verhaltensprobleme und Unterrichtsstörungen im Vorfeld minimieren können und dadurch idealerweise seltener in die Situationen kommen, auf Unterrichtsstörungen mit öffentlichem negativen Feedback zu reagieren. Borchert (2008) gibt in diesem Zusammenhang zu bedenken, dass Kinder im Durchschnitt nur eine 10 %ige Wahrscheinlichkeit haben, für unterrichtlich erwünschtes Verhalten belohnt zu werden. Positives Verhalten wird nach Borchert (2008) als selbstverständlich vorausgesetzt und daher nicht registriert, wohingegen

unverwünschtes Verhalten über intensive Beachtung verstärkt und aufrechterhalten wird. In einer kontrollierten Einzelfallstudie zum gezielten Einsatz von proaktiven Classroom-Management-Strategien im Rahmen einer SEL-Förderung im inklusiven Biologieunterricht konnten Ferreira González, Hövel, Hennemann und Schlüter (2019) positive Effekte auf das Unterrichtsverhalten von Schülerinnen und Schüler mit herausforderndem Verhalten feststellen. Auch bei der praktischen Umsetzung von Ben & Lee sollte daher intensiv auf ein proaktives Classroom-Management geachtet werden.

Literatur

- Barry, C. M. & Wentzel, K. R. (2006). Friend influence on prosocial behavior: The role of motivational factors and friendship characteristics. *Developmental Psychology*, 42, 153-163. doi:10.1037/0012-1649.42.1.153
- Beelmann, A., Pfost, M. & Schmitt, C. (2014). Prävention und Gesundheitsförderung bei Kindern und Jugendlichen. Eine Meta-Analyse der deutschsprachigen Wirksamkeitsforschung. *Zeitschrift für Gesundheitspsychologie*, 22, 1-14. doi: 10.1026/0943-8149/a000104.
- Borchert, J. (2008). Lehrer-Schüler-Interaktion. In B. Gasteiger-Klicpera, H. Julius, C. Klicpera, J. Borchert & H. Götzte (Hrsg.), *Handbuch Sonderpädagogik. 3. Sonderpädagogik der sozialen und emotionalen Entwicklung* (S. 740-751). Göttingen: Hogrefe.
- Crick, N. & Dodge, K. (1994). A Review and Reformulation of Social Information-Processing Mechanisms in Children's Social Adjustment. *Psychological Bulletin*, 115, 74-101. doi:10.1037/0033-2909.115.1.74
- Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D. & Schellinger, K. B. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions. *Child Development*, 82, 405-432. doi:10.1111/j.1467-8624.2010.01564.x
- Feinman, S. (1992). Social referencing and conformity. In S. Feinman, *Social Referencing and the Social Construction of Reality in Infancy* (S. 229-268). New York: Plenum Press. doi:10.1007/978-1-4899-2462-9_10
- Ferreira González, L., Hövel, D., Hennemann, T. & Schlüter, K. (2019). Auswirkungen des gezielten Einsatzes von Classroom Management Strategien im inklusiven Fachunterricht Biologie auf das Unterrichtsverhalten von Schülern unter erhöhten Risiken aus Perspektive der Lehrperson – Eine Einzelfallstudie. *Empirische Sonderpädagogik*, 1, 53-70.
- Finch, W., Bolin, J. & Kelley, K. (2014). *Multilevel Modeling Using R*. Boca Raton: CRC Press.
- Goodman, R. (1997). The Strengths and Difficulties Questionnaire: A Research Note. *Journal of Child Psychology and Psychiatry*, 38, 581-586. doi:10.1111/j.1469-7610.1997.tb01545.x
- Goodman, R. (2001). The psychometric properties of the Strengths and Difficulties Questionnaire. *Journal of the American Academy of Child and Adolescent Psychiatry*, 40(11), 1337-1345. doi:10.1097/00004583-200111000-00015
- Grob, A., Meyer, C. & Hagmann von Arx, P. (2009). *Intelligence and Development Scales*. Bern: Hans Huber, Hogrefe.
- Helmeke, A. (2009). *Unterrichtsqualität und Lehrerprofessionalität. Diagnose, Evaluation und Verbesserung des Unterrichts*. Seelze-Velber: Kallmeyer.
- Hövel, D., Hennemann, T. & Rietz, C. (2019). Meta-Analyse programmatischer-präventiver Förderung der emotionalen und sozialen Entwicklung in der Grundschule. *Emotionale und Soziale Entwicklung (ESE) in der Pädagogik der Erziehungshilfe*, 1, 38-55.
- Hox, J. (2010). *Multilevel Analysis. Techniques and Applications*. New York: Routledge.
- Huber, C. (2019). Ein integriertes Rahmenmodell zur Förderung sozialer Integration im inklusiven Unterricht. Sozialpsychologische Grundlagen, empirische Befunde und schulpraktische Ableitungen. *VHN*, 88, 27-43. doi:10.2378/vhn2019.art06d
- Huber, C., Gebhardt, M. & Schwab, S. (2015). Lehrkraftfeedback oder Spaß beim Spiel? Eine Experimentalstudie zum Einfluss von Lehrkraftfeedback auf die soziale Akzeptanz bei Grundschulkindern. *Psychologie in Erziehung und Unterricht*, 62, 51-64. doi:10.2378/peu2015.art04d
- Huber, C., Gerullis, A., Gebhardt, M. & Schwab, S. (2018). The impact of social referencing on social acceptance of children with disabilities and migrant background: an experimental study in primary school settings. *European Journal of Special Needs Education*, 33, 269-285. doi:10.1080/08856257.2018.1424778
- Karing, C., Beelmann, A. & Haase, A. (2015). Herausforderungen von Präventionsarbeit an Grundschulen. *Prävention und Gesundheitsförderung*, 229-234. doi:10.1007/s11553-015-0494-1
- Klipker, K., Baumgarten, F., Göbel, K., Lampert, T. & Hölling, H. (2018). Psychische Auffälligkeiten bei Kindern und Jugendlichen in Deutschland – Querschnittergebnisse aus KiGGS Welle 2 und Trends. *Journal of Health Monitoring*, 37-45. doi:10.17886/RKI-GBE-2018-077

- Lemerise, E. & Arsenio, W. (2000). An Integrated Model of Emotion Processes and Cognition in Social Information Processing. *Child Development*, 71, 107-118. doi:10.1111/1467-8624.00124
- McCormick, M., Neuhaus, R., Horn, E., O'Connor, E., White, H., Harding, S., . . . McCloskey, S. (2019). Long-Term Effects of Social-Emotional Learning on Receipt of Special Education and Grade Retention: Evidence From a Randomized Trial of INSIGHTS. *AEA Open*, 5(3), 1-21. doi:10.1177/2332858419867290
- Meyer, C. S., Hagmann-von Arx, P. & Grob, A. (2009). Die Intelligence and Development Scale Sozial-Emotionale Kompetenz (IDS-SEK). *Diagnostica*, 55, 234-244. doi:10.1026/0012-1924.55.4.234
- Müller, C. M. & Zurbriggen, C. (2016). An Overview of Classroom Composition Research on Social-emotional Outcomes – Introduction to the Special issue. *Journal of Cognitive Education and Psychology*, 15, 163-184. doi:10.1891/1945-8959.15.2.163
- Prinstein, M. J. (2007). Moderators of peer contagion: A longitudinal examination of depression socialization between adolescents and their best friends. *Journal of Clinical Child and Adolescent Psychology*, 36, 159-170. doi:10.1080/15374410701274934
- Rauer, W. & Schuck, K.-D. (2004). *FEES 1-2. Fragebogen zur Erfassung emotionaler und sozialer Schulerfahrung von Grundschulkindern erster und zweiter Klassen*. Göttingen: Beltz-Tests.
- Reicher, H. (2010). *Sozial-emotionales Lernen im Kontext inklusiver Pädagogik: Potenziale und Perspektiven*. Graz: Grazer Universitätsverlag.
- Saile, H. (2007). Psychometrische Befunde zur Lehrerversion des „Strengths and Difficulties Questionnaire“ (SDQ-L). Validierung anhand soziometrischer Indizes. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 39, 25-32. doi:10.1026/0049-8637.39.1.25
- Urban, M., Hövel, D. & Hennemann, T. (2018). *Ben & Lee 3. und 4. Klasse. Programm zur Förderung sozial-emotionaler Kompetenzen in Verbindung mit fachlichen Zielen des Deutsch- und Sachunterrichts*. Köln: hpa edition.
- Wilson, S. & Lipsey, M. (2007). School-based interventions for aggressive and disruptive behavior: Update of a meta-analysis. *American Journal of Preventive Medicine*, 33, 130-143. doi:10.1016/j.amepre.2007.04.011