

Wuntke, Lena Varuna

Schulische Förderung der emotionalen und sozialen Entwicklung von Kindern durch Erweiterung spezifischer Lehrkraftkompetenzen

Opladen • Berlin • Toronto : Budrich Academic Press 2024, 256 S. - (Dissertation, Universität Greifswald, 2022)



Quellenangabe/ Reference:

Wuntke, Lena Varuna: Schulische Förderung der emotionalen und sozialen Entwicklung von Kindern durch Erweiterung spezifischer Lehrkraftkompetenzen. Opladen • Berlin • Toronto : Budrich Academic Press 2024, 256 S. - (Dissertation, Universität Greifswald, 2022) - URN: urn:nbn:de:0111-pedocs-314531 - DOI: 10.25656/01:31453; 10.3224/96665091

<https://nbn-resolving.org/urn:nbn:de:0111-pedocs-314531>

<https://doi.org/10.25656/01:31453>

in Kooperation mit / in cooperation with:



<https://www.budrich.de>

Nutzungsbedingungen

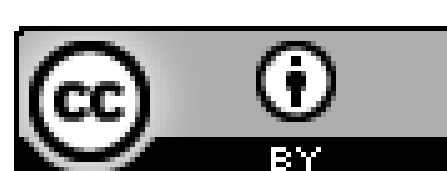
Dieses Dokument steht unter folgender Creative Commons-Lizenz: <http://creativecommons.org/licenses/by/4.0/deed.de> - Sie dürfen das Werk bzw. den Inhalt vervielfältigen, verbreiten und öffentlich zugänglich machen sowie Abwandlungen und Bearbeitungen des Werkes bzw. Inhaltes anfertigen, solange Sie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

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use

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

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



Kontakt / Contact:

peDOCS

DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation

Informationszentrum (IZ) Bildung

E-Mail: pedocs@dipf.de

Internet: www.pedocs.de

Mitglied der


Leibniz-Gemeinschaft



Schulische Förderung der emotionalen und sozialen Entwicklung von Kindern durch Erweiterung spezifi- scher Lehrkraftkompetenzen

Lena Varuna Wuntke

Schulische Förderung der emotionalen und
sozialen Entwicklung von Kindern durch
Erweiterung spezifischer Lehrkraftkompetenzen

Lena Varuna Wuntke

Schulische Förderung der
emotionalen und sozialen
Entwicklung von Kindern durch
Erweiterung spezifischer
Lehrkraftkompetenzen

Budrich Academic Press
Opladen • Berlin • Toronto 2024

Bibliografische Information der Deutschen Nationalbibliothek
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de> abrufbar.

Die Open-Access-Publikation wurde durch die Universität Greifswald gefördert.

© 2024 Dieses Werk ist bei der Budrich Academic Press GmbH erschienen und steht unter der Creative Commons Lizenz Attribution 4.0 International (CC BY 4.0): <https://creativecommons.org/licenses/by/4.0/>
Diese Lizenz erlaubt die Verbreitung, Speicherung, Vervielfältigung und Bearbeitung unter Angabe der UrheberInnen, Rechte, Änderungen und verwendeten Lizenz. www.budrich-academic-press.de



Die Verwendung von Materialien Dritter in diesem Buch bedeutet nicht, dass diese ebenfalls der genannten Creative-Commons-Lizenz unterliegen. Steht das verwendete Material nicht unter der genannten Creative-Commons-Lizenz und ist die betreffende Handlung gesetzlich nicht gestattet, ist die Einwilligung des jeweiligen Rechteinhabers für die Weiterverwendung einzuholen. In dem vorliegenden Werk verwendete Marken, Unternehmensnamen, allgemein beschreibende Bezeichnungen etc. dürfen nicht frei genutzt werden. Die Rechte des jeweiligen Rechteinhabers müssen beachtet werden, und die Nutzung unterliegt den Regeln des Markenrechts, auch ohne gesonderten Hinweis.

Dieses Buch steht im Open-Access-Bereich der Verlagsseite zum kostenlosen Download bereit (<https://doi.org/10.3224/96665091>).
Eine kostenpflichtige Druckversion kann über den Verlag bezogen werden. Die Seitenzahlen in der Druck- und Onlineversion sind identisch.

ISBN 978-3-96665-091-5 (Paperback)
eISBN 978-3-96665-905-5 (PDF)
DOI 10.3224/96665091

Umschlaggestaltung: Bettina Lehfeldt, Kleinmachnow – www.lehfeldtgraphic.de
Satz: Angelika Schulz, Zülpich

Danksagung

Die vorliegende Doktorarbeit wurde in der Zeit von Mai 2018 bis Dezember 2022 am Institut für Erziehungswissenschaft der Universität Greifswald unter der Leitung von Prof. Dr. Kathrin Mahlau angefertigt.

An dieser Stelle möchte ich mich bei all denjenigen bedanken, die mich auf vielfältige Weise dabei unterstützt haben:

Mein außerordentlicher Dank gilt meiner Doktormutter Prof. Dr. Kathrin Mahlau, für die ausgezeichnete Betreuung, die konstante Unterstützung sowie für die stets geöffnete Tür bei der Umsetzung der gesamten Arbeit. Als besonders wertvoll habe ich dabei die Freiräume empfunden, die sie mir während des gesamten Forschungsprojektes gewährte, was wesentlich zum Gelingen dieser Arbeit beitrug.

Prof. Dr. Gino Casale danke ich für die Anfertigung des Zweitgutachtens, die konstruktive Begleitung sowie die wertvollen konzeptionellen und methodischen Anregungen.

Besonders bedanken möchte ich mich bei den Lehrkräften und Schulleitungen, die sich an der Studie beteiligt haben und ohne deren engagiertes Mitwirken diese Arbeit nicht möglich gewesen wäre.

Für ihren geschätzten Beitrag des Korrekturlesens bedanke ich mich bei meinen lieben Freund*innen Dr. Pavla Schäfer und Prof. Dr. Peter Meisel. Mein Dank gilt außerdem Dr. Mirjam Reiß, für die tatkräftige wie professionelle Unterstützung bei der Durchführung der Studie und ihren Beitrag bei meinem Einstieg in das Universum der Statistik.

Weiter bedanke ich mich bei meinem Bruder Leo Stippl, der mich auf meinem Weg mit produktiven Gesprächen und lieben Worten begleitet hat.

Von Herzen danken möchte ich außerdem meinem Ehemann Christian Wuntke. Sein geduldiges Zuhören, die unzähligen Gespräche, seine Hilfe in Wort und Tat, die liebevollen Aufmunterungen wie auch seine Rücksichtnahme haben mich während der gesamten Zeit getragen und gestützt.

Eine ganz besondere Danksagung gilt schließlich meiner Patentante Beate Welsch, die mich in meiner akademischen Laufbahn uneingeschränkt unterstützt hat und mir in so vielen Bereichen meines Lebens eine Mentorin ist und war. Ihr widme ich diese Arbeit.

Lena Varuna Wuntke

Inhalt

Danksagung	5
Abkürzungsverzeichnis	11
Abbildungsverzeichnis	13
Tabellenverzeichnis.....	15
1 Einleitung.....	19
2 Forschungsstand.....	23
2.1 Schulische Förderung der ESE von Kindern.....	23
2.1.1 Zentrale Komponenten der ESE	23
2.1.2 Verhaltensauffälligkeiten und emotional-soziale Entwicklungsstörungen.....	29
2.1.2.1 Definition	30
2.1.2.2 Klassifikation und Einteilung.....	33
2.1.2.3 Erklärungsansätze zur Entstehung von Verhaltensauffälligkeiten und emotional-sozialen Entwicklungsstörungen.....	36
2.1.2.4 Prävalenzen und Komorbiditäten.....	41
2.1.2.5 Prognose und spezifische Verlaufsformen.....	42
2.1.2.6 Zusammenfassung und Schlussfolgerungen	43
2.1.3 Schulische Förderung im Bereich ESE.....	43
2.1.3.1 Postulat der Prävention	44
2.1.3.2 Empirisch bewährte Handlungsansätze.....	48
2.1.3.3 Evidenzbasierte Praxis	50
2.1.3.4 Mehrstufige schulische Förderkonzeptionen	52
2.1.4 Zusammenfassung und Schlussfolgerungen für die schulische Förderung der ESE von Kindern.....	61
2.2 Professionalisierung von Lehrkräften für die schulische Förderung der ESE von Kindern.....	62
2.2.1 Das Konstrukt Lehrkraftprofessionalisierung.....	62
2.2.2 Professionelle Handlungskompetenz von Lehrkräften	66
2.2.2.1 Das COACTIV-Kompetenzmodell und seine Erweiterung.....	67

2.2.2.2	(Weiter-)Entwicklung professioneller Handlungskompetenz von Lehrkräften	78
2.2.3	Lehrkraftprofessionalisierung durch Fortbildung	80
2.2.3.1	Wirkungsweise und Wirksamkeit von Lehrkraftfortbildungen.....	83
2.2.3.2	Erkenntnisse zu Lehrkraftfortbildungen zur Förderung der ESE bei Schüler*innen.....	89
2.2.4	Zusammenfassung und Schlussfolgerungen zur Professionalisierung von Lehrkräften für die schulische Förderung der ESE von Kindern	94
2.3	Gesamtzusammenfassung zur schulischen Förderung der ESE durch Lehrkraftprofessionalisierung	96
3	Entwurf eines Professionalisierungskonzepts zur Förderung der ESE von Grundschüler*innen durch Erweiterung spezifischer Lehrkraftkompetenzen	99
3.1	Ziele und konzeptionelle Überlegungen.....	99
3.2	Handlungsfelder und Professionalisierungselemente.....	104
3.2.1	Handlungsfeld A: Erweiterung des Lehrkraftwissens.....	104
3.2.1.1	Professionalisierungselement I: Grundlagen der ESE von Kindern und etwaige Störungsformen	104
3.2.1.2	Professionalisierungselement II: Verhaltensverlaufdiagnostik.....	104
3.2.1.3	Professionalisierungselement III: Handlungsstrategien der Primären und Sekundären Prävention	105
3.2.2	Handlungsfeld B: Veränderung des unterrichtlichen Lehrkrafthandelns	106
3.2.2.1	Professionalisierungselement IV: Verhaltensmonitoring	106
3.2.2.2	Professionalisierungselement V: Einsatz interventiver Handlungsmöglichkeiten vor dem Hintergrund präventiver Handlungsstrategien	107
3.2.3	Handlungsfeld C: Reflexion durch Beratung.....	107
3.2.3.1	Professionalisierungselement VI: Psychosozial-horizontale Beratung.....	108
3.2.4	Zusammenfassung	108
3.3	Überlegungen zu den Stärken und Schwächen der Konzeption.....	109

4	Evaluation des entwickelten Professionalisierungskonzepts	115
4.1	Ziele der Evaluationsstudie und Vorbemerkungen	115
4.2	Fragestellungen und Hypothesen	116
4.2.1	Fragestellungen.....	116
4.2.2	Hypothesen	118
4.3	Methode	120
4.3.1	Studiendesign und Untersuchungsplan	120
4.3.2	Konkrete Durchführung der Evaluationsstudie.....	125
4.3.3	Bildung und Beschreibung der Untersuchungsgruppen.....	128
4.3.3.1	Lehrkräfte.....	130
4.3.3.2	Schüler*innen	130
4.3.4	Messverfahren	132
4.3.4.1	Erfassung der Schüler*innenoutcomes	133
4.3.4.2	Erfassung des unterrichtlichen Lehrkrafthandelns	137
4.3.4.3	Erfassung der Lehrkraftkognitionen	137
4.3.4.4	Erfassung der Lehrkraftreaktionen.....	138
4.3.5	Auswertungsverfahren, statistische Prüfgrößen und Interpretationsmaße	140
4.3.5.1	Verlaufsdiagnostische Auswertungsverfahren und Prüfgrößen	142
4.3.5.2	Statusdiagnostische Auswertungsverfahren und Prüfgrößen	150
4.3.5.3	Qualitative Auswertungsverfahren und Prüfgrößen	151
5	Ergebnisse	153
5.1	Hypothesenüberprüfung.....	153
5.1.1	Schüler*innenoutcomes.....	153
5.1.2	Lehrkrafthandeln.....	171
5.1.3	Lehrkraftkognitionen	176
5.1.4	Lehrkraftreaktionen	181
5.2	Ergebniszusammenfassung	189

6	Diskussion	193
6.1	Methodenkritische Reflexion der Ergebnisse.....	193
6.2	Diskussion der Ergebnisse	201
6.2.1	Schüler*innenoutcomes.....	201
6.2.2	Lehrkrafthandeln.....	206
6.2.3	Lehrkraftkognitionen	208
6.2.4	Lehrkraftreaktionen	213
6.3	Schlussfolgerungen für die Konzeption zukünftiger Professionalisierungsmaßnahmen	218
6.4	Ausblick für Forschung, Lehre und Praxis.....	222
7	Fazit.....	231
	Literaturverzeichnis	235

Abkürzungsverzeichnis

DBR	Direct Behavior Rating
EbP	Evidenzbasierte Praxis
ESE	Emotional-soziale Entwicklung
FSP ESE	Förderschwerpunkt Emotional-soziale Entwicklung
GPk-IT	General Pedagogical Knowledge for Inclusive Teaching
GPk-IT-SEL	General Pedagogical Knowledge for Inclusive Teaching focussing on Social and Emotional Learning
KMK	Kultusministerkonferenz
MBD	Multiple-Baseline-Design
MIS	Multi-Item-Scale
MZP	Messzeitpunkt
NAP	Nonoverlap of all Pairs
PAND	Percentage of all nonoverlapping Data
PET	Percentage of data points exceeding the trend
RTI	Response-to-Intervention
SDQ	Strenghts and Difficulties Questionnaire
SF ESE	Sonderpädagogischer Förderbedarf mit dem Schwerpunkt emotional-soziale Entwicklung
SIS	Single-Item-Scale
SKI	Sozial-kognitive Informationsverarbeitung
UN-BRK	Übereinkommen der Vereinten Nationen über die Rechte von Menschen mit Behinderung
WirkLehr	Skala Lehrer-Selbstwirksamkeitserwartung

Abbildungsverzeichnis

Abbildung 1:	Sozial-kognitive Informationsverarbeitung	39
Abbildung 2:	Ablauf eines evidenzbasierten Entscheidungsprozesses	51
Abbildung 3:	Das Response-to-Intervention-Modell mit den nach Intensität und Spezifität gestuften evidenzbasierten Maßnahmen zur emotional-sozialen Unterstützung.....	54
Abbildung 4:	Modell professioneller Handlungskompetenz.....	68
Abbildung 5:	Professionelle Handlungskompetenz als Kontinuum.....	70
Abbildung 6:	Professionelles Wissen von Lehrkräften und Anforderungen der Inklusion.....	72
Abbildung 7:	Modellierung der Lehrkraftkompetenzen im Bereich des pädagogisch-psychologischen Professionswissens.....	77
Abbildung 8:	Erweitertes Angebots- und Nutzungsmodell zur Erklärung der Wirksamkeit von Lehrkraftfortbildungen.....	82
Abbildung 9:	Allgemeine und spezifische, auf den FSP ESE bezo- gene Kriterien erfolgreicher Lehrkraftfortbildungen, die als Grundlage für die Konzipierung von Professio- nalisierungsmaßnahmen im Kontext Inklusion dienen können/sollten.....	95
Abbildung 10:	Gesamtzusammenfassung des Forschungsstands zur schulischen Förderung der ESE von Kindern durch Lehrkraftprofessionalisierung; Zentrale Argumenta- tionslinien und Ableitung des entwickelten Professionalisierungskonzepts für Grundschullehrkräfte zur Erweiterung spezifischer Lehrkraftkompetenzen	97
Abbildung 11:	Das im Rahmen des entwickelten Professionali- sierungskonzepts zu erweiternde Professionswissen der Lehrkräfte gegliedert nach Wissensbereichen, Wissensfacetten und Schüler*innenoutcomes	101
Abbildung 12:	Handlungsfelder im Professionalisierungskonzept mit den dazugehörigen Professionalisierungselementen.....	103

Abbildung 13:	Aufbau und Ablauf des Professionalisierungskonzepts für Grundschullehrkräfte zur Förderung der ESE ihrer Schüler*innen (Schuljahr 2018/19) sowie empirisch untersuchte Wirkungsebenen.....	110
Abbildung 14:	Studiendesign und Untersuchungsplan zur evaluativen Begleitung der Professionalisierungsmaßnahmen (Treatment) für Grundschullehrkräfte zur Förderung der ESE ihrer Schüler*innen.....	124
Abbildung 15:	Bildung der Untersuchungsgruppen der Lehrkräfte und Schüler*innen anhand von Stichprobenziehungen, Dropouts und Ausschlussverfahren	129
Abbildung 16:	Bestimmung der Überlappungsfreiheit im Zuge der PAND-Berechnung.....	146
Abbildung 17:	Regressionsmodell von Huitema und McKean (2000) zur Einzelfallanalyse.....	149
Abbildung 18:	Auflistung der während der Intervention eingesetzten Handlungsmöglichkeiten nach Erklärungs- und Handlungsansatz sowie die Abstufung der Einsatzhäufigkeit..	171
Abbildung 19:	Einsatzhäufigkeit der <i>84 Handlungsmöglichkeiten</i> während der Interventionsphase für jeden Einzelfall und im Durchschnitt über alle Schüler*innen	173
Abbildung 20:	Zentrale Schlüsselfaktoren für die Professionalisierung von Lehrkräften für die schulische Inklusion im Bereich ESE.....	221

Tabellenverzeichnis

Tabelle 1:	Das schrittweise Erwerben von Schlüsselfertigkeiten zur Entwicklung von Emotionswissen.....	24
Tabelle 2:	Beispiele für Strategien der reflexiven Emotionsregulation und deren Initiatoren.....	26
Tabelle 3:	Entwicklungsspannen mit Entwicklungsaufgaben von der frühen Kindheit bis in das Grundschulalter	28
Tabelle 4:	Klassifikation von Verhaltensstörungen bei Kindern und Jugendlichen.....	34
Tabelle 5:	Störungsspezifische Merkmale von externalisierenden Verhaltensauffälligkeiten und emotional-sozialen Entwicklungsstörungen	35
Tabelle 6:	Kriterien effektiven Classroom Managements	46
Tabelle 7:	Pädagogisch-psychologisches Basiswissen sowie sonderpädagogisches Spezialwissen für den inklusiven Unterricht ..	74
Tabelle 8:	Wirkungsebenen von Lehrkraftfortbildungen	84
Tabelle 9:	Befunde zu strukturellen und inhaltlichen Merkmalen von Fortbildungsmaßnahmen im FSP ESE und zu deren Wirksamkeit	91
Tabelle 10:	Durchführung der Evaluationsstudie anhand des inhaltlichen Ablaufs des Professionalisierungskonzepts, der Professionalisierungselemente, der evaluativen Begleitung und der eingesetzten Messverfahren	127
Tabelle 11:	Untersuchungsgruppe im Vergleich zur Gesamtstichprobe anhand deskriptiver Daten und der Rohwerte der Subskalen des Kurzscreenings schulbezogenen Verhaltens .	132
Tabelle 12:	Operationalisierte Items der Multi-Item-Scales, geordnet nach deren Häufigkeit.....	134
Tabelle 13:	Auswertungsverfahren und statistische Prüfgrößen, geordnet nach den untersuchten Evaluationsebenen, den abhängigen Variablen und den eingesetzten Messverfahren	141

Tabelle 14:	Vorgehen bei der Auswertung der kontrollierten Einzel- fallstudie durch die Kombination deskriptiver Interpreta- tionsmaße, Effektstärkemaße und inferenzstatistischer Interpretations- und Prüfgrößen der Stückweisen Regressionsanalyse.....	143
Tabelle 15:	Deskriptive Ergebnisse der spezifischen Verhaltensauffälligkeiten für jedes Kind und über alle Schüler*innen	154
Tabelle 16:	Überlappungsmaße hinsichtlich der spezifischen Verhaltensauffälligkeiten für jedes einzelne Kind und gemittelt über alle Schüler*innen	156
Tabelle 17:	Stückweise Regressionsanalyse hinsichtlich der spezifischen Verhaltensauffälligkeiten für jedes Kind	157
Tabelle 18:	Mehrebenenanalyse hinsichtlich der spezifischen Verhaltensauffälligkeiten über alle Schüler*innen.....	158
Tabelle 19:	Deskriptive Ergebnisse des Regelverhaltens für jedes Kind	160
Tabelle 20:	Überlappungsmaße hinsichtlich des Regelverhaltens für jedes Kind und über alle Schüler*innen	162
Tabelle 21:	Stückweise Regressionsanalyse hinsichtlich des Regelverhaltens für jedes Kind.....	163
Tabelle 22:	Mehrebenenanalyse hinsichtlich des Regelverhaltens über alle Schüler*innen	164
Tabelle 23:	Mehrebenenanalyse hinsichtlich der spezifischen Verhaltensauffälligkeiten mit lerntheoretischer Intervention	165
Tabelle 24:	Mehrebenenanalyse hinsichtlich der spezifischen Verhaltensauffälligkeiten bei kombinierter Intervention	166
Tabelle 25:	Mehrebenenanalyse hinsichtlich des Regelverhaltens mit lerntheoretischer Intervention	167
Tabelle 26:	Mehrebenenanalyse hinsichtlich des Regelverhaltens bei kombinierter Intervention.....	167
Tabelle 27:	Deskriptive Statistik und Ergebnisse des Mittelwert- vergleichs über drei MZP hinsichtlich der Verhaltensprobleme der Schüler*innen.....	169

Tabelle 28:	Deskriptive Statistik und Ergebnisse des Mittelwert- vergleichs über drei MZP hinsichtlich des prosozialen Verhaltens der Schüler*innen, erhoben mit dem Strengths and Difficulties Questionnaire.....	170
Tabelle 29:	Interaktionseffekte zwischen der Einsatzhäufigkeit der <i>84 Handlungsmöglichkeiten</i> und den spezifischen Verhaltensauffälligkeiten aller Schüler*innen anhand einer Mehrebenenanalyse	175
Tabelle 30:	Interaktionseffekte zwischen der Einsatzhäufigkeit der <i>84 Handlungsmöglichkeiten</i> und dem Regelverhalten aller Schüler*innen anhand einer Mehrebenenanalyse.....	176
Tabelle 31:	Ergebnisse der Angaben zum Selbstwirksamkeits- empfinden der Lehrkräfte vor und nach der Intervention sowie mit und ohne Beratung	178
Tabelle 32:	Interaktionseffekte zwischen dem Selbstwirksamkeits- empfinden der Lehrkräfte und den spezifischen Verhaltensauffälligkeiten aller Schüler*innen anhand einer Mehrebenenanalyse	179
Tabelle 33:	Interaktionseffekte zwischen dem Selbstwirksamkeits- empfinden der Lehrkräfte und dem Regelverhalten aller Schüler*innen anhand einer Mehrebenenanalyse.....	179
Tabelle 34:	Gemittelte und kategorisierte Lehrkräfteeinschätzungen bezüglich der Professionalisierungsmaßnahme nach deren Beendigung.....	183
Tabelle 35:	Gemittelte Lehrkräfteeinschätzung bezüglich der durchlaufenen Fortbildungen und des eingesetzten Verhaltensmonitorings nach Beendigung der Professionalisierungsmaßnahme	186
Tabelle 36:	Gemittelte Lehrkräfteeinschätzung bezüglich der erhaltenen systemisch-kooperativen Beratungssitzung nach Beendigung der Professionalisierungsmaßnahme.....	187
Tabelle 37:	Zusammenfassende Darstellung der Ergebnisse der Evaluationsstudie zum Professionalisierungskonzept, geordnet nach den Wirkungsebenen von Lehrer*innenfortbildungen, den abhängigen Variablen/ Untersuchungsschwerpunkten und den Analyseverfahren ...	190

1 Einleitung

Die Arbeit von Lehrkräften ist seit jeher kontinuierlichen und vielfältigen Veränderungsprozessen unterworfen (Herzog & Makarova, 2012). Um Schüler*innen auf sich wandelnde Ansprüche in Gesellschaft und Beruf vorzubereiten und gleichzeitig die sich stetig verändernden Anforderungen an die eigene berufliche Tätigkeit erfolgreich bewältigen zu können, ist berufliches Lernen von Lehrkräften unabdingbar. Im besonderen Maße trifft dieser Umstand auf die sich seit der Ratifizierung des Übereinkommens der Vereinten Nationen über die Rechte von Menschen mit Behinderungen (UN-Behindertenrechtskonvention [UN-BRK], 2009) beschleunigenden Anstrengungen hin zu einer inklusiven Gesellschaft mit einem dementsprechenden Schulsystem zu. Der Auftrag inklusiver Bildung umfasst die wirksame Unterstützung aller Lernenden gemäß ihren individuellen Bedürfnissen, insbesondere auch derer mit Beeinträchtigungen: „States Parties shall ensure that persons with disabilities receive the support required, within the general education system, to facilitate their effective education“ (United Nations 2006, Art. 24 Nr. 2 d). Obwohl schulische Inklusion in Deutschland in den Bundesländern sehr unterschiedlich umgesetzt wird, ist ein Grundkonsens erkennbar: Die größtmögliche gemeinsame Beschulung von Kindern mit und ohne sonderpädagogischen Förderbedarf im *gemeinsamen Unterricht (GU)* der Regelschule und damit die Erhöhung der sogenannten *Inklusionsquote* (Klemm, 2018). In einigen Bundesländern führt/e dies zur schrittweisen Schließung bestimmter Förderschularten. In Mecklenburg-Vorpommern, wo das vorliegende Professionalisierungskonzept durchgeführt und evaluiert wurde, wird die flächendeckende inklusive Beschulung von Schüler*innen in den *Förderschwerpunkten Lernen und Sprache* sowie maßgeblich auch im *Förderschwerpunkt Emotional-soziale Entwicklung (FSP ESE)* bis zum Schuljahr 2027/2028 angestrebt (Ministerium für Bildung und Kindertagesförderung, 2022a). Kinder, die Verhaltensauffälligkeiten und emotional-soziale Entwicklungsstörungen in der Schule zeigen und damit im weitesten Sinne dem *FSP ESE* zugeordnet werden können, wurden aber auch schon vor den bildungspolitischen Bemühungen im Rahmen der Umsetzung der UN-BRK größtenteils an Regelschulen unterrichtet. Stein und Müller (2015, S. 39f.) konstatieren, dass Inklusion für diese Kinder vielmehr hieße, „sie überhaupt wahrzunehmen, mit sonderpädagogischer Professionalität zu begleiten, Regelschullehrkräfte hinsichtlich des Umgangs mit ihnen zu beraten und eine Exklusion aus dem Regelschulwesen zu verhindern“.

Durch die mit den Inklusionsbestrebungen einhergehende wachsende Heterogenität der Schülerschaft sowie die sich verändernden Lernvoraussetzungen und Bedürfnisse der Lernenden entsteht die Notwendigkeit einer qualitativen Weiterentwicklung des Unterrichts und des schulischen Angebots insgesamt. Um Schüler*innen in den einzelnen Förderschwerpunkten anhand pädagogischer Prävention und Intervention effektiv und nachhaltig fördern zu können, benötigen Lehrkräfte spezifische professionelle Kompetenzen. Das sich neuformierende „inklusive“ Professionalisierungsprofil, welches mehr und mehr in Studien- und Prüfungsordnungen der ersten Phase der Lehrer*innenbildung abgebildet ist (u.a. Prüfungs- und Studienordnung für den Studiengang Grundschullehramt an der Universität Greifswald 2020, S. 5; Studiengangsspezifische Prüfungs- und Studienordnung für den Studiengang Lehramt an Grundschulen der Universität Rostock 2021, S. 21), weisen im Schuldienst befindliche Regelschullehrkräfte meist (noch) nicht auf, weil sie dafür nicht ausgebildet wurden (Melzer et al., 2015). Eine geeignete Möglichkeit, diese Lehrkräfte für den inklusiven Unterricht (weiter) zu professionalisieren, sind berufsbegleitende Fortbildungen (Darling-Hammond et al., 2017; Lipowsky & Rzejak, 2012; Timperley et al., 2007).

Im Vergleich zu anderen Förderschwerpunkten fordert die inklusive Beschulung im FSP ESE die Lehrkräfte und Schulen in besonderem Maße heraus – national (Ahrbeck, 2011; Ellinger & Stein, 2012) wie international (Dyson, 2010; Van Mieghem et al., 2020). Nachweislich nehmen Regelschullehrkräfte Schüler*innen mit auffälligem, externalisierendem Verhalten vielfach als sehr belastend wahr und zeigen negative inklusionsbezogene Einstellungen hinsichtlich dieser Schüler*innen (Avramidis et al., 2000; Urbutt, 2001; Van Mieghem et al., 2020). Im Umgang mit herausfordernden Verhaltensweisen benötigen Lehrkräfte somit qualitativ hochwertige Fortbildungsangebote, in deren Rahmen sie zum einem lernen, welche Bedürfnisse Kinder mit Verhaltensauffälligkeiten haben, welche Gründe es für ihr Verhalten geben kann und mit welchen pädagogischen Maßnahmen sie diese Schüler*innen fördern können. Neben deklarativem und prozeduralem Wissen um Fördermöglichkeiten sollten Lehrkräfte zum anderen ein Handlungsrepertoire erlernen und beherrschen, welches sie in konkreten schulischen Situationen kompetent und situativ einsetzen können. Dies ist umso relevanter als der FSP ESE mit ca. 1.3 % den zweitgrößten Förderschwerpunkt in Deutschland darstellt (Kultusministerkonferenz 2018, S. XVI), wobei der Anteil förderbedürftiger Kinder und Jugendlicher ohne offizielle Feststellung des Förderbedarfs noch um ein Vielfaches größer sein dürfte. So kann man davon ausgehen, dass in Schulklassen mit einer Klassenstärke von 22 Schüler*innen im Durchschnitt mindestens zwei Kinder Auffälligkeiten im Verhalten zeigen (Methner et al. 2017, S. 20). Die angemessene schulische Förderung der ESE dieser Kinder stellt daher eine besondere pädagogische Herausforderung dar und an der effektiven Integration dieser Schüler*innengruppe „entscheidet

sich womöglich letztlich auch, wieweit [sic] die Vision der inklusiven Schule trägt“ (Willmann 2015, S. 63).

Hier setzt die vorliegende Arbeit an. Anknüpfend an die Notwendigkeit einer qualitativ hochwertigen Förderung der Schüler*innen im Bereich ESE im gemeinsamen Unterricht an Regelschulen und das beschriebene Professionalisierungsdefizit der im Schuldienst befindlichen Lehrkräfte wurde ein Professionalisierungskonzept für Grundschullehrkräfte entworfen und im Schuljahr 2018/2019 an Grundschulen in M-V eingeführt sowie begleitend evaluiert. Anhand des Forschungsstandes zur schulischen Förderung der ESE von Kindern sowie zur Professionalisierung von Lehrkräften wurden zunächst Qualitätsmerkmale abgeleitet, die dann in eine komplexe Konzeption einfließen. Die Arbeit ist in vier große Themenfelder unterteilt. Der erste Teil befasst sich mit dem Forschungsstand zur schulischen Förderung der ESE von Kindern sowie mit der dafür notwendigen Professionalisierung von Lehrkräften (Kap. 2). Anhand der theoretischen Einordnung und der wissenschaftlichen Fundierung wird dann im zweiten Teil ein Professionalisierungskonzept für Grundschullehrkräfte zur Förderung der ESE ihrer Schüler*innen abgeleitet und umfassend erläutert (Kap. 3). Im dritten Teil der Arbeit wird die Evaluationsstudie vorgestellt, in deren Rahmen dieses Professionalisierungskonzept wissenschaftlich untersucht worden ist (Kap. 4). Nach der Vorstellung der Ergebnisse (Kap. 5) und deren Diskussion (Kap. 6) werden abschließend Schlussfolgerungen für Forschung und Praxis gezogen (Kap. 7).

2 Forschungsstand

2.1 Schulische Förderung der ESE von Kindern

Um ein wirksames Professionalisierungskonzept für Lehrkräfte im Bereich ESE zu entwerfen, ist es zunächst notwendig, die Zielgruppe der Kinder mit Verhaltensauffälligkeiten und emotional-sozialen Entwicklungsstörungen genauer zu beschreiben sowie Möglichkeiten der schulischen Förderung darzulegen. Anhand der zentralen Komponenten der ESE von Kindern wird zunächst die Genese emotional-sozialer Kompetenzen dargestellt. Darauf aufbauend werden atypische Verläufe umfassender erläutert. Abschließend werden schulische Fördermöglichkeiten und -konzeptionen im Bereich ESE vorgestellt und Schlussfolgerungen für die Professionalisierung von Lehrkräften gezogen.

2.1.1 Zentrale Komponenten der ESE

Im Zuge der ESE bilden sich sowohl emotionale als auch soziale Kompetenzen heraus, deren Vorhandensein Menschen zu sozial kompetentem Handeln befähigen. Beide Kompetenzbereiche sind dabei in vielfältiger Weise miteinander verknüpft und bedingen sich gegenseitig (Petermann & Wiedebusch, 2008), wobei man heute davon ausgeht, dass die emotionale Kompetenz eine Teilmenge der sozialen Kompetenz auf der Ebene der Fertigkeiten darstellt (Hennemann et al. 2017, S. 21).

Unter *emotionaler Kompetenz* versteht man den adäquaten Umgang mit Emotionen. Sie setzt sich aus den Komponenten Emotionsausdruck (mimisch und sprachlich), Emotionsverständnis und Emotionsregulation zusammen und lässt sich in weitere Unterkomponenten und Schlüsselfertigkeiten gliedern (Denham, 1998). Das Konzept des Emotionswissens (Pons et al., 2004) greift konkrete Schlüsselfertigkeiten emotionaler Kompetenz nach Saarni (1999) auf und bringt diese in eine entwicklungschronologische Reihenfolge, wobei die einzelnen Fertigkeiten nicht streng aufeinander aufbauen müssen, wie in Tabelle 1 dargestellt wird, sondern sich auch parallel zueinander entwickeln können. Janke (2008) konnte im deutschsprachigen Raum nachweisen, dass sich die ersten Schlüsselfertigkeiten bereits im Kindergartenalter und die letzten erst in der späten Grundschulzeit entwickeln.

Tabelle 1: Das schrittweise Erwerben von Schlüsselertigkeiten zur Entwicklung von Emotionswissen nach Pons et al. (2004)

Emotionswissen	
Mimik	Erkennen und Benennen verschiedener Emotionsausdrücke anhand der Mimik
Anlässe	Zuordnung einer Emotion zu einem konkreten Anlass/ einer bestimmten Situation
Wünsche	Erwartungen und Wünsche werden als Emotionsauslöser verstanden
Emotionsperspektive	Fähigkeit zur emotionalen Perspektivübernahme der Gefühle anderer (Theory of Mind)
Erinnerung	Verständnis dafür, dass abgespeicherte Erinnerungen erneut Emotionen auslösen können
Regulation	Fähigkeit, das eigene emotionale Befinden zu beeinflussen
Verbergen	Fähigkeit, die eigenen Emotionen zu verbergen
Multiple Emotionen	Verständnis dafür, dass man mehrere Emotionen gleichzeitig haben kann
Gewissen	Verständnis für Emotionen, die sich aus dem sozialen Kontext ergeben

Die Basis für die Erklärung emotionaler Kompetenz stellen die Emotionen selbst dar. Unter Emotionen versteht man die „zeitlich begrenzte Konfiguration von Körperreaktionen (z.B. Herzschlagänderung), Ausdrucksformen und subjektivem Gefühlszustand“ (Holodynski et al. 2013, S. 196). Diese haben somit objektive und subjektive Komponenten und lassen sich in Basis- und Sekundäremotionen unterteilen, deren primäre Funktion die Handlungsanleitung ist. Sie stellen sicher, dass „die eigenen Handlungsziele, -ergebnisse und -folgen und der situative [...] Kontext in ihrer Beziehung zum Grad der Befriedigung der eigenen Motive“ gegeben ist (Holodynski 2006, S. 41). Basis- oder Grundemotionen wie Freude, Ärger/Wut, Trauer und Angst sind universell und kulturübergreifend bei allen Menschen vertreten. Bereits Neugeborene verfügen über sogenannte „Vorläuferemotionen“, die sich anhand interpersonaler Regulation durch Bezugspersonen während der ersten sechs

Lebensmonate zu den vier Basisemotionen entwickeln (Holodynski & Friedlmeier 2006, S. 120). Sekundäremotionen wie Schuldbewusstsein, Scham, Verlegenheit, Neid, Stolz und Empathie sind selbstbezogen. Sie bilden sich etwa zwischen dem 18.-24. Lebensmonat durch das Erkennen der eigenen Person als körperlich eigenständiges Wesen heraus und treten dann innerhalb sozialer Interaktionen auf. Im Laufe der Kleinkind- und Vorschulzeit lernen Kinder, dass sie diese selbst aktiv lenken können. Große Schritte zur Regulierung und zum Verbergen von Emotionen vollziehen sie im Alter von 6-7 Jahren. Mit dem Übergang vom Kindergarten zur Grundschule erhöht sich die soziale Komplexität und entsprechend auch die Bandbreite sozialer Emotionen. Der Erwerb von Kompetenzen, die dazu ermächtigen, gemischte Emotionen zu verstehen und Sekundäremotionen in die soziale Interaktion mit einzubeziehen, erstreckt sich bis ins Jugendalter (z.B. Janke, 2007, 2008). Nach Holodynski (2006) kann die emotionale Entwicklung in die drei Phasen Vorläuferemotionen, Entstehung von funktionstüchtigen Emotionen und Entstehung der intrapersonalen Regulation (selbstständige Regulation von Emotionen und Handlungen) eingeteilt werden. Dabei verschiebt sich die Gewichtung von der interpsychischen, bei der die Umwelt die Emotion des Individuums noch stark beeinflusst, hin zur intrapsychischen Regulation, bei der das Kind zunehmend selbstständig seine Emotionen steuert. Entsprechende Fähigkeiten basieren auf einem Repertoire an effektiven Strategien wie der Nutzung von Sprache als Mittel der psychologischen Distanzierung, der Anwendung exekutiver Funktionen auf die eigenen Emotionen und dem Aufschub von Befriedigung ausgelöster Emotionen durch Zukunftsperspektiven (Holodynski et al., 2013). Da die Emotionsregulation zentral für die ESE von Kindern ist, soll dieses im Folgenden noch eingehender erläutert werden.

Emotionen kommen selten bis gar nicht in unreguliertem Zustand vor. Einerseits leiten sie Handlungen an und regulieren diese, andererseits regulieren Handlungen Emotionen und verstärken oder mindern deren Intensität, bevor sie überhaupt ausgebildet werden (Holodynski et al., 2013). Das Generieren von Emotionen und deren Regulierung sind demnach zwei nicht voneinander trennbare Vorgänge (Campos et al., 2004; Kullik & Petermann, 2012). Es können drei Strategiebereiche unterschieden werden: reflexive Emotionsregulationsstrategien, volitionale sowie habituelle Regulationen von Handlungen (Seitz-Stein et al. 2012, S. 198f.). Die Entwicklung vollzieht sich dabei von interpersonaler Regulation mithilfe von Bezugspersonen zu einer intrapersonalen, die das Kind eigenständig umsetzen kann. Physische Strategien entwickeln sich immer mehr hin zu mentalen (Holodynski & Friedlmeier 2006, S. 136), wie es in Tabelle 2 anhand von Beispielen illustriert wird.

Tabelle 2: Beispiele für Strategien der reflexiven Emotionsregulation und deren Initiatoren (Holodynski & Friedlmeier 2006, S. 136)

Initiator der Emotionsregulation			
Art der Strategie	Bezugsperson initiiert interpersonale Regulation für das Kind	Kind initiiert interpersonale Regulation durch Bezugsperson	Kind initiiert intrapersonale Regulation
Verhaltensstrategie (physisch)			
Beruhigung	Wiegen, Streicheln, Körperkontakt	Kind sucht Nähe der Bezugsperson, um sich beruhigen zu lassen	Kind beruhigt sich selbst durch Handlungen wie Saugen, Streicheln
Symbolische Strategien (mental)			
Beruhigung, Trost	Verbaler Trost und Beruhigung	Verbale Aufforderung an Bezugsperson, zu trösten	Kind beruhigt sich durch (positive) Selbstinstruktion
Antezedente Strategien			
Diskurs über Regulation von Emotionen	Bezugsperson spricht mit Kind über Emotionen und deren Regulation	Kind fragt Bezugsperson über Emotionen und deren Regulation	Gedankliche Auseinandersetzung über potentielle Emotionsepisoden

Reflexive Emotionsregulationsstrategien sind mentale und kognitive Prozesse zur Steuerungs- und Kontrollfunktion von Verhalten und Aufmerksamkeit, etwa dem Aufschub der Motivbefriedigung oder der Anpassung des Emotionsausdruckes an kulturelle Darbietungsformen (Holodynski et al., 2013). Bei der volitionalen Regulation von Handlungen werden Ziele antizipiert und die dafür benötigten Handlungen geplant. Reflexive und volitionale Strategien sind sich sehr ähnlich. Bei letzterer gelingen Planung und Antizipation

jedoch bereits so vorausschauend, dass eine zielhinderliche emotionale Reaktion verhindert werden kann. Die folgenden dafür notwendigen Fähigkeiten bilden sich in Grundzügen zwischen dem dritten und sechsten Lebensjahr aus:

- ein Repertoire an Regulationsstrategien wie z.B. Ablenkung, Umdeutung, Änderung der situativen Bedingungen, usw.,
- die Fähigkeit, Sprache als Mittel der psychologischen Distanzierung nutzen zu können, um Anlässe, Ausdrucks- und Erlebensmuster sowie Konsequenzen von Emotionen und deren Regulation mit Hilfe von Regulationsstrategien repräsentieren und mit diesem Wissen willentlich operieren zu können,
- die Fähigkeit, mental auf Zeitreise zu gehen, sich zukünftig motivbefriedigende Situationen willentlich vorzustellen,
- die Fähigkeit, exekutive Funktionen auf die eigenen Emotionen anwenden zu können, d.h., einen Handlungsimpuls willentlich zu hemmen, um eine subdominante, aber zielführende Handlung auszuführen (Holodynski et al. 2013, S.199f.).

Habituelle Regulationen von Handlungen sind schließlich erlernte Handlungs-Ziel-Verknüpfungen, die als Schema zur Verfügung stehen und eine automatisierte Regulation ermöglichen.

Wie bereits herausgestellt wurde, stellen Emotionen und das Erlernen von Regulationsstrategien wesentliche Bedingungsfaktoren für die Entwicklung emotional-sozialer Kompetenzen dar. Deren Herausbildung wird während der Kindheit und Jugend durch spezifische Entwicklungsaufgaben und die dabei auftretenden Herausforderungen, Chancen und Risiken beeinflusst (Hennemann et al. 2017, S. 30ff.). Grundsätzlich kann festgehalten werden: Übergänge spielen bei der Entwicklung der ESE eine wesentliche Rolle, da sie immer „Prozesse mit offenem Ausgang“ (Hennemann et al. 2017, S. 30) sind. Befunde der Entwicklungspsychologie (Masten & Powell, 2003) stellen für bestimmte Entwicklungsspannen und damit einhergehende Transitionen vordergründige Entwicklungsaufgaben heraus. Da im Rahmen dieser Arbeit die ESE bis in die Grundschulzeit relevant ist, werden die Entwicklungsaufgaben von der frühen Kindheit bis ins Grundschulalter im Folgenden eingehender dargestellt (Tabelle 3).

Anhand der Übersicht wird deutlich, dass Kinder in den ersten drei Entwicklungsspannen neben emotionalen auch zunehmend soziale Kompetenzen entwickeln.

Soziale Kompetenz beschreibt die Fähigkeit, in sozialen Situationen persönliche Ziele zu erreichen und gleichzeitig positive Beziehungen zu anderen aufrecht zu erhalten (Siegler et al., 2008). Sie ist die „Gesamtheit des Wissens, der Fähigkeiten und Fertigkeiten einer Person, welche die Qualität eige-

Tabelle 3: Entwicklungsspannen mit Entwicklungsaufgaben von der frühen Kindheit bis in das Grundschulalter (Berk 2011, S. 246ff.; Koglin & Petermann 2008, S. 82; Vonderlin & Pauen 2013, S. 83)

Entwicklungsspanne	Entwicklungsaufgaben
0-3 Jahre	– Etablierung der Grundlagen für die Selbstregulation der körperlichen Prozesse, der Aufmerksamkeit und der eigenen Impulse – Annäherung an und Rückzug von emotionsauslösende/n Reize/n – initialisiert emotionale Entwicklung und Ausbildung des prä-frontalen Kortex – Entwicklung v.a. durch Mutter-Kind-Beziehung
Kindergartenalter	– Selbstständigkeit in Alltagsanforderungen – Selbstregulation und Frustrationstoleranz – Fantasie- und Spieltätigkeit – Normen, Grenzen und Regeln akzeptieren – soziale Integration in die Gleichaltrigengruppe
Grundschulalter	– angemessenes Verhalten in der Schule zeigen – allgemeine Verhaltensregeln zu Hause, in der Schule und in der Öffentlichkeit befolgen – mit Gleichaltrigen in der Schule zurechtkommen und Freundschaften schließen

nen Sozialverhaltens – im Sinne sozial kompetenten Verhaltens – fördert“ (Kanning 2009, S. 15). Sozial kompetentes Verhalten ist entsprechend die sichtbare Manifestation sozialer Kompetenz in einer konkreten Situation und vollzieht sich in der Wechselwirkung zwischen emotional-sozialen Kompetenzen, etwa der Fähigkeit angemessene Gefühle entwickeln zu können, kognitiven sozialen Kompetenzen, z.B. die gute Einschätzung und/oder Antizipation der für die Interaktion relevanten Faktoren, sowie aktionalen sozialen Kompetenzen, etwa anhand eines vorhandenen Repertoires an verbalen und nonverbalen Verhaltensfertigkeiten (Döpfner, 1989).

Zusammenfassend kann festgehalten werden: Emotionsregulationsstrategien bilden die Basis für kompetentes soziales Handeln und tragen zur psychischen Gesundheit bei. Die hierfür bedeutsamen Kompetenzen erlernen Kinder in frühen sozialen Interaktionserfahrungen. Bestenfalls entsteht dabei ein „Engelskreis“ (Castello 2017, S. 101): Kinder entwickeln durch positive Bindungserfahrungen prosoziales Verhalten, können dadurch später stabilere

Freundschaften schließen und so wiederum ihre sozialen Kompetenzen trainieren. Ist die Emotionsregulation hingegen ungünstig ausgeprägt, so stellt sie einen Risikofaktor für die Entstehung von Verhaltensauffälligkeiten und emotional-sozialen Entwicklungsstörungen bei Kindern dar (Koglin et al., 2013). In einem „Teufelskreis“ sammeln Kinder mit einem schwer regulierbaren Temperament weniger positive soziale Erfahrungen, entwickeln unangemessene Erwartungen an soziale Beziehungen sowie ungünstige emotional-soziale Strategien, die dann wiederum zur sozialen Ablehnung führen (Castello 2017, S. 101). Eine ausgeprägte Fähigkeit zur Emotionsregulation ist demnach ein wichtiger Schutzfaktor für die emotional-soziale Entwicklung (Lösel & Bender, 2008). Diese zu fördern ist eine der wesentlichsten Entwicklungsaufgaben im Bereich ESE (Koglin et al., 2013; Petermann & Barnow, 2013).

2.1.2 Verhaltensauffälligkeiten und emotional-soziale Entwicklungsstörungen

..., böse, sittlich verwildert, moralisch schwachsinnig, entartet, psychopathisch, erziehungsschwierig, gemeinschaftsschwierig, fehlangepasst, führungsrésistent, verhaltensbehindert, verhaltensgestört, verhaltensauffällig, ...

Im Laufe der Zeit wurden und werden sogenannte *schwierige* Kinder auf unterschiedlichste Weise gekennzeichnet und beschrieben (Hillenbrand 2008, S. 29; Myschker & Stein 2014, S. 46; Willmann 2015, S. 47ff.). Anhand dieser – nur kleinen – Auswahl an Bezeichnungen wird deutlich: Die Begriffe und Deutungsmuster haben sich über die Zeit verändert und können nicht losgelöst von den allgemeinen gesellschaftlichen Entwicklungen und den jeweiligen zeitgenössischen Vorstellungen über Erziehung und Kindheit betrachtet werden (Willmann 2015, S. 46). Der Gegenstandsbereich und seine sprachlichen Mittel setzen vielmehr ein soziales Bezugssystem voraus, in dem Normalitätsvorstellungen sowie die Bewertung und der Umgang mit Abweichung (Devianz) sich stetig verändern, was eine exakte begriffliche Operationalisierung erschwert (Willmann, 2015). Denn: Die Abgrenzung zwischen dem, was wir als normales Erleben und Verhalten einerseits und als Störung andererseits bezeichnen, ist durchaus fließend (Lenhard 2021, S. 18).

Verschiedene Disziplinen wie Pädagogik, Psychologie, Medizin, Rechtswissenschaft oder Soziologie nahmen sich schrittweise der Thematik an und erforsch(t)en diese mit ihren fachspezifischen Methoden und Zugängen. Als Folge haben sich differente Kategorien- und Klassifikationssysteme herausgebildet. Unterschiedlich ausgerichtete Erklärungsansätze wie tiefenpsychologische, behaviorale, konstruktivistische Annahmen zur Erklärung der Phänomene und Symptome führen ebenfalls zu unterschiedlichen Begriffsbildungen und Definitionen (Hillenbrand, 2008; Myschker & Stein, 2014). Gemein

ist den Begriffen im Feld: Sie zielen auf die Beschreibung emotionaler und verhaltensbezogener Phänomene ab, die meist eine eher negative Bedeutungsgestalt haben. Heute besteht weitgehende Einigkeit in der Forderung, „die Begriffe nicht als Wertung einer Person, seiner Charaktereigenschaften und Persönlichkeit zu verstehen, sondern als zusammenfassende Kennzeichnung von Problemlagen, die sich in Verhaltensweisen ausdrücken können“ (Hillenbrand 2008, S. 29).

2.1.2.1 Definition

Begriffe im Feld der Verhaltensauffälligkeiten und emotional-sozialen Entwicklungsstörungen ziehen unterschiedliche Phänomene zu einem Kontraktionsbegriff (Hillenbrand, 2008) zusammen. In der Fachdiskussion in Deutschland können vor allem drei Begriffsvarianten herausgestellt werden: *emotional-sozialer Förderbedarf* (schulischer Bereich), *Verhaltensstörung* (klinischer Bereich) und *seelische Behinderung* (Jugendhilfe). Da im Kontext von schulischer Inklusion und damit (sonder-)pädagogischer Förderung schulische und klinisch-therapeutische Perspektiven relevant sind und nicht getrennt voneinander betrachtet werden können (Willmann 2015, S. 45), werden beide Perspektiven bei der nachfolgenden definitorischen Verortung berücksichtigt.

Eine weitreichende Definition von Verhaltensstörungen als phänomenologischem Oberbegriff zur Kennzeichnung einer Vielzahl unterschiedlicher Erscheinungsformen aus einer eher klinisch-therapeutischen Sicht, die bis heute breite Verwendung findet, stammt von Myschker. Hiernach ist eine Verhaltensstörung

ein von den zeit- und kulturspezifischen Erwartungen abweichendes maladaptives Verhalten, das organogen und/oder milieureaktiv bedingt ist, wegen der Mehrdimensionalität, der Häufigkeit und des Schweregrades die Entwicklungs-, Lern- und Arbeitsfähigkeit sowie das Interaktionsgeschehen in der Umwelt beeinträchtigt und ohne besondere pädagogisch-therapeutische Hilfe nicht oder nur unzureichend überwunden werden kann (Myschker 2009, S. 49).

International ist der Begriff der *Gefühls- und Verhaltensstörung* weit verbreitet, der vom US-amerikanischen Fachverband Council for Children with Behavior Disorders (CCBD) eingeführt wurde und die Möglichkeit bietet, sich interdisziplinär besser zu verständigen:

Der Begriff der Gefühls- und Verhaltensstörungen beschreibt eine Beeinträchtigung (disability), die in der Schule als emotionale Reaktionen und Verhalten wahrgenommen werden und sich von altersangemessenen, kulturellen oder ethnischen Normen so weit unterscheiden, dass sie auf die Erziehungserfolge des Kindes oder Jugendlichen einen negativen Einfluss haben. Erziehungserfolge umfassen schulische Leistungen, soziale, berufsqualifizierende und persönliche Fähigkeiten. Eine solche Beeinträchtigung ist

- mehr als eine zeitlich begrenzte, erwartbare Reaktion auf Stresseinflüsse in der Lebensumgebung;
- tritt über einen längeren Zeitraum in zwei verschiedenen Verhaltensbereichen (settings) auf, wobei mindestens einer dieser Bereiche schulbezogen ist;
- ist durch direkte Intervention im Rahmen allgemeiner Erziehungsmaßnahmen insofern nicht aufhebbar, als diese Interventionen bereits erfolglos waren oder erfolglos sein würden;

Gefühls- und Verhaltensstörungen können im Zusammenhang mit anderen Behinderungen auftreten und erfordern für ihre Beschreibung Informationen aus verschiedenen Quellen und Messverfahren (Opp 2003, S. 509f.).

Wegen der starken Betonung von Normabweichungen steht der Terminus *Störung* jedoch immer wieder in der Kritik. So würden dabei Umfeld- und Risikofaktoren vernachlässigt werden sowie der Fokus zu stark auf die Kinder und Jugendlichen und dabei eher interventiv als präventiv ausgerichtet sein (u.a. Stein & Müller, 2015). Andere Begrifflichkeiten wie etwa *Verhaltensoriginalität* haben sich bis heute jedoch nicht durchgesetzt.

Im schulischen Bereich hat sich seit den KMK-Empfehlungen 1994 und 2000 vor allem die Bezeichnung *Förderbedarf in der emotionalen und sozialen Entwicklung* etabliert, der wie folgt definiert wird:

Sonderpädagogischer Förderbedarf ist bei Kindern und Jugendlichen mit Beeinträchtigungen der emotionalen und sozialen Entwicklung, des Erlebens und der Selbststeuerung anzunehmen, wenn sie in ihren Bildungs-, Lern- und Entwicklungsmöglichkeiten so eingeschränkt sind, dass sie im Unterricht der allgemeinen Schule auch mit Hilfe anderer Dienste nicht hinreichend gefördert werden können (Kultusministerkonferenz 2000, S. 10f.).

Auch diese Definition wurde kritisiert, da sie Verhaltensweisen als ursächlich personalisiert darstellt und dadurch eine interaktionistische Sichtweise, etwa anhand der Betrachtung der Lebensumstände und des Umfeldes, verhindert. Außerdem bleiben kognitive Kompetenzen, wie die Wahrnehmung sozialer Situationen, die Fähigkeit zum sozialen Problemlösen oder die differenzierte Reflexion von Handlungsfolgen, gänzlich unberücksichtigt (Stein & Müller, 2014). Da die vorliegende Arbeit die (präventive) schulische Förderung bei Verhaltensauffälligkeiten und emotional-sozialen Entwicklungsstörungen in den Blick nimmt, orientieren sich die Darstellungen jedoch vor allem an dem durch die KMK bestimmten schulischen FSP ESE. Auffälligkeit, Störung und Behinderung werden dabei nicht (nur) als intraindividuelle Personenmerkmale verstanden. Daneben spielen vor allem Beiträge der Situation und des Umfeldes, also ungünstige Mensch-Umwelt-Interaktionen bei der Entwicklung emotional-sozialer Entwicklungsstörungen eine tragende Rolle, die oftmals in einem komplexen, wechselseitigen Verhältnis stehen und zu beeinträchtigten Teilhabemöglichkeiten der Betroffenen führen (u.a. Gottfried et al., 2021). Ein solches interaktionistisches Verständnis ist auch anschluss-

fähig an die aufkommenden Anforderungen eines stärker inklusiven Schulsystems, weil über dieses Verständnis emotional-soziale Entwicklungsstörungen nicht erst identifiziert werden, wenn sie verfestigt sind, sondern bereits dann, wenn sie sich – etwa als Verhaltensauffälligkeiten von Schüler*innen – anbahnen (Stein & Müller, 2015).

Innerhalb der vorliegenden Ausarbeitung werden der Begriff Verhaltensauffälligkeiten und dessen Äquivalente Verhaltensprobleme, herausforderndes/schwieriges Verhalten, Auffälligkeiten im Sozialverhalten, (nicht-)regelkonformes Verhalten und Probleme im Bereich ESE als Vorform von emotional-sozialen Entwicklungs- und Verhaltensstörungen verwendet (Blumenthal et al. 2020, S. 12f.). Synonym zu Letzterem wird der Begriff des Sonderpädagogischen Förderbedarfs mit dem Schwerpunkt emotionale und soziale Entwicklung (SF ESE) genutzt, der sich vor allem auf besonders stark ausgeprägte emotional-soziale Entwicklungs- und Verhaltensstörungen bezieht, die häufig durch eine Kumulation von schwerwiegenden Entwicklungsrisiken begünstigt werden. Von Verhaltensauffälligkeiten wird gesprochen, wenn die Beeinträchtigungen von geringerer Intensität sind, eher temporär auftreten und mit pädagogischen Mitteln bewältigt werden können, ohne dass sie in einem SF ESE münden (Ahrbeck 2017, S. 2).

Herausfordernd für eine Klassifikation von Verhaltensauffälligkeiten und emotional-sozialen Entwicklungsstörungen ist der Umstand, dass es sich zwar um ein ‚hartes‘ Problemfeld mit teilweise gravierenden Ausprägungen und Folgen für die Betroffenen und ihr Umfeld handelt, dieses gleichzeitig aber auch ‚weich‘ ist, weil keine objektive Art der Schädigung, wie etwa bei einer Sinnesbeeinträchtigung, vorliegt. Vielmehr werden

unterschiedliche Bezugsmaßstäbe angelegt, die anhand von sozial konstruierten Normen und Kriterien festlegen, ob ein Verhalten als auffällig oder gar gestört einzuschätzen ist. Implizite und explizite sozio-kulturelle Normen, die in einer Gesellschaft gelten, werden dabei als Maßstab bei der Beurteilung herangezogen. Dies können Gesetze, institutionelle Regeln wie Schulordnungen, aber auch weniger regulierte Normen wie Klassenregeln oder andere Verhaltensregeln an öffentlichen Orten sein (Stein & Müller 2015, S. 29).

Trotz dieser ‚relativen‘ Beschaffenheit der Phänomene Verhaltensauffälligkeiten und emotional-soziale Entwicklungsstörungen sind je nach Disziplin und wissenschaftlicher Ausrichtung verschiedene Modelle der Klassifikation bzw. Einteilung entwickelt worden, von denen einige, die für diese Arbeit relevant sind, im Folgenden vorgestellt werden sollen.

2.1.2.2 Klassifikation und Einteilung

Um Hilfeleistungen im Rahmen der sehr großen Breite dessen, was als Verhaltensauffälligkeiten und emotional-soziale Entwicklungsstörungen bezeichnet wird, zu spezifizieren, werden die Symptome über internationale Klassifikationssysteme diagnostisch bewertet und kategorial geordnet. Dabei finden überwiegend die *International Classification of Diseases (ICD 10)* der Weltgesundheitsorganisation sowie das *Diagnostic and Statistical Manual of mental Disorders 5 (DSM V)* der American Psychiatric Association Anwendung. Bei beiden Klassifikationssystemen müssen für die Zuordnung einer Verhaltensstörung mindestens folgende drei Kriterien zutreffen:

- Intensität: längere Zeitdauer, hoher Schweregrad
- Ökologie: in mindestens zwei Settings, eines davon Schule
- Integration: Erfordernis spezifischer Hilfen für Teilhabe an der Gesellschaft

Unterschiede von Verhaltensauffälligkeiten und klinisch relevanten emotional-sozialen Entwicklungsstörungen nach ICD-10 liegen hauptsächlich in der Häufigkeit und Intensität des Auftretens der Symptome. Emotional-soziale Entwicklungsstörungen bei Kindern und Jugendlichen werden hier in folgende Subtypen eingeteilt:

- *hyperkinetische Störungen* (z.B. Aktivitäts- und Aufmerksamkeitsstörung, hyperkinetische Störung des Sozialverhaltens)
- *Störungen des Sozialverhaltens* (z.B. Störung des Sozialverhaltens in der Familie/bei fehlenden sozialen Bindungen/oppositionellem, aufsässigem Verhalten)
- *kombinierte Störung des Sozialverhaltens und der Emotionen* (z.B. Störung des Sozialverhaltens mit depressiver Störung)
- *emotionale Störungen des Kindesalters* (z.B. emotionale Störung mit Trennungsangst, phobische emotionale Störung, Störung mit sozialer Überempfindlichkeit, emotionale Störung mit Geschwisterrivalität)
- *Störungen sozialer Funktionen mit Beginn in der Kindheit und Jugend* (z.B. reaktive Bindungsstörung des Kindesalters, Bindungsstörung des Kindesalters mit Enthemmung)
- *Tic-Störungen* (z.B. vorübergehende/chronische motorische/vokale Ticstörung, kombinierte vokale und multiple motorische Tics [Tourette-Syndrom])
- *andere Verhaltens- und emotionale Störungen mit Beginn in der Kindheit und Jugend* (z.B. Enkopresis/Einkoten, Enuresis/Einnässen, Balbuties/Stottern)

Tabelle 4: Klassifikation von Verhaltensstörungen bei Kindern und Jugendlichen nach Myschker und Stein (2014, S. 58)

Verhaltensstörungen	
externalisierende Störungen	Aggression, Hyperaktivität, Aufmerksamkeitsstörung, Impulsivität
internalisierende Störungen	Angst, Minderwertigkeit, Trauer, Interessenlosigkeit, Schlafstörungen, somatische Störungen
sozial unreifes Verhalten	Konzentrationsschwäche, altersunangemessenes Verhalten, leicht ermüdbar, leistungsschwach, nicht belastbar
sozialisiert delinquentes Verhalten	Gewalttätigkeit, Reizbarkeit, Verantwortungslosigkeit, leichte Erregbarkeit und Frustration, Beziehungsstörungen, niedrige Hemmschwelle

Eine weitere, insbesondere in der Schulforschung häufig rezipierte empirische Klassifikation nehmen Myschker und Stein (2014) vor. Sie gliedern Verhaltensstörungen bei Kindern in *externalisierende* und *internalisierende Störungen*, *sozial unreifes* sowie *sozialisiert delinquentes Verhalten*. Diese sind anhand ihrer Erscheinungsbilder in Tabelle 4 dargestellt.

Empirisch sind insbesondere unter- (externalisierende) und überkontrollierte (internalisierende) Verhaltensweisen gut belegt (Myschker & Stein, 2014) und somit auch für den schulischen Bereich besonders relevant. Im Rahmen dieser Arbeit stehen externalisierende Verhaltensauffälligkeiten und emotional-soziale Entwicklungsstörungen im Fokus, die Schulen und Lehrkräfte vor besondere Herausforderungen stellen, da sie nach außen und gegen die Umwelt gerichtet sind. Diese lassen sich im Wesentlichen anhand dreier Dimensionen beschreiben: *Störungen mit oppositionellem Trotzverhalten*, *Störungen des Sozialverhaltens* und die *Aufmerksamkeitsdefizit-/Hyperaktivitätsstörung*. Die *Störung mit oppositionellem Trotzverhalten* stellt eine impulsiv-aggressive Störungsform dar. Es besteht eine geringe Frustrationstoleranz sowie verzerrte Wahrnehmung der Bedrohung oder Benachteiligung. Bei *Störungen des Sozialverhaltens* handeln die Kinder und Jugendlichen instrumentell-aggressiv. Sie möchten Macht und Dominanz herstellen, um gezielt eigene Interessen durchzusetzen, dabei sind sie furchtlos, emotional unempfindsam und unempathisch (Linderkamp, 2022). In Tabelle 5 werden die drei Varianten externalisierenden Problemverhaltens anhand ihrer störungsspezifischen Erscheinungsformen charakterisiert.

Tabelle 5: Störungsspezifische Merkmale von externalisierenden Verhaltensauffälligkeiten und emotional-sozialen Entwicklungsstörungen nach Linderkamp (2022, S. 34)

externalisierende Verhaltensauffälligkeiten und emotional-soziale Entwicklungsstörungen	
Störungen mit oppositionellem Trotzverhalten	– das Kind ist andauernd (seit wenigstens sechs Monaten), beständig und in einem mit dem Entwicklungsstand des Kindes nicht zu vereinbarenden und unangemessenen Ausmaß aufbrausend, empfindlich und uneinsichtig für eigenes Fehlverhalten – es streitet sich häufig mit Erwachsenen – es widersetzt sich häufig Anweisungen oder Regeln von Erwachsenen – es verärgert und beleidigt andere absichtlich – es bestehen bedeutsame Beeinträchtigungen in sozialen oder schulischen Funktionsbereichen
Störungen des Sozialverhaltens	– Brechen gesellschaftlicher Normen und Regeln – wiederholtes und anhaltendes Verletzen der Rechte anderer – aggressives Verhalten gegenüber Menschen und Tieren (wie Schlägereien, Tierquälerei, Einsatz von Waffen, Erzwingen sexueller Handlungen) – Brandstiftung – Zerstören des Eigentums anderer – Einbrüche – Betrügen – Stehlen – wiederholt über Nacht weg bleiben vor dem 13. Lebensjahr – Schule schwänzen – bedeutsame Beeinträchtigungen in sozialen oder schulischen Funktionsbereichen
Aufmerksamkeitsdefizit/Hyperaktivitätsstörungen	– Kardinalsymptomatiken Unaufmerksamkeit und/oder Hyperaktivität/Impulsivität – müssen bereichsübergreifend (also in Kindergarten/Schule und Elternhaus) vorliegen – müssen seit wenigstens sechs Monaten zu beobachten sein – müssen in einem verglichen mit dem Alter und Entwicklungsstand des Kindes deutlich unangemessenen Ausprägungsgrad vorliegen – verursachen einen deutlichen Leidensdruck bzw. eine manifeste Beeinträchtigung der sozialen und schulischen Funktionsfähigkeit

Für den Bereich der Pädagogik liegen keine strengen Klassifikationssysteme vor (Stein & Müller, 2015). Bedenkt man den präventiven Förderauftrag von Schule und „das pädagogische Anliegen, Kinder und Jugendliche als einzigartige Individuen anzusehen und sie entsprechend zu begleiten“, argumentieren Stein und Müller (2015, S. 13), so „ist zumindest fraglich, ob es dieser bedarf“. Die Autoren schlagen stattdessen den Begriff *Einteilung* vor. Hierbei wären etwa die Einteilung nach dem Erscheinungsbild (deskriptiv) und den Erklärungen der Problematiken oder nach den indizierten Fördermaßnahmen denkbar. Pädagogisch besonders relevant erscheint die Einteilung nach Erklärungs- und Handlungsansätzen, wie sie Blumenthal et al. (2020) herausarbeiten.

2.1.2.3 Erklärungsansätze zur Entstehung von Verhaltensauffälligkeiten und emotional-sozialen Entwicklungsstörungen

Verhaltensauffälligkeiten und emotional-soziale Entwicklungsstörungen sind multifaktoriell bedingt, denn bei der Entstehung spielen meist mehrere Ursachen eine Rolle. Rekurrierend auf ein von Speck (1990) gefordertes *bio-sozial-interaktionales Erklärungssystem*, welches Risikofaktoren, ihr mögliches Zusammenwirken sowie mögliche Symptome bzw. Symptomverbindungen aufzeigt, identifizieren Myschker und Stein (2014, S. 91) zunächst folgende beeinflussende Faktoren, die auf heranwachsende junge Menschen einwirken und je nach Ausprägung zu Problemlagen führen können:

- die individuellen Anlagen und Informationsaufnahme- und -verarbeitungsmuster
- verschiedene lernbiografisch bedingte Eigenheiten
- früh wirksam werdende Selbstbestimmungs- und Selbstorganisationstendenzen
- übergeordnete soziokulturelle Gegebenheiten
- vielfältige größere und kleinere soziale Systeme, wie Familie, Kindergarten, Schule, peer-group, etc.

Verhaltensstörungen entwickeln sich nach Myschker und Stein (2014, S. 89f.) in drei Phasen. In der Anfangsphase kommt es zu Beeinträchtigungen durch Problemkonstellationen der oben genannten Faktoren (Anlage, Umwelt, Selbstbestimmung). In der zweiten Phase formt sich das maladaptive Verhalten heraus und in der dritten oder Endphase verfestigt sich dieses anhand einer habitualisierten Störung. Erklärungsansätze zur Entstehung von Verhaltensauffälligkeiten und emotional-sozialen Entwicklungsstörungen können allerdings immer nur Hinweischarakter haben, da verschiedene Ursachen zu gleichen Erscheinungsformen führen und gleiche Ursachen sehr unterschiedliche Erscheinungsformen hervorbringen können (Myschker & Stein 2014, S. 90). Umso relevanter für die Betrachtung der Ätiologie er-

scheint es deshalb, Forschungsergebnisse und Theorien aus unterschiedlichen Disziplinen wie der Medizin, der Psychologie, der Soziologie und der (Sonder-)Pädagogik miteinzubeziehen. Mit Blick auf die Förderung der ESE in der Schule soll im Folgenden konkret auf drei ausgewählte Perspektiven näher eingegangen werden, die empirisch arbeitende Erziehungs- und Sozialwissenschaftler*innen (u.a. Blumenthal et al., 2020; Hennemann et al., 2017) als besonders relevant für die Entstehung von Verhaltensauffälligkeiten und emotional-sozialen Entwicklungsstörungen herausstellen: lerntheoretische, kognitionspsychologische und entwicklungspsychologische Erklärungsansätze.

Lerntheoretische Perspektive

Grundlage der Lern- und Verhaltenstheorie ist die Annahme, dass Verhalten durch situative Reize erlern- und veränderbar ist und dabei von genetischer Anlage, Umweltfaktoren und kognitiven Prozessen beeinflusst wird. Danach führen Lernvorgänge in allen Altersphasen hauptsächlich nach den Prinzipien des *klassischen Konditionierens*, des *operanten Konditionierens* und des *Modelllernens* zum Aufbau und zur Modifikation von Verhaltensweisen (Hillenbrand 2008, S. 72).

Das *klassische Konditionieren* oder *Signallernen* nach Pawlow¹ gilt dabei als grundlegende Lernform. Einer natürlichen, meist angeborenen Reaktion wird durch Lernen eine neue, bedingte Reaktion hinzugefügt. Ein bekanntes Beispiel ist der Pawlowsche Hund: Wenn ihm Futter gereicht wurde, erklang zugleich stets ein Glockenton. Nach einigen solchen Futtergaben begann schon allein nach dem bekannten Glockenton der Speichel des Hundes zu fließen – der Glockenton als zunächst neutraler Reiz wurde dadurch zu einem konditionierten. Aus lerntheoretischer Perspektive wird dem klassischen Konditionieren eine wesentliche Beteiligung an der Übernahme von Normen bzw. der Gewissensbildung zugesprochen. Aber auch Angstproblematiken können auf dem Wege des klassischen Konditionierens entstehen, indem es zu einer Verknüpfung eines traumatischen Ereignisses oder mehrerer sub-traumatischer Ereignisse mit einem neutralen Reiz kommt (Myschker & Stein 2014, S. 124).

Beim *operanten Konditionieren* nach Skinner² können Konsequenzen, die auf ein bestimmtes Verhalten folgen, Verhaltensänderungen herbeiführen. Bei der *positiven Verstärkung* oder dem *Verstärkungslernen* folgen positive und bei der *negativen Verstärkung* negative Konsequenzen auf ein bestimmtes Verhalten. Positive Verstärkung erhöht durch angenehme Konsequenzen die Verhaltensrate. Negative Verstärkung erhöht die Verhaltensrate, indem ein unangenehmer Zustand oder Reiz beendet wird (Blumenthal et al. 2020,

1 Iwan Petrowitsch Pawlow (1849-1936) war ein russischer Mediziner und Physiologe.

2 Burrhus Frederic Skinner (1904-1990) war ein US-amerikanischer Psychologe.

S. 25). Aus lerntheoretischer Sicht werden sowohl externalisierende als auch internalisierende Verhaltensprobleme durch operantes Konditionieren entwickelt und aufrechterhalten. Gerade externalisierende Problematiken wie Aggressivität entstehen „häufig über Prozesse des operanten Konditionierens, indem sie positiv verstärkt werden, beispielsweise durch die Aufmerksamkeit oder Anfeuerung von Klassenkameraden“ (Myschker & Stein 2014, S. 126).

Beobachtungslernen nach Bandura³ postuliert, dass Verhalten über *Imitationslernen* oder *Lernen am Modell* nachgeahmt und dadurch in das eigene Verhaltensrepertoire übernommen wird. Dabei werden insbesondere solche Verhaltensweisen aufmerksam beobachtet und gemerkt, die beliebte Personen zeigen, etwa Geschwister, Freunde, Eltern, Lehrkräfte, aber auch Menschen des öffentlichen Lebens wie Stars (Hillenbrand 2008, S. 72; Myschker & Stein 2014, S. 126).

Kognitionspsychologische Perspektive

Die kognitive Psychologie, die sich seit Ende der 1960er Jahren entwickelte, erklärt die psychischen Prozesse, die zwischen Reizaufnahme und dem daran anschließenden Erleben und Verhalten stattfinden (Hillenbrand 2008, S. 72). Dabei ist die stete Wechselwirkung zwischen Intuition, Reflexion und Prozessen der Handlungsplanung besonders relevant. Für die Entstehung von Verhaltensauffälligkeiten und emotional-sozialen Entwicklungsstörungen geraten damit bestimmte kognitive Prozesse in den Fokus. Neben pädagogisch bedeutsamen Theorien und Modellen wie dem *Fähigkeitsselbstkonzept*, dem Erleben von Selbstwirksamkeit in sozialen und Leistungssituationen, den Erkenntnissen der Gedächtnispsychologie und den daraus resultierenden Modellen über das Lernen als Informationsverarbeitung (Blumenthal et al. 2020, S. 27), hat sich in den letzten Jahren ein Modell als besonders relevant für die Erklärung der Entwicklung des emotional-sozialen Erlebens und Verhaltens und damit auch für etwaige Störungsformen erwiesen: das Modell zur *sozial-kognitiven Informationsverarbeitung* (SKI; Crick & Dodge, 1994). Lemerise und Arsenio (2000, S. 112ff.) betonen in einer Weiterentwicklung des SKI-Modells den Einfluss von Emotionen auf den sechs-phasigen Prozess der sozial-kognitiven Informationsverarbeitung. Während die einzelnen Phasen der Informationsverarbeitung durchlaufen werden, erfolgt der Rückgriff auf frühere soziale Erfahrungen wie erlernte Regeln, soziale Schemata und Wissen zu sozialen Situationen, deren Summe als *Database* bezeichnet wird. Diese Informationsquelle wird laufend durch neue Erfahrungen aktualisiert. Emotionsprozesse, die zwischen der Database und den konkreten Anforderungen einer sozialen Situation vermitteln, basieren auf dem genetisch be-

3 Albert Bandura (1925-2021) war ein kanadischer Psychologe.

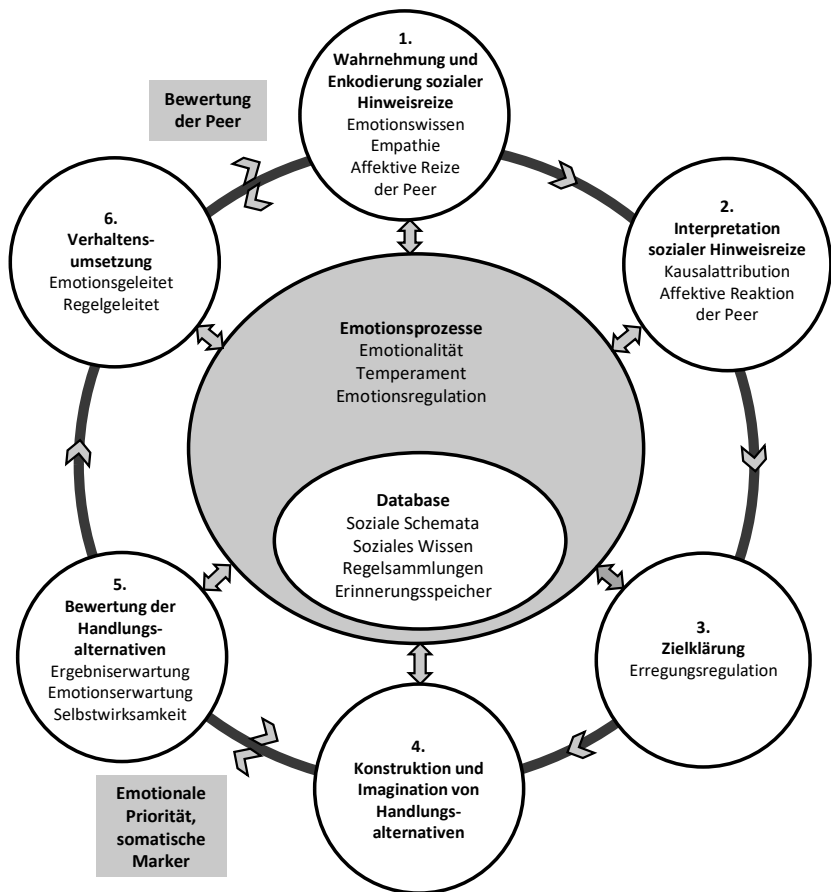


Abbildung 1: Sozial-kognitive Informationsverarbeitung (Lernerise & Arsenio, 2000; modifiziert nach Hövel, 2014; erstellt nach Hennemann et al. 2017, S. 46)

dingten Temperament, der Emotionalität und den erlernten Emotionsregulationsstrategien/-kompetenzen. Letztere entwickeln sich vorrangig innerhalb früher Erfahrungen der Eltern-Kind-Interaktion, bei der die elterliche Sensitivität, Responsivität und die Bindungsqualität eine herausragende Rolle spielen (Castello 2017, S. 104; vgl. Kap. 2.1.1). Das Modell konzipiert die sozial-kognitive Informationsverarbeitung in sechs Phasen, deren Ablauf in Abbildung 1 dargestellt wird. In der ersten Phase des SKI-Modells werden soziale Hinweisreize, u.a. auch die Wahrnehmung der peer-group, unter Einfluss der eigenen Emotionen wahrgenommen – hierbei kann der soziale Hinweisreiz

die vorhandene Stimmung verfestigen oder ändern. In der zweiten Phase werden die Zusammenhänge der wahrgenommenen Situation und die mit der Handlung der anderen Person verbundenen Intentionen interpretiert (Kausalattribution der sozialen Hinweisreize). In der dritten Phase werden die eigenen Ziele in Bezug zur eben erfolgten Interpretation der Situation festgelegt. In der vierten und fünften Phase werden mögliche Handlungsalternativen aus der Database generiert und mit der in der Vergangenheit erfahrenen Selbstwirksamkeit abgestimmt. In der sechsten Phase wird die Handlung ausgeführt. Während der Ausführung werden weitere soziale Hinweisreize aufgenommen.

Unterschiedliche Studien haben gezeigt, dass sozial kompetente Kinder und Jugendliche die Handlungsausführung während der einzelnen Phasen variieren können (Hillenbrand & Hennemann, 2006). Kinder und Jugendliche mit emotional-sozialen Entwicklungsstörungen hingegen haben bei der Verarbeitung eingehender sozialer Hinweisreize und der Bewältigung der einzelnen Schritte zum Teil große Probleme (Beelmann & Raabe, 2007; Crick & Dodge, 1994). Eine verzerrt sozial-kognitive Informationsverarbeitung gilt mittlerweile als „ein bedeutsamer, proximaler Risikofaktor für die kindliche Entwicklung, Stabilisierung und Automatisierung speziell aggressiver Verhaltensprobleme“ (Beelmann & Raabe 2007, S. 75).

Perspektive der Resilienzforschung

Resilienz (lat. resilire = zurückspringen, abprallen), auch Anpassungsfähigkeit, bezeichnet den Prozess und die Fähigkeit einer Person, sich an Probleme und Veränderungen anzupassen. Dieser Prozess umfasst Auslöser/Stressoren, die Resilienz erfordern, Ressourcen, die Resilienz begünstigen und Konsequenzen, die Veränderungen im Verhalten oder in den Einstellungen hervorrufen. Brezinka (2003) identifiziert drei Risikofaktoren, die als einflussreiche Stressoren Auswirkungen auf die kindliche ESE haben können:

- *biologische Risikofaktoren:* Geschlecht (bei Jungen 3-4-mal größeres Risiko), pränatale Risiken (Drogen etc.), genetische Ausstattung und hormonelle Besonderheiten (erhöhter Testosteron-, geringer Cortisolspiegel), strukturelle Besonderheiten des zentralen Nervensystems (Störungen des Frontallhirnlappens), Besonderheiten in den autonomen Funktionen (geringer Herzrate, geringer Hautleitwiderstand)
- *psychische Risikofaktoren:* Aufmerksamkeits-, Hyperaktivitäts- und Impulsivitätsprobleme, unzureichende Impulskontrolle und Emotionsregulation (dadurch Entwicklung angemessenen Sozialverhaltens gehindert), verzerrte kognitive Informationsverarbeitung, überzogene Selbsteinschätzung, unzureichendes Einfühlungsvermögen
- *sozial-ökologische Risikofaktoren:* unsichere Bindung an die Eltern im Kleinkindalter, restriktive, inkonsistente Erziehung, Mangel an emotiona-

ler Wärme und Unterstützung für das Kind, soziale Ablehnung durch Gleichaltrige

Loeber (1990) identifiziert verstärkende Bedingungen für die Ausprägung von emotional-sozialen Entwicklungsstörungen je mehr Risikofaktoren vorliegen und kumulieren, je früher ein Problemverhalten gezeigt wird, je häufiger ein Problemverhalten auftritt, je länger ein Problemverhalten bereits gezeigt wurde, je vielfältiger das Problemverhalten ist (z.B. Hyperaktivität, Impulsivität, oppositionelles und aggressives Verhalten, Delinquenz und Kriminalität) und je verschiedener die Kontexte (Familie, Kindergarten, Schule, usw.) sind, in denen das Problemverhalten manifest wird.

Die Resilienzforschung nimmt an, dass den Risikofaktoren Schutzfaktoren gegenüberstehen. So können biologische, psychologische und soziale Faktoren eine protektive Wirkung auf die ESE von Kindern entfalten (Masten & Powell, 2003). Nach Laucht et al. (1997) sind diese schützenden Ressourcen in zwei Bereiche einteilbar: *personale Ressourcen* wie ein positives Selbstwertgefühl, eine hohe Selbstwirksamkeitserwartung, ein positives Sozialverhalten und ein hohes Maß an aktiver Stressbewältigung sowie *soziale Ressourcen* wie das Klima genereller sozialer Unterstützung, positive Freundschaften mit Gleichaltrigen, eine positiv wahrgenommene Schulerfahrung, die Lehrkraft als eine emotional vertraute, wertschätzende Bezugsperson und gleichzeitig als positives Modell (Hennemann et al. 2017, S. 37). Die Erkenntnisse zur Resilienzforschung unterstützen somit ein wechselseitiges Entwicklungsverständnis zwischen Stressoren und Ressourcen, Risiko- und Schutzfaktoren sowie Vulnerabilität und Resilienz. Je nach Störungsart, Personen- und Kontextmerkmalen haben Risiko- und Schutzfaktoren unterschiedliche Auswirkungen (Lösel & Bender 2008, S. 71), sodass es bislang unmöglich ist, generelle kausale Prozesse zwischen ihnen auszumachen. In Studien zur Resilienz konnten jedoch konkrete protektive Faktoren anhand operationalisierter Trainingsziele nachgewiesen werden, wie eine angemessene sozial-kognitive Informationsverarbeitung (Crick & Dodge, 1994), eine gute Impulskontrolle (Gottfredson & Hirschi, 1990) und positives Sozialverhalten (Wilson et al., 2001). Für die pädagogische Förderung ergeben sich daraus vor allem präventive Aufgaben, die sich am konkreten Entwicklungsverlauf des Kindes orientieren.

2.1.2.4 Prävalenzen und Komorbiditäten

Etwa 1.3 % aller Schüler*innen in Deutschland erhalten die Diagnose SF ESE (Kultusministerkonferenz, 2018). Von diesen nehmen bundesweit etwa die Hälfte am Unterricht der Regelschule teil. National und international werden die klinisch relevanten Prävalenzraten für emotional-soziale Entwicklungsstörungen bei Kindern und Jugendlichen in westlichen Ländern allerdings valide mit 12 bis 15 % angegeben (Ihle & Esser, 2008). In Deutschland

betragen diese rund 17.2 % und im Alter zwischen 7 und 10 Jahre sogar 19.8 % (Klasen et al., 2017). Diese Zahlen deuten darauf hin, dass knapp ein Fünftel aller Grundschulkinder in Deutschland Verhaltensauffälligkeiten und/oder emotional-soziale Entwicklungsstörungen aufweist. Man kann daher annehmen, dass der Anteil dieser Schüler*innen den SF ESE um ein Vielfaches übersteigt. Betrachtet man die in dieser Arbeit im Fokus stehenden externalisierend emotional-sozialen Entwicklungsstörungen im Grundschulalter, weisen internationale Studien auf folgende Prävalenzen hin: Eine *Störung mit oppositionellem Trotzverhalten* liegt bei 7.2 % der 4-11-Jährigen vor (Georgiades et al., 2019), eine *Störung des Sozialverhaltens* weisen 3.7 % der 6-11-Jährigen auf (Ghandour et al., 2019) und eine *Aufmerksamkeitsdefizit-/Hyperaktivitätsstörung* tritt bei 6-8-Jährigen mit 2.1 % und bei 9-11-Jährigen mit 6.1 % auf (Göbel et al., 2018).

Häufig sind mehrere Formen von emotional-sozialen Entwicklungsstörungen zugleich zu beobachten (Komorbidität). Insbesondere zwischen den verschiedenen externalisierenden Verhaltensstörungen bestehen beträchtliche, zum Teil deutlich über 50 % liegende Überschneidungen (Salmanian et al., 2019). Aber auch externalisierende und internalisierende Störungen treten bis zu 36.6 % komorbid auf (Ghandour et al., 2019). Darüber hinaus wird das gleichzeitige Auftreten von Lern- und Verhaltensstörungen bei ca. 50 % der betroffenen Kinder beobachtet (Klauer & Lauth, 1997; Visser et al., 2019).

2.1.2.5 Prognose und spezifische Verlaufsformen

Emotional-soziale Entwicklungsstörungen bei Kindern und Jugendlichen sind oft wiederkehrend und anhaltend, und die Entwicklung von Komorbiditäten sowie chronischen Beeinträchtigungen im Erwachsenenalter sind häufig. Beelmann und Raabe (2007) sprechen daher von verfestigenden und kumulierenden Risikokonstellationen im Bereich ESE. So zeigt u.a. die dritte BELLA-Studie⁴, dass 40.7 % der auffälligen Kinder auch nach sechs Jahren weiterhin Verhaltensauffälligkeiten und/oder emotional-soziale Entwicklungsstörungen aufweisen (Klasen et al., 2017). Die hohen Persistenzraten (Stabilität) führen wiederum zu ungünstigen Entwicklungsprognosen für die schulische, berufliche und persönliche Entwicklung (Ihle & Esser, 2008). Ein niedriger sozioökonomischer Status, die psychischen Belastungen der Eltern, die eigenen kognitiven Fähigkeiten, Mobbing-Erfahrungen und Drogenmissbrauch können als zentrale Risikofaktoren für die Chronifizierung von Verhaltensstörungen identifiziert werden (Otto et al., 2021), während ein positi-

4 Die „Befragung zum seelischen Wohlbefinden und Verhalten“ (BELLA) untersucht die psychische Gesundheit und gesundheitsbezogene Lebensqualität von Kindern und Jugendlichen in Deutschland. Die BELLA-Studie ist ein Zusatzmodul der Studie zur Gesundheit von Kindern und Jugendlichen in Deutschland (KIGGS), welche im Rahmen der Bundesgesundheitsberichterstattung am Robert Koch-Institut durchgeführt wird.

ves Familienklima und soziale Unterstützung sich als wichtige Ressourcen für die ESE erweisen (Klasen et al., 2017).

2.1.2.6 Zusammenfassung und Schlussfolgerungen

Anhand der Entstehung, der Prävalenzen, des Verlaufs und der Folgewirkungen von Verhaltensauffälligkeiten und emotional-sozialen Entwicklungsstörungen wird deutlich, was Forscher*innen schon lange fordern (u.a. Beelmann & Raabe, 2007; Heinrichs et al., 2013; Ihle & Esser, 2008): Um der Manifestierung von Störungen vorzubeugen und/oder um diese abzumildern, sollten anhand frühzeitig einsetzender Programme mit empirischen und theoretischen Qualitätskriterien Risikofaktoren gemindert und Ressourcen gestärkt werden, die insbesondere für Kinder mit niedrigem sozioökonomischen Status gezielte und niedrigschwellige Präventions- und Interventionsansätze bieten. Die Schule wird hierbei als „wichtigstes Setting präventiver Maßnahmen“ (Beelmann 2008, S. 442) angesehen. Der nächste Abschnitt widmet sich daher umfassend der Thematik schulischer Förderung im Bereich ESE.

2.1.3 Schulische Förderung im Bereich ESE

Frühzeitige präventive schulische Förderung wird in Deutschland durch zwei zentrale, eng miteinander verbundene bildungspolitische Praktiken erschwert, die insbesondere im Zuge der inklusiven Ausrichtung der Schulsysteme in den einzelnen Bundesländern vermehrt kritisiert werden: das sogenannte *Wait-to-fail-Problem* und das *Etikettierungs-Ressourcen-Dilemma*. Das *Wait-to-fail-Problem* kann beschrieben werden als die Tendenz von Bildungssystemen, Unterstützungsleistungen für Schüler*innen erst zu gewähren, wenn die Problementwicklung schon sehr weit fortgeschritten ist (Hillenbrand 2015, S. 195). Das *Etikettierungs-Ressourcen-Dilemma* beschreibt den Umstand, dass zusätzliche Ressourcen im Bildungssystem vordergründig durch Etikettierungsprozesse, z.B. die Feststellung eines sonderpädagogischen Förderbedarfs, gewährt werden. Etikettierungsprozesse sind oftmals mit einem erheblichen Zeitbedarf verbunden, der eine frühzeitige Intervention erschwert (Hennemann et al. 2015, S. 117f.). Dabei sind Vorteile schulischer Prävention bei Verhaltensauffälligkeiten und emotional-sozialen Entwicklungsstörungen, wie Reicher und Jauk (2012) sie im Folgenden zusammenfassen, klar ersichtlich:

- Im Setting Schule sind durch die Schulpflicht nahezu alle Kinder und Jugendlichen erreichbar.
- Im Vergleich zu anderen Settings ist die Schule durch vorhandene Logistikkressourcen wie Pädagog*innen, Räume und entsprechendes Material ein guter Standort für Prävention.

- Der schulische Bildungsauftrag sieht die Förderung sozialen Lernens und die Prävention emotional-sozialer Entwicklungsstörungen vor.

Der zuletzt genannte Bildungsauftrag von Schule findet sich in den Lehrplänen der Bundesländer vor allem in den Fächern Sprachen, Sachunterricht, Religion, Kunst und Sport wieder (Hennemann et al. 2017, S. 72f.). Die Ständige Konferenz der Kultusminister der Länder formuliert im Rahmen der Empfehlungen zur sonderpädagogischen Förderung im FSP ESE einen eigenen Präventionsauftrag für die Schule als präventiven Förderort: „Durch vorbeugende Maßnahmen können die Verfestigung sozial unangemessener Handlungsmuster frühzeitig verhindert, erwünschte angebahnt und dadurch die schulische Entwicklung positiv beeinflusst werden“ (KMK 2000, S. 3). Somit kommt insbesondere auch aus sonder- und inklusionspädagogischer Sicht der Prävention von Verhaltensauffälligkeiten und emotional-sozialen Entwicklungsstörungen eine zentrale Bedeutung zu. An diese geforderten Entwicklungsaufgaben von Schule anknüpfend, werden im Folgenden Prinzipien, Strukturen und konkrete Umsetzungsmöglichkeiten präventiver schulischer Förderung im Bereich ESE aufgezeigt, die teilweise in anderen Ländern, wie den USA oder Finnland, im Zuge der Einführung inklusiver Schulstrukturen bereits großflächig umgesetzt werden (Jahnukainen & Itkonen, 2015).

2.1.3.1 Postulat der Prävention

Unter Prävention (lat. *praevenire* = zuvorkommen, verhüten, vorbeugen) bzw. präventiven Maßnahmen werden alle Interventionen verstanden, die das Auftreten eines bestimmten Problems vermeiden oder reduzieren sollen (Brezinka, 2003). *Pädagogische Prävention* im Bereich ESE definieren Hennemann et al. als

systematisches, theoretisch und empirisch gut begründetes pädagogisches Handeln, das zum Ziel hat, mit Hilfe erzieherischer Mittel und erzieherischen Einwirkens in schulischen und außerschulischen pädagogischen Settings Kinder und Jugendliche so zu stärken, dass Risiken und Folgen einer belastenden sozial-emotionalen Entwicklung vermieden, gemildert oder bewältigbar werden (2017, S. 55).

Die konsequente schulische Prävention im Bereich ESE impliziert demnach, sich anbahnende negative Entwicklungstendenzen – sichtbar etwa anhand nicht-regelkonformen, störenden und anderweitig schwierigen oder auffälligen Verhaltens – so früh wie möglich mit präventiv-pädagogischen Maßnahmen zu begegnen oder deren Entstehung gar zu verhindern. Nach der international anerkannten Klassifikation des amerikanischen *Institute of Medicine* lassen sich präventive Maßnahmen in drei Formen einteilen (Hillenbrand 2015, S. 187). Während sich *universelle* Präventionsmaßnahmen an die Gesamtpopulation richten, fokussieren gezielte Präventionsmaßnahmen, zu

denen *selektive* und *indizierte* Maßnahmen gehören, bestimmte Zielgruppen. Selektive Präventionsmaßnahmen richten sich an Personen mit erhöhtem biologischen, psychischen oder sozialen Risiko für die Entwicklung psychischer Störungen (z.B. Kinder psychisch erkrankter Eltern, Kinder aus sozialen Problemmilieus) und indizierte Maßnahmen sprechen Personen mit einem ausgeprägten und hohen Risiko an, die bereits erste Symptome von Störungsformen aufweisen (Beelmann & Raabe 2007, S. 132; Heinrichs et al., 2013).

In ihrer umfassenden Metaanalyse über schulbasierte Prävention psychischer Störungen identifizieren Durlak et al. (2011) zwei grundlegende effektive schulische Maßnahmen:

1. die Etablierung positiver, sicherer Lernumgebungen und
2. die gezielte Förderung sozial-emotionaler Fertigkeiten in der Schule.

Bei der Etablierung positiver, sicherer Lernumgebungen haben sich national wie international Ansätze der effektiven Klassenführung, des *Classroom Managements*, bewährt. Kein anderes Unterrichtsmerkmal ist so eindeutig und konsistent mit dem Lernfortschritt von Schulklassen verknüpft wie die Klassenführung (Hattie, 2018). Gutes Classroom Management ist die transparente und konsistente Planung, Organisation und motivierende Gestaltung des Unterrichts, die eine möglichst optimale Lernaktivierung und Zeitznutzung bei den Schüler*innen zur Folge hat (Evertson & Emmer, 2009). Dies wiederum führt zu Lernfortschritten, einem insgesamt hohen Leistungsniveau bei den Schüler*innen und einer geringen Anzahl von Unterrichtsstörungen (Helmke, 2010). Es können proaktive und reaktive Kriterien effektiven Classroom Managements unterschieden werden. Diese sind in Tabelle 6 aufgelistet.

Der zweite von Durlak et al. (2011) identifizierte effektive Ansatz schulbasierter Prävention ist die gezielte Förderung sozial-emotionaler Fertigkeiten durch spezielle Förder- und Trainingsprogramme. „Der Grundgedanke besteht in dem Ziel, durch die Förderung sozialer und emotionaler Kompetenzen potentiell Problemverhalten zu vermeiden“ (Hillenbrand 2015, S. 191). Gezielte Förder- und Trainingsprogramme können sowohl auf der universellen, der selektiven als auch, nach entsprechender Anpassung, der indizierten Präventionsebene Anwendung finden. Ob sie im Rahmen schulbasierter Prävention auch wirksam sind, hängt maßgeblich von zwei Faktoren ab: der *theoretischen Fundierung* und der *Evidenzbasierung*. Förder- und Trainingsprogramme sind dann theoretisch fundiert, „wenn die Inhalte und Methoden von mindestens einem wissenschaftlichen Modell, das Lernen und Entwicklung von Jugendlichen erklärt, ausgehend entwickelt und ausgewählt wurden“ (Blumental et al. 2020, S. 22). Evidenzbasierung meint die „Integration des aktuellen wissenschaftlichen Erkenntnisstand[s] in die Expertise der Professionellen zur bestmöglichen Ausrichtung des Handelns an der individuellen Situation der Zielgruppe bzw. Lernenden“ (Hillenbrand 2015, S. 208).

Tabelle 6: Kriterien effektiven Classroom Managements nach Evertson und Emmer (2009) und Helmke (2010); entnommen aus Hennemann et al. (2017, S. 83)

Kriterien effektiven Classroom Managements	
proaktiv	reaktiv
– Klassenraum vorbereiten – Regeln und Verfahrensweisen planen und unterrichten – Konsequenzen festlegen – Positives (Lern-)Klima schaffen – Beaufsichtigen und Überwachen – Vorbereiten des Unterrichts – Verantwortlichkeit der Schüler*innen – Unterrichtliche Klarheit – Kooperative Lernformen	– Unterbindung unangemessenen Schüler*innenverhaltens – Strategien für potentielle Probleme

Leider erreichen präventive Maßnahmen aufgrund der Art ihrer Umsetzung in Schulen oft nicht ihre optimal möglichen Effekte. Für die effektive Implementation pädagogisch-psychologischer Präventionsprogramme in der Schule können jedoch konkrete Anforderungen herausgestellt werden. So extrahieren Hennemann et al. (2017, S. 57ff.) aus Metaanalysen zur schulischen Prävention von emotional-sozialen Entwicklungsstörungen (Beelmann, 2008; Durlak et al., 2011; Sklad et al., 2012), systematischen Reviews (Brezinka, 2003; Casale et al., 2014; Heinrichs et al., 2002) und der Zusammenfassung zu den Kriterien effizienter SEL⁵-Ansätze (Denham & Burton, 2003; Reicher & Jauk, 2012) zehn Indikatoren, die die Wirkung präventiver Förderung moderieren.

1. Theoretische Fundierung: Programme sollten von einem theoretisch fundierten Ansatz ausgehen, der auf einer multidisziplinären Sichtweise und entwicklungsadäquaten Strategien beruht.
2. Instruktion und Lebensweltbezug: Die Individualisierung von Programmetechniken und deren Zielgruppenspezifität sind wichtig, da die Profile sozial-emotionaler Stärken und Risiken extrem heterogen sind. Instruktionsbasierte Elemente sollten mit persönlichem Bedeutungsgehalt und Lebensweltbezug verknüpft werden.
3. Klare Zielformulierung: Die klare Definition dessen, was gefördert werden soll, unterstützt die Wirksamkeit. Je konkreter der Trainingsgegenstand formuliert ist und je eindeutiger das zu verändernde Verhalten antizipiert wird, desto größer sind die zu erwartenden Präventionseffekte.

5 SEL = social and emotional learning (sozial-emotionales Lernen)

4. Transfer in den Alltag: Die Inhalte und vermittelten Fertigkeiten sollen in Unterrichtsinhalte, in das Unterrichtsgeschehen und in die Beziehungsgestaltung zwischen Lehrkräften und Schüler*innen einfließen. Das alltägliche Miteinander in der Schule ist das ‚real life-Setting‘, um sozial-emotionale Kompetenzen zu lehren, zu lernen, zu modellieren und zu reflektieren.
5. Vom Verhalten zum Setting: Die Schul- und Klassenumgebung muss die erwünschten Verhaltensweisen anhand eines fürsorglichen, unterstützen, abwechslungsreichen und respektvollen Klimas unterstützen und so die Beziehungsfähigkeit und Eigenverantwortlichkeit der Schüler*innen stärken.
6. Früher Beginn und längerer Zeitraum: Die Reduktion von Verhaltensauffälligkeiten gestaltet sich bei jüngeren Kindern und somit während der Primarstufe effektiver, als bei älteren. Präventionszeiträume von mindestens einem Jahr zeigen höhere Effekte als kürzere Programme.
7. Erziehungspartnerschaften mit den Eltern/multisystemische Ausrichtung: Programme sollten nicht isoliert in einem Klassenraum oder Fach ablaufen, sondern multimodal konzipiert und auf alle im System Agierenden ausgeweitet werden, um längerfristige Effekte erzielen zu können. Programme, die die Eltern für einen längeren Zeitraum mit einbeziehen, sind erfolgversprechender.
8. Reflexion der Theorie durch die Pädagog*innen: Es ist essentiell, dass die Pädagog*innen die Durchführung des Trainings und die Erfahrungen, die sie dabei sammeln, kontinuierlich vor dem Hintergrund des theoretischen Wissens reflektieren. Aspekte, die sich aus der Umsetzung ergeben, sollten supervisorisch oder anhand eines unterstützenden Austausches bearbeitet werden können.
9. Begleitende Evaluation: Begleitstudien sind wichtig, um Erfolge zu belegen und die Nachhaltigkeit von Programmen zu sichern.
10. Gute Implementation: Wirksame Präventionsprogramme legen Wert auf eine gute Implementation. Die Gelingensbedingungen einer guten Implementation sind dabei vielfältig mit den bereits genannten Wirksamkeitsfaktoren verknüpft.

Durlak und DuPre (2008) beschreiben ebenfalls konkrete Faktoren für eine gelingende Implementation. Im Wesentlichen beziehen sich diese auf die Art der Durchführung des Präventionsprogramms durch die entsprechenden Pädagog*innen. Letztere beeinflussen somit maßgeblich die Wirksamkeit der Maßnahme, etwa durch „das persönliche Engagement während der Durchführung, eine positive Einstellung zum Präventionsprogramm, die Motivation, das Programm hochwertig durchzuführen und vor allem die Vertrautheit mit dem Präventionsansatz“ (Hennemann et al. 2017, S. 62). Als stärkster Faktor für die empirische Evidenz zeigte sich außerdem das sogenannte *monitoring*,

hier die Begleitung der Lehrkräfte bei der Umsetzung des Programms (Novins et al., 2013).

Für die effektive Implementation pädagogisch-psychologischer Präventionsprogramme in der Schule kann somit festgehalten werden, dass die diesbezügliche Professionalisierung von Pädagog*innen und deren Begleitung während des Prozesses essentielle Bedingungen für den Erfolg darstellen. Im Folgenden werden zunächst drei empirisch bewährte Handlungsansätze vorgestellt, auf deren theoretischer Basis viele Konzepte und Maßnahmen im Bereich ESE fußen und die an die in Kap. 2.1.2.3 vorgestellten Erklärungsansätze zur Entstehung von Verhaltensauffälligkeiten und emotional-sozialen Entwicklungsstörungen anknüpfen. Das darauffolgende Kapitel widmet sich der *evidenzbasierten Praxis* und den sich daraus ableitenden Anforderungen für eine gelingende präventive Förderung in der Schule.

2.1.3.2 Empirisch bewährte Handlungsansätze

Im Bereich ESE haben sich empirisch insbesondere lerntheoretische, kognitionspsychologische und entwicklungspsychologische Handlungsansätze bewährt (Blumenthal et al., 2020; Hillenbrand, 2015).

Lerntheoretisch fundierte Handlungsmöglichkeiten oder Methoden der *Verhaltensmodifikation* basieren auf den Modellen der Lern- und Verhaltenstheorie (vgl. Kap. 2.1.2.3) und damit der Grundannahme, dass Verhalten (v)erlernbar ist. Theoriegeleitete und lösungsorientierte Verhaltensmodifikation ermöglicht die Veränderung von Verhalten durch eine Prozessorientierung in fünf Schritten (Blumenthal et al., 2020):

1. Problembestimmung und Problemformulierung (Welches Verhalten ist in welchem Ausmaß problematisch?)
2. Problemanalyse (Welche individuellen bzw. Umweltfaktoren begünstigen das Verhalten?)
3. Bestimmung und Formulierung des erwünschten Verhaltens (Wie würde sich das gewünschte Verhalten bemerkbar machen?)
4. Planung und Durchführung der Modifikation (Welche lerntheoretischen Handlungsmöglichkeiten können das Problem verbessern?)
5. Prüfung des Erfolgs (Wie erkenne ich, ob sich das Verhalten verändert hat? Gegebenenfalls Neubestimmung des Problems)

Strategien der Verhaltensmodifikation sind u.a. Konditionierungstechniken, wie Token-Systeme und Response-Cost-Verfahren, oder kognitiv ausgerichtete Maßnahmen, wie Selbstinstruktionstechniken und Rollenspiele. Die genannten Maßnahmen führen nachweislich zur Verbesserung des unterrichtsbezogenen Verhaltens, der Schüler*innenleistungen und zur Verringerung auffälliger Verhaltensweisen (Hulac & Briesch 2017, S. 85). Auch im Bereich der Kinder- und Jugendpsychotherapie stellen sich behaviorale Interventionen anhand internationaler Metaanalysen, insbesondere für den Bereich

externalisierender Störungsformen, als besonders wirksam dar (Döpfner 2003, S. 263). Bei der Prävention von Verhaltensauffälligkeiten und emotional-sozialen Entwicklungsstörungen hat sich auf lange Sicht die Kombination von kognitiven und behavioralen Verfahren als besonders wirksam erwiesen. In der Metaanalyse von 84 kindorientierten Maßnahmen zur Prävention antisozialen Verhaltens von Lösel und Beelmann (2003) wird deutlich, dass reine Verhaltenstrainings zwar kurzfristig (Post-Test) gesehen robuste Effekte erzielen, diese Wirkung jedoch längerfristig (Follow-Up) nicht aufrechterhalten. Kognitive bzw. kognitiv-behaviorale Trainings wiederum zeigten insbesondere bei den Follow-Up-Erhebungen deutliche Wirksamkeit (Beelmann 2006, S. 159; Lösel & Beelmann, 2003; Wilson & Lipsey, 2007). Dies leitet über zum zweiten empirisch bewährten Handlungsansatz, der Förderung der ESE aus kognitionspsychologischer Perspektive.

Empirisch bewährte *kognitionspsychologisch fundierte Handlungsansätze* beruhen, neben der bereits erwähnten kognitiv-behavioralen Ausrichtung, vorrangig auf dem Modell sozial-kognitiver Informationsverarbeitung (SKI; vgl. Kap. 2.1.2.3). Blumenthal et al. (2020, S. 29) stellen auf Basis des erweiterten Modells von Lemerise und Arsenio (2000) folgende zentrale kognitionspsychologisch orientierte Förderschwerpunkte für die Schule heraus:

- Differenzierung sozialer Wahrnehmung
- Emotionen bewusst machen und einen angemessenen Umgang üben
- Erweiterung des Repertoires zur emotionalen Steuerung
- Ziele klären
- prosoziale Handlungsalternativen erlernen
- sozial angemessene Handlungsmöglichkeiten bewerten
- Umsetzung sozial angemessener Handlungsmöglichkeiten einüben

Entwicklungspsychologisch fundierte Handlungsmöglichkeiten knüpfen an die Resilienzforschung (vgl. Kap. 2.1.2.3) und die darin bedeutsamen Schutzfaktoren an. Die nach Laucht et al. (1997) in personal und sozial gegliederten Schutzfaktoren können im pädagogischen Kontext nutzbar gemacht werden. Als konkrete Trainingsziele können beispielsweise eine angemessene SKI, eine gute Impulskontrolle und ein positives Sozialverhalten anvisiert werden (Blumenthal et al. 2020, S. 32). Entwicklungspsychologische Handlungsansätze stellen grundsätzlich die Minimierung von Risikofaktoren und die systematische Stärkung von Schutzfaktoren in den Vordergrund (s. Kap. 2.1.2.3).

Neben der *theoretischen Fundierung* von Konzepten und Maßnahmen zur Förderung der emotional-sozialen Kompetenzen und des Verhaltens, wie sie in diesem Kapitel vorgestellt wurden, ist für den effektiven pädagogischen Einsatz deren wissenschaftlich und praktisch erprobte Wirksamkeit und Anwendbarkeit wesentlich. Dieser sogenannten *evidenzbasierten Praxis* widmet sich das nächste Kapitel.

2.1.3.3 Evidenzbasierte Praxis

Der Begriff *Evidenzbasierung* (engl. *evidence-based* = auf Beweise/Nachweise gegründet, nachweisorientiert) beschreibt zunächst allgemein die Berücksichtigung und Nutzung der besten verfügbaren Informationen, um Entscheidungen zu treffen oder Empfehlungen zu geben. Dazu werden gezielt identifizierte empirische Nachweise als Grundlage von Entscheidungen eingefordert. In der Pädagogik bezieht sich die Evidenzbasierung vorrangig auf die Frage, ob ein vorteilhafter Effekt durch eine pädagogische Maßnahme wissenschaftlich mittels empirischer Methodik nachgewiesen werden kann, „da pädagogische Fachkräfte ansonsten einer erhöhten Gefahr fehlerhafter Schlussfolgerungen ihrer praktischen Implikationen ausgesetzt sind“ (Hennemann et al. 2017, S. 66). Wichtig zu betonen ist: Evidenzbasierte Verfahren erhöhen die Wahrscheinlichkeit, eine positive Wirkung auszulösen, können dies aber nicht garantieren. Mit den durch Evidenzbasierung „operationalisierte[n], replizierbare[n] Handlungsformen, die einer kritischen Prüfung durch wissenschaftliche Forschung unterzogen wurden, und daher belegbare, nachgewiesene Fakten ihrer positiven Wirksamkeit vorlegen können“ ist aber keineswegs eine standardisierte, technologische, rezeptartige Vorgehensweise gemeint, sondern das Anliegen, auf der Basis „professioneller Expertise und Erkenntnis des konkreten Einzelfalls nach dem besten wissenschaftlichen Erkenntnisstand zu handeln“ (Hillenbrand 2015, S. 174f.). Neben der Identifikation von evidenzbasierten Verfahren anhand der besten verfügbaren wissenschaftlichen Erkenntnis ist vor allem auch die individuelle Expertise der Anwendenden bezogen auf die Zielgruppe und -person entscheidend. Auf diesen Überlegungen fußend hat sich in der Pädagogik das sogenannte Trias-Modell *Evidenzbasierter Praxis (EbP)* etabliert (Blumenthal & Mahlau, 2015; Grosche, 2017). Ziel evidenzbasierter Praxis ist es, „die bestmögliche und geeignetste Handlungsstrategie für die beteiligten Personen zu finden“ (Blumenthal & Mahlau 2015, S. 409). Das Modell stellt die Auswahl von pädagogischen Maßnahmen und Konzepten auf insgesamt drei Grundprinzipien und fordert deren gleichrangige (Grosche, 2017) Berücksichtigung:

- *interne Evidenz*: praktische Expertise und Professionalität der Anwender*innen
- *soziale Evidenz*: Werte, Wünsche und Bedarfe der Schüler*innen
- *externe Evidenz*: aus hochwertigen empirischen Studien wissenschaftlich belegte Wirksamkeit der Fördermethode

Die Überprüfung der externen Evidenz kann auf zwei Arten erfolgen (Casale et al., 2015a). Um die Frage nach der Wirksamkeit beantworten zu können, sollten zum einen experimentelle Gruppenuntersuchungen und/oder Einzelfallstudien durchgeführt werden. Zum anderen sollte die Evidenzbasierung im Einzelfall überprüft werden, das bedeutet, die förderbegleitende, regel-

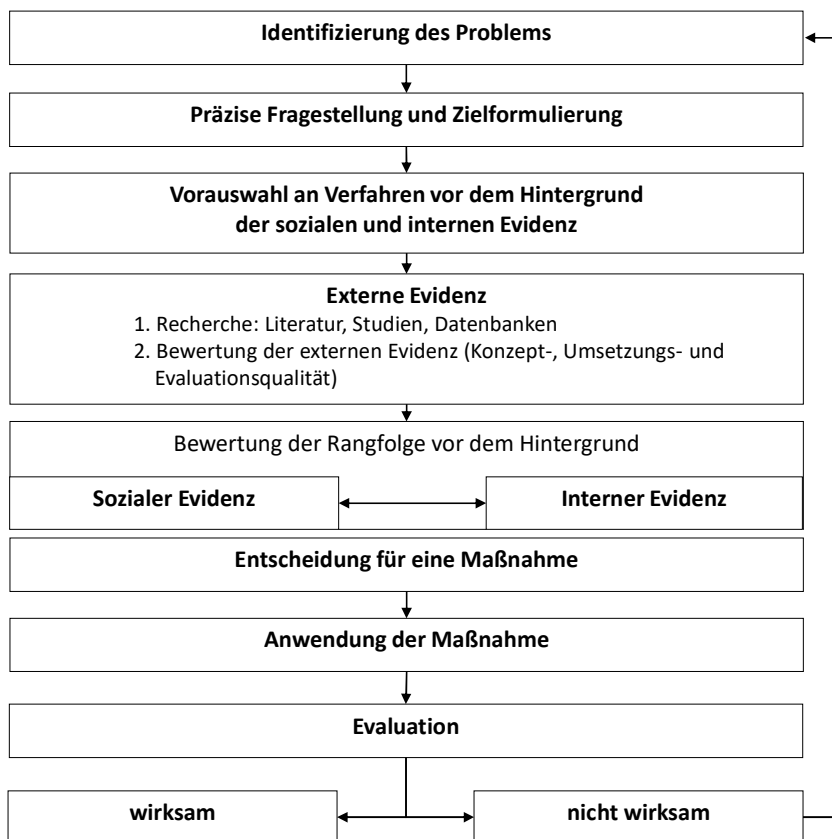


Abbildung 2: Ablauf eines evidenzbasierten Entscheidungsprozesses; entnommen aus Blumenthal und Mahlau (2015, S. 410)

mäßige Überprüfung des individuellen Lern- und Fördererfolgs am einzelnen Kind. Die soziale Evidenz berücksichtigt die spezifischen Bedürfnisse der Schüler*innen bzw. deren Erziehungsberechtigter. Das bedeutet, dass *die* Fördermethoden ausgewählt werden, die zu den Schüler*innen der Klasse am besten passen. Die interne Evidenz bezieht sich auf die Expertise der pädagogischen Fachkräfte, und damit auf deren Berufserfahrung, Professionswissen (Baumert & Kunter, 2006) und (Selbst-)Wirksamkeitserfahrungen in Bezug auf spezifische Förderkonzepte, -verfahren und -materialien. Im Bereich der internen Evidenz entscheidet sich die Lehrkraft für die Fördermethode, die am besten zu ihr selbst passt (Blumenthal & Mahlau 2015, S. 407). In der EbP wird grundsätzlich ein stufenweises Vorgehen empfohlen, um die Ent-

scheidung für ein geeignetes Interventionsverfahren oder Fördermaterial gut treffen zu können. Abbildung 2 visualisiert den Ablauf dieses evidenzbasierten Entscheidungsprozesses, der die drei erläuterten Grundprinzipien Schritt für Schritt mit einbezieht.

Resümierend kann festgehalten werden: Für die präventive schulische Förderung im Bereich ESE ist eine pädagogische Praxis, im Sinne einer EbP, notwendig, bei der theoretisch fundierte und evidenzbasierte Verfahren und Fördermaterialien Anwendung finden, die auf die Schüler*innen abgestimmt und von den Pädagog*innen professionell umgesetzt werden (können). Im Folgenden soll anhand des Ansatzes mehrstufiger Förderkonzeptionen dargestellt werden, wie die präventive Förderung auf schulorganisatorischer Ebene etabliert und die beschriebenen Dimensionen und Konzeptionen präventiven Arbeitens darin verortet werden können.

2.1.3.4 Mehrstufige schulische Förderkonzeptionen

Ausgehend von der Klassifikation präventiver Maßnahmen anhand der Ebenen universelle, selektive und indizierte Prävention (s. Kap. 2.1.3.1) wurden schulische Rahmenkonzepte entwickelt, die als *mehrstufige schulische Förderkonzeptionen* bezeichnet werden. Eines der prominentesten Beispiele eines solchen Mehrebenen-Modells ist das *Response-to-Intervention*-Konzept, welches für die schulische Prävention von emotional-sozialen Entwicklungsstörungen zunehmend Beachtung findet. *Response-to-Intervention (RTI)* ist ein Modell zur schulischen Prävention und frühen Intervention bei Lernbeeinträchtigungen und Verhaltensproblemen aus den USA (Blumenthal et al., 2014; Cheney et al., 2008; Gresham et al., 2005). Es sieht ein gestuftes System aus Diagnose und Förderung vor, in dem durch kleinschrittige Lernverlaufsdiagnostik und Verhaltensfortschrittsmessungen die Wirksamkeit einer Intervention bei Lern- und Verhaltensproblemen überprüft und anhand der Ergebnisse die Förderung zeitnah anpasst werden kann. Das RTI-Konzept setzt an der individuellen Lernausgangslage des Kindes an und fokussiert auf den förderbedürftigen Verhaltensbereich. Das übergeordnete Ziel ist dabei der bestmögliche Fördererfolg der Schüler*innen (*Response*) durch eine bestmögliche individuelle Förderung (*Intervention*). Weist ein Kind etwa einen erhöhten Bedarf an Unterstützung auf, wird die Förderung auf einer nächsten Förderebene intensiviert, spezifiziert und weiter individualisiert. Die Anzahl der Förderebenen ist nicht vorgegeben und kann je nach Bedarf differenziert werden. Oft wird allerdings ein dreistufiges System aus Diagnostik und Förderung umgesetzt (Mahlau et al., 2016). Im Folgenden werden die drei Stufen für den Förderbereich ESE konkretisiert.

Auf der *Förderebene I (universelle Prävention)* werden alle Schüler*innen einer Lerngruppe durch schul- und klassenweite Maßnahmen wie qualitativ hochwertigen Unterricht, soziales Lernen und effektives Classroom Management einbezogen und angesprochen. Konkrete evidenzbasierte Maß-

nahmen zur emotional-sozialen Unterstützung auf Förderebene I sind etwa Peer Tutoring, Aktivierung der Lernenden, klare Kommunikation, Strukturierung von Lernprozessen und Klassenraum, Modellierung positiven Verhaltens, Lehrkraft-Schüler*in-Beziehung und Routinen (Hillenbrand 2015, S. 199). Zur Diagnostik werden in regelmäßigen Abständen, etwa dreimal jährlich, universelle Screenings eingesetzt, die eine große Bandbreite unterrichtsrelevanter Verhaltensweisen erfassen, beispielsweise anhand des *Kurzscreenings schulbezogenen Verhaltens* (Blumenthal & Blumenthal, in Vorb.). Weisen die Ergebnisse des Screenings auf Verhaltensprobleme bei einzelnen Schüler*innen hin, so ist dies ein Indiz dafür, dass sie von den Fördermaßnahmen auf Ebene I nicht genügend angesprochen wurden. Laut internationaler Studien betrifft dies in der Regel etwa 10 bis 15 % aller Schüler*innen (Hillenbrand 2015, S. 197). Für diese Kinder wird dann eine differenzierte Diagnostik und intensivere Förderung auf *Förderebene II (selektive Prävention)* angestrebt. Hierbei werden neben den klassenweiten Maßnahmen zusätzlich evidenzbasierte Maßnahmen in Kleingruppen oder als unterrichtsintegrierte Individualförderung realisiert. Durch engmaschige Verhaltensverlaufsdiagnostik wird etwa einmal täglich überprüft, inwiefern sich eine Verhaltensverbesserung bei einem Kind einstellt. Konkrete evidenzbasierte Maßnahmen zur emotional-sozialen Unterstützung auf Förderebene II sind beispielsweise Regeln, Verstärkersysteme, Token-Programme, Verhaltenskontrakte, gezielte Förderprogramme, systematisches Elterninformationssystem, Interventionstechniken sowie Monitoring und die Erstellung von Förderplänen. Sollten die Schüler*innen auf diese Förderung ansprechen, so wird die intensivere Förderung auf Stufe II beendet. Sind jedoch keine Verhaltensverbesserungen erkennbar, so wird die Diagnostik und Förderung auf der nächsten Präventionsebene weiter spezifiziert und intensiviert. *Indizierte Prävention* auf der *Förderebene III* ist für Schüler*innen unter Hoch-Risikobedingungen vorgesehen. Dies betrifft etwa 1 bis 5 % aller Schüler*innen (Hillenbrand 2015, S. 197). Für diese Kinder wird eine hochindividualisierte Einzelfallförderung umgesetzt, die durch eine engmaschige, mehrmals täglich stattfindende Verhaltensverlaufsdiagnostik zur Überprüfung des Fördererfolgs begleitet wird. Oft findet die Unterstützung in spezialisierten Settings, etwa Kleingruppen oder speziellen Klassen statt. Daneben ist die intensive Kooperation mit dem Bezugssystem (Elternhaus, Wohngruppe, o.Ä.), weiteren professionellen Fachkräften und spezialisierten Fachdiensten, wie der Kinder- und Jugendhilfe sowie Kliniken, notwendig. Evidenzbasierte Maßnahmen auf Förderebene III können etwa Selbst-Monitoring, differentielle Verhaltensdiagnostik, individualisierte Förderprogramme oder die Aufstellung eines Krisen-/Interventionsplans sein (Hillenbrand 2015, S. 197). In Abbildung 3 sind evidenzbasierte Maßnahmen zur emotional-sozialen Unterstützung strukturiert nach dem RTI-Ansatz dargestellt.

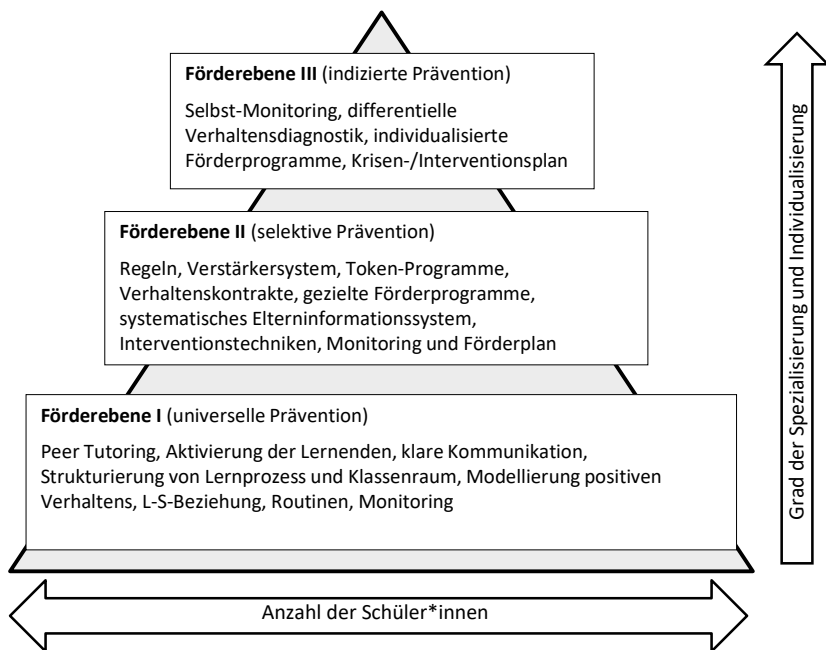


Abbildung 3: Das Response-to-Intervention-Modell mit den nach Intensität und Spezifität gestuften evidenzbasierten Maßnahmen zur emotional-sozialen Unterstützung (Blumenthal et al. 2020, S. 50; Hillenbrand 2015, S. 199)

Die Umsetzung des RTI-Modells ist auch in Deutschland erprobt worden (Mahlau et al., 2016). Im Rahmen des *Rügener Inklusionsmodells (RIM)* (Hartke, 2017) wurde der Ansatz für das deutsche Schulsystem adaptiert und dessen Umsetzung mehrere Jahre lang wissenschaftlich begleitet. Nach vier Schuljahren konnte in der Experimentalgruppe eine Reduktion des Problemverhaltens der Schüler*innen, ein verbessertes prosoziales Verhalten und eine als besser eingeschätzte emotional-soziale Schulerfahrung (Klassenklima, soziale Integration und Selbstkonzept der Schulfähigkeit) nachgewiesen werden (Blumenthal et al., 2020; Voß et al., 2016). Zwei zentrale Kernelemente von RTI sind

- der Einsatz evidenzbasierter Verfahren im Unterricht und
- eine entwicklungsbezogene, engmaschige Begleitdiagnostik (Hillenbrand, 2015).

Da beide Elemente auch im vorliegenden Professionalisierungskonzept umgesetzt wurden, sollen sie im Folgenden anhand eines konkreten Beispiels für die Förderebene II im Bereich ESE exemplarisch vorgestellt werden.

Einsatz evidenzbasierter Verfahren: 84 Handlungsmöglichkeiten

Evidenzbasierte Verfahren auf der Förderebene II werden unterrichtsintegriert oder im Kleingruppensetting durchgeführt. Eine kompakte Sammlung und Übersicht über theoretisch fundierte, praxisnahe Fördermaßnahmen, die sich ausdrücklich für die unterrichtsintegrierte selektive Prävention eignen (Blumenthal et al. 2020, S. 64ff.; Hartke et al. 2018, S. 8), bietet das Buch „Schwierige Schüler – Förderschule: 84 Handlungsmöglichkeiten bei Verhaltensauffälligkeiten und sonderpädagogischem Förderbedarf (1. bis 10. Klasse)“ von Hartke et al. (2018) – im Folgenden *84 Handlungsmöglichkeiten* genannt. Bei dem Fördermanual handelt es sich um ein Gesamtkonzept aus Diagnostik, Förderplanung und Evaluation zur Gestaltung und Anleitung eines unterstützenden Lehrkraftverhaltens. Anhand seiner Konzeption vereinen die *84 Handlungsmöglichkeiten* beide Faktoren schulbasierter Prävention, die theoretische Fundierung und die evidenzbasierte Praxis. Mittels des standardisierten Einschätzungsbogens *Schulische Einschätzung des Verhaltens und der Entwicklung (SEVE)* wird der*die Schüler*in zunächst eingeschätzt. Die Auswertung des Diagnosebogens führt dann direkt zu einem der Förderbereiche *Arbeitsverhalten, Aggression, Angst, Schulabsentismus* und *soziale Integration* sowie zu konkreten Handlungsmöglichkeiten für diese Förderbereiche. Letztere basieren auf den theoretisch fundierten pädagogisch-psychologischen Handlungsansätzen der Lerntheorie („lerntheoretische Handlungsmöglichkeiten“), Kognitionspsychologie („kognitionspsychologische Handlungsmöglichkeiten“) und der humanistischen Psychologie⁶ („schüler*innenzentrierte Handlungsmöglichkeiten“). Die Auswahl einer Handlungsmöglichkeit sollte dabei explizit immer mit ‚pädagogischem Augenmaß‘ erfolgen (interne und soziale Evidenz): „Die Umsetzung einer konkreten Maßnahme im Unterricht verlangt einen planenden und verantwortlich handelnden Praktiker, der die einzelne Situation mit den speziellen Bedingungen erfasst und berücksichtigt“ (Hartke et al. 2018, S. 8ff.). Im Folgenden werden einige der lerntheoretisch und kognitionspsychologisch begründeten Handlungsmöglichkeiten, die den Großteil der Fördermaßnahmen im Manual darstellen, exemplarisch vorgestellt. Diese wurden Blumenthal et al. (2020, S. 64ff.) entnommen.

- 6 Die Humanistische Psychologie gilt als eine der drei Hauptströmungen in der Psychologie. Im Kontrast zu den Konzepten der Psychoanalyse und des Behaviorismus, welche sich in erster Linie auf einzelne psychische Störungen, ihre Ursachen und Behandlungsmöglichkeiten konzentrieren, betrachtet die humanistische Psychologie den Menschen als Ganzes. Demnach beeinflusst die gesamte Lebensumgebung eines Menschen seine Psyche, so etwa seine Lebensumstände, frühere Erfahrungen, seine Bedürfnisse, Umweltfaktoren und das soziale Umfeld (Mutzeck 2008, S. 48 ff.).

Beispiele lerntheoretisch begründeter Handlungsmöglichkeiten
(Hartke et al., 2018)

- Anregung positiver selbstständiger Aktivitäten: Können sich Schüler*innen selbstständig unterrichtsbezogen in einem anregungsreichen Klassenraum beschäftigen, ist die Wahrscheinlichkeit gering, dass sie Langeweile entwickeln und unerwünschte Verhaltensweisen zeigen. Zeigt sich das durch die Anregungen intendierte Verhalten, ist es positiv zu verstärken. Im günstigsten Fall wirkt es selbstverstärkend, da es den Schülerinnen und Schülern Spaß bringt und Erfolgserlebnisse entstehen.
- Klare Instruktion: Mithilfe präzise formulierter Aufforderungen wird Schüler*innen ein Zielverhalten nahegebracht. Klare Instruktionen können sowohl in Unterrichtssituationen, wie bei einem Wechsel von Arbeitsphasen, bei Arbeitsanweisungen oder dem Erklären von Regeln, als auch in Konfliktsituationen eingesetzt werden. Oft ist es hilfreich, das beschriebene Verhalten probeweise ausführen zu lassen, ggf. ist es dabei zu korrigieren und anschließend durch soziale Anerkennung zu verstärken.
- Das Kind erwischen, wenn es gut ist: Zeigen Schüler*innen während des Unterrichts oder der Pause ein positives Verhalten, wird dieses wertgeschätzt. So werden sie angeregt, diese Verhaltensweise häufiger zu zeigen bzw. beizubehalten. Auf diese Weise können Selbstbewusstsein und Selbstbild, das Verhalten in Lern-, Regel- oder sozialen Situationen sowie Kreativität positiv beeinflusst werden.
- Logische Konsequenzen: Stört ein Kind den Unterricht oder fällt es durch destruktive oder aggressive Verhaltensweisen auf, sind logische Konsequenzen ein Mittel, um ein unerwünschtes Verhalten abzubauen. Durch einen offensichtlichen Zusammenhang zwischen Verhalten und der unmittelbaren Reaktion der Lehrkraft wird das Kind darin unterstützt, dieses Verhalten zu beenden und zu reflektieren. Dies und die Klarheit über zukünftige Konsequenzen unterstützen die Schüler*innen dabei, dieses Fehlverhalten zukünftig zu unterlassen.
- Geplantes Ignorieren: Unerwünschtem Verhalten, welches spontan und aus einer Situation heraus entsteht, wie bspw. Schnipsen, um Aufmerksamkeit zu erregen, kann mit geplantem Ignorieren begegnet werden. Durch die Nichtachtung kann verhindert werden, dass sich das Fehlverhalten verfestigt. Zudem kann ein Verhalten, das fälschlicherweise positiv verstärkt wurde (z.B. das Eingehen auf Wortbeiträge ohne Beachtung von Gesprächsregeln), durch ein geplantes Ignorieren wieder abgebaut werden.
- Positive Konditionierung des Klassenraums: Durch bewusst herbeigeführte angenehme Ereignisse im Klassenraum, Erlebnisse voller Entspannung und Freude, kann dieser im Lauf der Zeit zu einem Ort werden, der eine positive Grundstimmung hervorruft.

- Lehrkraft als positives Verhaltensmodell: Indem die Lehrkraft das Kind über das erwartete Verhalten aufklärt, eine bestimmte Verhaltensweise explizit demonstriert und Übungssituationen schafft, kann sie dieses Verhalten sukzessiv aufbauen und festigen. Dieses Vorgehen eignet sich z.B. zur Vermittlung von Arbeitsabläufen oder von sozial angemessenem Rückmeldeverhalten.

Beispiele kognitionspsychologisch begründeter Handlungsmöglichkeiten (Hartke et al., 2018)

- Selbstkontrollkarten für den Unterricht: Selbstkontrollkarten unterstützen Schüler*innen in ihrer Selbstwahrnehmung und -beobachtung und helfen ihnen, sich gewünschte Verhaltensweisen anzueignen. Vorab werden ein Ziel sowie der Zeitraum der Beobachtung auf einer Kontrollkarte notiert (z.B. pro Unterrichtsstunde mindestens zwei Wortmeldungen im Unterrichtsgespräch). Die Beobachtungen werden anschließend von der Schülerin bzw. dem Schüler ebenfalls auf der Karte (z.B. in Tabellenform) festgehalten und hinsichtlich der Zielerreichung bewertet. Am Ende des Beobachtungs- oder Bewertungszeitraums (z.B. ein Schultag oder eine Schulwoche) verstärkt der Schüler oder die Schülerin sich selbst entsprechend vorher festgelegter Regeln oder es findet eine Verstärkung durch die Lehrkraft statt.
- Kognitives Modellieren und Selbstinstruktion einer allgemeinen Problemlösungsstrategie: Erfordert die Lösung einer komplexen Aufgabe ein gut überlegtes Lösungsverhalten, zeigen sich schnell die Grenzen ungestümer Arbeitsweisen und gering entwickelter Problemlösekompetenzen. Für die (Weiter-)Entwicklung absichtsvoller Handlungsmuster demonstriert die Lehrkraft eine Tätigkeit und begleitet diese sprachlich durch erläuternde Kommentare. Im Anschluss löst die Lehrkraft mit dem Kind entsprechend dem demonstrierten Lösungsweg eine ähnliche Aufgabe. Dabei wird eine Formulierung zur Beschreibung des Lösungsweges gesucht, die als Grundlage für die selbstständige Ausführung der Problemlösestrategie durch den*die Schüler*in dienen kann. Ist die ‚Problemlöseformel‘ gefunden, übt das Kind mit ihr, Aufgaben zu lösen. Hierbei ist oft eine Visualisierung der Problemlöseformel hilfreich.
- Die Stopptechnik: Die Stopptechnik ist eine Wenn-Dann-Vereinbarung mit sich selbst. Für Schüler*innen, die schnell wütend und handgreiflich werden, könnte sie lauten: „Wenn ich merke, ich werde ärgerlich, dann sage ich zu mir Stopp und zähle ganz leise und langsam bis zehn, beruhige mich und überlege, was ich tun kann, um eine friedliche Lösung zu finden.“ Für Schüler*innen, die sich gern unerlaubt Sachen anderer Schüler*innen aneignen, könnte die Vereinbarung lauten: „Wenn ich etwas unerlaubt in die Hand nehmen möchte, dann sage ich zu mir innerlich Stopp und stecke die Hände in die Hosentasche.“

- Reflektierendes Krisengespräch: Ein reflektierendes Krisengespräch kann nach der Verletzung einer Regel sowie nachfolgenden haltlosen Begründungen seitens eines Schülers/einer Schülerin hilfreich sein. Die Lehrkraft thematisiert das vorangegangene Verhalten und überlegt mit dem Kind bzw. Jugendlichen, welches Benehmen angebracht gewesen wäre. Problematische Sichtweisen, die als Erklärung für ein Fehlverhalten dienen, werden in diesem Zusammenhang ebenfalls diskutiert. Das Resultat des Gesprächs sollen alternative Handlungsmuster sein, mit denen die beteiligten Schüler*innen zukünftig eine Konfliktsituation bewältigen.

Die Wirksamkeit einer Vorgängerversion der *84 Handlungsmöglichkeiten* (Hartke & Vrbán, 2011) konnte in einer Studie mit 180 Schüler*innen der dritten Klasse nachgewiesen werden. Im Rahmen der Studie wurden vor allem lerntheoretische Ansätze wie etwa die Verhaltensmodifikation mit Verstärkersystemen und Verhaltensverträge eingesetzt und deren Effektivität in Bezug auf die schulische Förderung überprüft. Nach drei Monaten zeigten die Kinder einen deutlichen Rückgang von Verhaltensproblemen sowie bessere Lernfortschritte, insbesondere im Lesen. Eine erste empirische Basis für die externe Evidenz der *84 Handlungsmöglichkeiten* ist somit gelegt. Weitere empirische Studien müssen die Datenlagen jedoch ergänzen.

Entwicklungsbezogene Begleitdiagnostik: Direct Behavior Rating

Die entwicklungsbezogene, engmaschige Begleitdiagnostik ist ein zweites wesentliches Element des RTI-Ansatzes. Eine Prozessdiagnostik, die den Fördererfolg präventiver bzw. interventiver Maßnahmen während ihres Einsatzes überprüft, ermöglicht die direkte Anpassung der pädagogischen Maßnahmen. Für den Bereich Verhalten existieren bisher kaum veränderungssensitive Instrumente, die Fortschritte im Verhalten adäquat abbilden. Das sogenannte *Direct Behavior Rating* (DBR; Christ et al., 2009) ist ein solches Verfahren zur Verhaltensfortschrittsdiagnostik, das als hoch kompatibel mit RTI-basierten Mehrebenen-Förderprogrammen eingeschätzt wird (Daniels et al., 2017). Das Messverfahren stellt eine Kombination aus direkter systematischer Verhaltensbeobachtung und einer Verhaltensbeurteilung mit Hilfe einer Ratingskala dar. Die Skala kann nominalskalierte, ordinalskalierte oder intervall- bzw. verhältnisskalierte Merkmale abbilden und umfasst in der Regel 6-11 Stufen (Chafouleas, 2011; Huber, 2015a). DBR-Skalen können den prozentualen Anteil, die Qualität oder die Häufigkeit eines Verhaltens abbilden. Die Wahl der Ratingskalen ist vor allem von dem zu beobachtenden Zielverhalten abhängig zu machen. Da das DBR-Verfahren effizient und flexibel ist, eignet es sich gut für die Messung von Verhaltensfortschritten im Rahmen kontrollierter Einzelfallstudien. Die Beurteilung eines Verhaltens erfolgt dabei immer in oder direkt nach der Situation, in der das Verhalten beobachtet werden soll (Huber & Rietz, 2015), wobei die Beobachtungszeit-

räume zwischen wenigen Sekunden bis hin zu einem Tag andauern können. Somit kann auch mehrfach täglich ein gut beobachtbares Verhalten eingeschätzt werden (Christ et al., 2009).

Grundsätzlich lassen sich zwei Formen von DBR unterscheiden: *Single-Item-Scales* und *Multi-Item-Scales* (Christ et al., 2009). *Single-Item-Scales* (SIS) erfassen mit nur einem einzigen Item eine übergeordnete Verhaltensdimension, etwa den globalen Verhaltensausschnitt Störverhalten oder Arbeitsverhalten. Dies bietet sich besonders an, wenn ein eher breit gefächertes Verhaltensproblem beurteilt werden soll. Sehr konkrete und spezifische Verhaltensweisen („arbeitet in Stillarbeitsphasen ruhig und konzentriert“ oder „meldet sich vor Wortbeiträgen“) sind durch SIS hingegen schwer zu erfassen. Mitunter sind solche konkreten Verhaltensinformation aber notwendig, um den individuellen Fördererfolg beurteilen zu können (Casale et al., 2017) bzw. bisher verdeckt gebliebene Verhaltensmuster sichtbar zu machen (Mahlau & Erdmann, 2020). Für diese Zwecke eignet sich die *Multi-Item-Scale* (MIS) besser, da diese in der Regel drei bis fünf spezifische und konkrete Verhaltensweisen zur Erfassung einer übergeordneten Verhaltensdimension operationalisiert (Casale et al., 2017).

Die Validität und Reliabilität von DBR hängen in erheblichem Maße von dem zu beobachtenden Zielverhalten ab. Festhalten lässt sich zunächst: Mit DBR lassen sich keine soliden statusdiagnostischen Aussagen treffen, denn DBR-Rater*innen schwanken erheblich in ihren Einschätzungen über ein und denselben Verhaltensbereich und weichen außerdem im Hinblick auf ihr Urteil von professionellen Verhaltensbeobachter*innen ab (Christ et al., 2009; Huber, 2015a; Volpe & Briesch, 2012). Stattdessen eignen sich DBR aber gut für die Verlaufsdagnostik, da die längsschnittlichen Veränderungen sowohl von DBR-Anwender*innen als auch professionellen Verhaltensbeobachter*innen vergleichbar beurteilt werden (Huber & Rietz, 2015; Riley-Tillman et al., 2011). Wichtig ist jedoch, dass die Verhaltensbeurteilungen immer von den gleichen Personen stammen, nur dann spiegeln die Veränderungen der Beobachtungsdaten eine gute Schätzung der tatsächlichen Verhaltensentwicklung wider (Huber, 2015a).

Bei den SIS, die einen übergeordneten Verhaltensbereich messen, lag die höchste Interrater-Reliabilität bisher bei den Beobachtungsbereichen *Teilnahme am Unterricht*, *störendes Verhalten* und *respektvolles Verhalten* (Chafouleas, 2011; Chafouleas et al., 2013; Christ et al., 2009; Riley-Tillman et al., 2008, 2009, 2011), die aus diesem Grund auch als die *Big 3* der SIS (Chafouleas, 2011) bezeichnet werden. Die Beurteilung mit hoch-inferenten Items, die ein hohes Ausmaß an Schlussfolgerungen auf Seiten der Beobachter*innen erfordern (Lotz et al., 2013), ist jedoch per se weniger zuverlässig als mit wenig-inferenten Items (Cone, 1977); dies zeigt sich – wenn auch nicht widerspruchsfrei – beim Vergleich zwischen SIS und MIS. Während die Studie von Riley-Tillman et al. (2009) konstatiert, dass die Formulierung

der Items keinen substanziellen Einfluss auf die Zuverlässigkeit der Messung hat, wurde in anderen Studien genau dies widerlegt (Casale et al., 2015, 2017; Christ et al., 2009) bzw. dahingehend differenziert, dass eine zuverlässige Einschätzung nur dann geleistet werden kann, wenn die Rater*innen genau wissen, wie sich die Verhaltensweisen äußern (Conley et al., 2014) und demnach die globalen Verhaltensbereiche, zumindest implizit, operationalisiert sind. Die MIS sind grundsätzlich stärker operationalisiert und somit niedrig-inferent. Zahlreiche Studien weisen auf eine gute bis sehr gute Genauigkeit, Reliabilität, Validität und Änderungssensitivität der DBR-MIS hin (Casale et al., 2021; Chafouleas et al., 2012; Matta et al., 2020; Smith et al., 2018; Volpe & Briesch, 2015). Die psychometrische Qualität von DBR-MIS konnte insbesondere für externalisierendes Verhalten nachgewiesen werden (Casale et al., 2017). Für die Operationalisierung des zu beurteilenden Verhaltens gilt sowohl für SIS als auch für MIS: Zur Erhöhung der Beobachtungsgüte sollte vorrangig die Anwesenheit und weniger die Abwesenheit eines Verhaltensausschnitts definiert werden (Chafouleas et al., 2013; Huber, 2015a). Soll beispielsweise der Bereich *störendes Verhalten* beobachtet werden, so ist demnach eine negative Verhaltensformulierung zu wählen (Mahlau & Erdmann, 2020), z.B. „Ruft in den Unterricht herein, ohne sich zu melden“.

Die Angaben über die erforderliche Anzahl an Beurteilungen zur Erlangung einer akzeptablen Beobachtungsgüte schwanken stark. Huber und Rietz (2015) konnten im Rahmen eines systematischen Reviews unterschiedlicher methodischer Studien zu DBR zeigen, dass bei weniger als fünf Messzeitpunkten die Messwerte zu stark zwischen den Rater*innen streuen und somit mehr Varianz durch die Rater*innen erklärt wird als durch die Verhaltensunterschiede der beobachteten Personen. Bei entscheidungsrelevanten Fragestellungen empfehlen einige Autor*innen sicherheitshalber sogar die Zusammenfassung von bis zu 20 Messwerten (Volpe & Briesch, 2012). Casale et al. (2017) konnten in Bezug auf externalisierendes Verhalten aufzeigen, dass sich bereits nach 13 Messungen mit DBR-MIS zuverlässige Aussagen über die Verhaltensentwicklung treffen lassen. Ein systematisches Review von Studien zu DBR (Huber & Rietz, 2015) zeigt außerdem, dass sich Validität und Reliabilität von DBR weiter steigern lassen, wenn die Rater*innen zuvor ein kurzes Training durchlaufen haben. Hinweise auf eine Verbesserung der Beobachtungsgüte durch intensive Trainingsphasen gab es bislang aber nicht.

2.1.4 Zusammenfassung und Schlussfolgerungen für die schulische Förderung der ESE von Kindern

Aufgrund der hohen Prävalenz-, Persistenz- und Komorbiditätsraten sowie sich verfestigender, kumulierender Risikokonstellationen bei emotional-sozialen Entwicklungsstörungen (u.a. Beelmann & Raabe, 2007; Ihle & Esser, 2008; Klasen et al., 2017; Otto et al., 2021; Visser et al., 2019) stehen bei der schulischen Förderung der ESE präventive Maßnahmen im Vordergrund. Schulen bieten sich für die präventive Förderung besonders an. Dies gilt insbesondere für Kinder unter Risikobedingungen, die anhand schulischer Maßnahmen wesentlich besser erreicht werden können als durch außerschulische. Eine Diagnostik erfolgt jedoch in aller Regel erst dann, wenn eine erhebliche Problematik deutlich wurde (*Wait-to-fail-Problem* und *Etikettierungs-Ressourcen-Dilemma*). Wünschenswert wäre, dass es gar nicht erst zur Verfestigung von Problemen kommt. Das Ziel schulischer Förderung im Bereich ESE muss demnach sein, im Schulalltag präventiv emotional-soziale Kompetenzen zu fördern, um sich anbahnende Entwicklungsstörungen zu vermeiden oder bereits bestehende zu vermindern oder ganz abzubauen. Die dafür erforderliche stark präventive pädagogische Arbeit im emotional-sozialen Bereich, zu der die Früherkennung sich anbahnender Probleme und die daran anknüpfende Förderplanung, -durchführung und -evaluation gehören, macht eine erhebliche Sensibilität aller Pädagog*innen für ihre Schüler*innen erforderlich (Stein & Müller, 2015). So fordert die evidenzbasierte Praxis neben der externen, wissenschaftlichen Evidenz von Fördermaßnahmen und deren individuelle Passung an die Werte, Wünsche und Bedarfe der Schüler*innen (soziale Evidenz), vor allem auch die praktische Expertise der Anwender*innen (interne Evidenz). Als eines der Desiderate zur Implementation von EbP identifiziert auch Hillenbrand die Professionalität der Lehrkräfte. Er argumentiert, dass die Umsetzung eines potentiellen Wissens nicht von selbst gelingt, sondern

ein reflektiertes, wissenschaftlich geschultes und selbstkritisches Vorgehen in der Informationsgewinnung wie auch in der Implementation evidenzbasierter Verfahren [erfordert]. Angesichts des Kenntnisstands von (potentiellen) Lehrkräften zeigen sich hier große Herausforderungen für die zukünftige Lehrerbildung in allen Phasen (Hillenbrand 2015, S. 209).

Die wirksame Umsetzung präventiver schulischer Förderung der ESE von Kindern ist demnach eng an die professionellen Kompetenzen der Lehrkräfte (Baumert & Kunter, 2006) geknüpft. Diese müssen durch Aus- und Weiterbildung gestärkt werden (Stein & Müller 2015, S. 33).

2.2 Professionalisierung von Lehrkräften für die schulische Förderung der ESE von Kindern

Im folgenden Kapitel soll es um die Frage gehen, wie Regelschullehrkräfte in der dritten Phase der Lehrer*innenbildung für inklusiven Unterricht im Bereich des FSP ESE erfolgreich fortgebildet werden können. Ausgehend von einer kurzen etymologisch-historischen und soziologischen Einordnung der Begriffe *Profession*, *Professionalität* und *Professionalisierung*, soll das Konstrukt der *Professionalisierung von Lehrkräften* genauer betrachtet werden. Kompetenzorientierte Ansätze von Professionalisierung stehen hierbei im Fokus. Anhand des Modells *professioneller Handlungskompetenz* (Baumert & Kunter, 2006) und dessen Erweiterungen für inklusiven Unterricht (Blömeke et al., 2015; Gottfried et al., 2021; König et al., 2019; Leko & Roberts, 2014; Osher et al., 2014) wird der Handlungsrahmen definiert, unter dem sich die Konzipierung des in dieser Arbeit im Fokus stehenden Professionalisierungskonzepts für Grundschullehrkräfte zur Förderung der ESE ihrer Schüler*innen einordnen lässt. Neben der Modellierung professioneller Handlungskompetenz anhand des in diesem Zusammenhang relevanten Wissensbereichs *pädagogisch-psychologisches Wissen* von Lehrkräften wird auch dargestellt, wie sich Lehrkraftkompetenz im Zuge beruflichen Lernens (weiter-)entwickelt (Berliner, 2004; König, 2010; Kunter et al., 2011) und wie die spezifischen Kompetenzbereiche im Rahmen von Lehrkraftfortbildungen gezielt gefördert und untersucht werden können (u.a. Lipowsky, 2010). Der theoretischen Einordnung folgen empirische Erkenntnisse zu wirksamen Fortbildungsangeboten und deren Reichweite (u.a. Darling-Hammond et al., 2017; Guskey, 2014; Lipowsky, 2010; Lipowsky & Rzejak, 2012; Timperley et al., 2007). Dies soll Aufschluss darüber geben, wie Fortbildungsangebote grundsätzlich konzipiert sein sollten, welche generellen Wirkfaktoren auszumachen sind und wie Fortbildungen auf ihre Wirksamkeit hin überprüft werden können. Abschließend werden mit Blick auf Untersuchungen zur Wirksamkeit von Lehrkraftfortbildungen im Kontext der Inklusion im FSP ESE (u.a. Leidig, 2019; Leidig et al., 2016) konkrete Qualitätsmerkmale von Fortbildungsangeboten für dieses Professionalisierungsfeld herausgestellt.

2.2.1 Das Konstrukt Lehrkraftprofessionalisierung

Eine Möglichkeit, sich dem Konstrukt der Lehrkraftprofessionalisierung zu nähern, ist die etymologisch-historische sowie soziologische Betrachtung der Konzepte Profession, Professionalität und Professionalisierung. Der Wortstamm aller drei Begriffe geht auf das lateinische Wort *professio* zurück und bedeutet so viel wie Beruf, Gewerbe, Berufung (Pfeifer, 1993). Der klassi-

sche Begriff der Profession im soziologischen Sinn stellt einen Ordnungsmodus moderner Gesellschaften dar, dem je nach Ausrichtung unterschiedliche Eigenschaften zugeschrieben wurden (Berkemeyer et al. 2012, S. 118). Das traditionelle berufssoziologische Professionen-Modell⁷ grenzt die Profession („free professions“) – dazu zählten vorrangig Ärzt*innen, Pfarrer*innen, Jurist*innen und Architekt*innen – von den gewöhnlichen Berufen („semi-professional“) ab. Der Begriff der Profession wurde maßgeblich dafür genutzt, Berufe anhand bestimmter Kriterien von anderen Berufen zu unterscheiden (Horn 2016, S. 154). Als Merkmale der Profession wurden die akademische Ausbildung und der Erwerb spezifischen Wissens, das Befassen mit komplexen und riskanten technischen, wirtschaftlichen, sozialen und/oder humanen Problemlagen sowie mit spezifischen Dienstleistungen für Individuen und die Gesellschaft herausgestellt. Ferner galten die Existenz eigener Organisationen und Verhaltensregeln sowie die Autonomie gegenüber Staat und Klientel als typische Merkmale der Profession (Horn, 2016). In diesem Kontext bezeichnet Professionalisierung den sozialen Durchsetzungsprozess des Hinaufsteigens eines gewöhnlichen Berufs in den Status einer Profession (Terhart 2011, S. 203). Pädagogische Berufe, und insofern auch der Beruf der Lehrkraft, wurden als „semi-professional“ eingestuft, da sie die Kriterien für eine „free profession“ nicht vollständig erfüllten (Horn, 2016).

In der aktuellen erziehungswissenschaftlichen Diskussion treten viele der soziologischen Grundannahmen, bspw. über Professionalisierung als kollektiven Prozess, in den Hintergrund, da das Verständnis von Professionen und Professionalität – nicht zuletzt aufgrund der realen Entwicklungen in den ehemals klassischen Professionen – als veraltet gilt:

Die traditionellen Professionen haben sich in großen Teilen zu gewöhnlichen Berufen mit straffer organisatorischer Gängelung und typischer Angestelltenmentalität zurückentwickelt (Deprofessionalisierung), andere Berufe haben sich in neue Statusdimensionen, Tätigkeitsfelder, Ausbildungsformen etc. hinein bewegt, die mit dem klassischen Professionenkonzept analytisch nicht zu fassen sind. Der ununterbrochene Wandel der Struktur und des Status von Berufen bzw. des gesamten Berufssystems selbst zwingt grundsätzlich zu einer ständigen Anpassung berufssoziologischer Modelle, Typenbildungen und Kategorisierungen (Terhart 2011, S. 203).

Die jüngere Professionstheorie geht davon aus, dass „Professionelle spezifische Aufgaben lösen müssen, für die sie eine an Standards orientierte Expertise entwickeln und dabei gleichzeitig ein professionelles Selbstverständnis oder professionelles Ethos ausbilden“ (Berkemeyer et al. 2012, S. 118). Da-

7 Es ist darauf zu verweisen, dass zwischen der anglo-amerikanischen und der kontinentaleuropäischen Tradition sowohl bezüglich der Berufsgeschichte als auch der Begriffe in diesem Feld Unterschiede existieren (Horn, 2016).

nach kann auch der Lehrberuf als Profession verstanden werden (Kunter et al. 2011, S. 30), zeichnet er sich gegenüber anderen Berufen doch dadurch aus, dass sich die Berufspraxis – neben einem hohen Grad an Freiheitsspielraum und Eigenverantwortung – nicht in routinemäßigen Handlungen in vorhersehbaren Situationen erschöpft. Stattdessen sind hohe Reflexionsfähigkeit, Expertise, eine theoretische Wissensbasis sowie spezialisierte Fertigkeiten nötig, um das (Professions-)Wissen auf nur begrenzt vorhersehbare, sich stetig wandelnde Situationen angemessen anwenden zu können (Kunter et al., 2011; Shulman, 1998). Der Begriff der Professionalität kann deshalb als ein dauerhafter Anspruch an den Lehrkraftberuf und der der Professionalisierung als „die dauerhafte, dort hinführende Aufgabe“ (Hugger et al. 2021, S. 11) verstanden werden.

Wendet man sich dem Konstrukt der *Lehrkraftprofessionalisierung* genauer zu, so wird ersichtlich, dass sich hierbei unterschiedliche Konzepte herausgebildet haben. Im deutschsprachigen Raum bezeichnet Professionalisierung recht allgemein den individuellen berufsbezogenen Entwicklungs- und Lernprozess von Lehrkräften im Sinne des beruflichen Lernens, der grundsätzlich in allen Phasen der Lehrer*innenbildung stattfindet (Gräsel & Trempler 2017, S. 2). In der englischsprachigen Literatur haben sich im Zuge des Professionalisierungsdiskurses die Begriffe *Professional Learning* und *Professional Development* herausdifferenziert (Durksen et al., 2017). Auch wenn beide Begrifflichkeiten teils synonym verwendet werden, so können maßgeblich folgende Unterschiede herausgestellt werden: Professional Learning referiert auf das individuelle Lernen der Lehrkräfte als einen selbst verantworteten und gestalteten aktiven Prozess sowie dessen Produkt und findet vorwiegend in informellen Lerngelegenheiten als situativer Lernprozess statt (Durksen et al. 2017, S. 53f.; Timperley 2011, S. 4). Professional Development bezieht sich auf Lernangebote, die für die Lernenden entwickelt und diesen angeboten werden als ein „set of coordinated, comprehensive, and intensive activities designed to enhance educators‘ knowledge, beliefs, skills, and practices for the purpose of improving student outcomes“ (Leko & Roberts 2014, S. 43). Professional Development kann in formalen (z.B. Weiterbildungskurse und/oder Fortbildungsveranstaltungen⁸) oder non-formalen (z.B. kollegiale Unterrichtshospitation, gegenseitige Qualifizierung im Team, Onlinekurse, Literaturstudium, etc.) Situationen realisiert werden (Fussangel et al., 2016; Lipowsky, 2014; Lipowsky & Rzejak, 2015). Der in dieser Arbeit verwendete Begriff der Professionalisierung impliziert sowohl das Kon-

8 In Deutschland wird in der Regel zwischen Fort- und Weiterbildung unterschieden (z.B. Richter 2016, S. 246). Während eine Weiterbildung mit dem Ziel der formalen Erweiterung der Berufsqualifikation (weiteres Unterrichtsfach, Qualifikation für eine Leitungsfunktion) verbunden ist (Fussangel et al., 2016, S. 362; Richter, 2016, S. 246), fokussieren Fortbildungen auf die „Unterstützung der Lehrkräfte in ihrer aktuellen Berufspraxis und auf Qualifikationserhaltung“ (Fussangel et al., 2016, S. 362).

zept des Professional Learning als Prozess und Produkt beruflichen Lernens von Lehrkräften als auch den des Professional Development, wobei letzterer sich auf konkrete Lernangebote in Form von Fortbildungen⁹ bezieht. Im Folgenden werden einzelne formale Fortbildungsangebote als *Fortbildungsveranstaltungen*, mehrere gezielt gebündelte formale Fortbildungsangebote als *Fortbildungseinheit* und/oder *-modul* und alle zusammengefassten Fortbildungsangebote im Rahmen eines Projekts oder zeitlich abgesteckten Rahmens als *Fortbildungsmaßnahme/n* bezeichnet.

Das Verständnis von *Lehrkraftprofessionalisierung* und die Forschung dazu unterscheiden sich vor dem Hintergrund der jeweiligen theoretischen Perspektive (Gräsel & Trempler 2017, S. 2). Lange Zeit galt der Lehrberuf als Kunst oder Handwerk, für das Lehrkräfte bestimmte Persönlichkeitseigenschaften, -merkmale oder Talente mitbringen mussten. Etwa seit den 1990er Jahren wird jedoch zunehmend betont, dass berufsspezifische und v.a. auch erlernbare Merkmale das Handeln und den Erfolg von Lehrkräften bestimmen und dass der Aufbau dieser Merkmale ein professionsspezifischer Lernprozess ist (Wirtz, 2021). Neben einer Reihe von unterschiedlichen Positionen¹⁰ haben sich, speziell für den Lehrkraftberuf, aktuell vor allem drei Ansätze als wirkmächtig herauskristallisiert, um die Professionalität von Lehrkräften zu beschreiben: der *berufsbiographische*, der *strukturtheoretische* und der *kompetenzorientierte* Bestimmungsansatz (vgl. Bonnet & Hericks, 2014; Terhart, 2011). Der *berufsbiographische* Professionsansatz konzeptualisiert Professionalität als „berufsbiographisches Entwicklungsproblem“ (Terhart 2011, S. 56). Er richtet seinen Blick auf individuelle Professionalisierungsprozesse von Lehrpersonen sowie zum Teil auch auf deren Wechselwirkungen mit dem gesamten Lebenslauf, um so in „longitudinaler Perspektive Entwicklungsverläufe des Aufbaus, der Weiterentwicklung und des Verlusts von Professionalität zu rekonstruieren“ (Bonnet & Hericks 2014, S. 7). Der *strukturtheoretische* Professionsansatz (Helsper, 2014) legt besonderes Augenmerk auf das Verhältnis von Organisation/Institution zu individuellem Handeln. Er schärft den Blick für die Krisenhaftigkeit und Widersprüchlichkeiten (Antinomien) des Lehrkrafthandelns, etwa wenn Schüler*innen unter Zwang zur Selbstbestimmung befähigt werden sollen (Autonomie-Antinomie) oder Schüler*innen gefördert, aber gleichzeitig durch die Notenbewertung im Bildungssystem selektiert werden sollen (Förderung-Selektion-Antinomie). Der *kompetenzorientierte* Ansatz legt den Fokus auf die Entwicklung von „Struktur- und Entwicklungsmodelle[n]“ (Bonnet & Hericks

9 Fortbildung wird vielfach auch bedeutungsgleich mit Professional Development verwendet (z.B. Fussangel et al., 2016; Lipowsky, 2014; Lipowsky & Rzejak, 2015).

10 Tillmann (2011) etwa kontrastiert quantitativ orientierte Forschungsansätze mit psychologischem Hintergrund (Persönlichkeitsansatz und Expertenansatz bzw. kompetenztheoretischer Ansatz) und soziologisch geprägte qualitativ rekonstruierende Forschungsansätze (z.B. kultur- bzw. praxistheoretischer Ansatz, strukturtheoretischer Professionsansatz).

2014, S. 5), die die für den Lehrkraftberuf notwendige Handlungskompetenz modellieren. Der Ansatz setzt bei den Kompetenzen von erfolgreichen Lehrkräften und nicht bei den Problemen, Widersprüchlichkeiten und Defiziten des Lehrkraftberufs an. Unterricht – nicht die Erziehung – wird als das „Kerngeschäft“ (Tenorth 2006, S. 585) oder die „Kernaufgabe“ (Blömeke 2002, S. 42) der Lehrkräfte herausgestellt. Ausgehend von einer möglichst genauen Aufgabenbeschreibung des Berufs der Lehrkraft werden Kompetenzbereiche und Wissensdimensionen definiert, die für die Bewältigung dieser Aufgaben wichtig bzw. notwendig sind. Dabei werden Resultate der empirischen Forschung zum professionellen Wissen und Können von Lehrkräften ausgewertet und insbesondere auf ihren „empirisch nachzuweisenden Beitrag zum Erreichen des Zwecks der Institution Schule/Unterrichten bezogen“, nämlich den „nachweisbaren fachlichen und überfachlichen Lernerfolgen“ der Schüler*innen (Terhart 2011, S. 207).

Tillmann (2011) attestiert den unterschiedlichen Ansätzen klare Konfliktlinien und auch Terhart (2011) beschreibt deren Verhältnis als teilweise sich „berührend, [...], teilweise überlappend bzw. inhaltlich kongruent“, aber auch als „teilweise direkt-konfrontativ“ (S. 216). Angesichts der bedeutenden und notwendigen Veränderung des Lehrkraftberufs im Zuge der inklusiven Ausrichtung von Schule kann davon ausgegangen werden, dass ein „Arbeitsbündnis“ der unterschiedlichen Professionalisierungsansätze, so wie Tillmann (2011, S. 315) es fordert, wertvolle Impulse zur neuen Ausrichtung des Professionalisierungsprofils von Lehrkräften geben kann. Insbesondere könnte der strukturtheoretische, neben dem oft im Fokus stehenden kompetenzorientierten Ansatz „wertvolle Impulse für die Professionalisierung von Lehrkräften im Kontext herausfordernder Lehr-Lernsituationen“ (Leidig 2019, S. 18) liefern. Aufgrund der Schwerpunktsetzung der vorliegenden Arbeit steht jedoch der kompetenzorientierte Bestimmungsansatz besonders im Fokus und soll aus diesem Grund im folgenden Kapitel genauer erläutert werden.

2.2.2 Professionelle Handlungskompetenz von Lehrkräften

Der kompetenzorientierte Professionalisierungsansatz stützt sich vor allem auf die Expertise-Forschung (Berliner, 2004; König, 2010) sowie auf die Taxonomie der für den Lehrkraftberuf relevanten Wissensformen nach Shulman (2004) und wird insbesondere seit dem sogenannten PISA-Schock Anfang des 21. Jahrhunderts intensiv diskutiert. Neben kognitiven Kompetenzbereichen berücksichtigt er im Sinne eines weiteren Verständnisses von Kompetenz auch nicht-kognitive Bereiche (Kunter et al., 2009). Die individuellen Voraussetzungen zur erfolgreichen Bewältigung spezifischer beruflicher Aufgaben, die neben Wissen, Motivation und Überzeugungen auch

Fähigkeiten zur Selbstregulation umfassen, werden als professionelle Handlungskompetenz bezeichnet (Baumert & Kunter, 2006; Kunter et al., 2009). Ausgehend vom Modell *professioneller Handlungskompetenz* (Baumert & Kunter, 2006) ist insbesondere das sogenannte *Professionswissen* oder *professionelle Wissen* von Lehrkräften (nach Shulman, 1987) für die vorliegende Studie von Bedeutung und soll im Folgenden vorgestellt werden.

2.2.2.1 Das COACTIV-Kompetenzmodell und seine Erweiterung

Das *COACTIV*¹¹-Kompetenzmodell, das von Baumert und Kunter im Rahmen der PISA-Studie entwickelt wurde (2006), modelliert das Konstrukt professioneller Kompetenz von Lehrkräften (s. Abbildung 4). Hierbei wurden verschiedene, bereits bestehende Konzepte zur Professionalisierung und Qualifizierung von Lehrkräften (Bromme, 1992, 1997; Shulman, 1987; Weinert, 2001) in ein übergreifendes theoretisches Modell zusammengeführt. Das Modell stützt sich auf die Annahme des deutschen Psychologen Franz Weinert (2001), dass Kompetenzen grundsätzlich veränder- und erlernbar sind. Professionelle Kompetenz nach Baumert und Kunter (2006) äußert sich in professionellem Handeln und lässt sich durch vier Dimensionen charakterisieren: Wissen und Können (Professionswissen), Überzeugungen und Werthaltungen, Motivation sowie Selbstregulation. Überzeugungen und Werthaltungen sind Teil des professionellen Handelns und beinhalten subjektive Theorien über Lehren, Lernen und Unterrichtsziele im jeweiligen Fachbereich. Die Dimension Selbstregulation bezieht sich auf den effektiven und schonenden Umgang mit den eigenen Ressourcen. Als Kern der Professionalität von Lehrkräften und fundamental für erfolgreiches Unterrichten erachten Baumert und Kunter (2006, S. 483ff.) die Dimension des Professionswissens, das sich auf das deklarative Wissen („Wissen, dass...“) und prozedurale Können („Wissen, wie...“) der Lehrkräfte bezieht (Bromme 2001, S. 122). Angelehnt an die Theorien des amerikanischen Unterrichtsforschers Lee Shulman und des deutschen Psychologen Rainer Bromme zum professionellen Wissen von Lehrkräften, stützt sich das Professionswissen nach Baumert und Kunter (2006) auf fünf große Kompetenzbereiche: das *Fachwissen*, das *fachdidaktische Wissen*, das *generische pädagogisch-psychologische Wissen*, das *Organisationswissen* und das *Beratungswissen* (s. Abbildung 4). Fundiertes *Fachwissen* wird als Voraussetzung dafür angesehen, fachdidaktisches Wissen zu entwickeln und zu vertiefen. Das *fachdidaktische Wissen* umfasst u.a. das Wissen um das didaktische und diagnostische Potential von Aufgaben, die curriculare Abstimmung der Inhalte, Kenntnis über typische Fehler der Schü-

11 COACTIV steht für das Forschungsprojekt „Cognitive Activation in the Classroom: The Orchestration of Learning Opportunities for the Enhancement of Insightful Learning in Mathematics“.

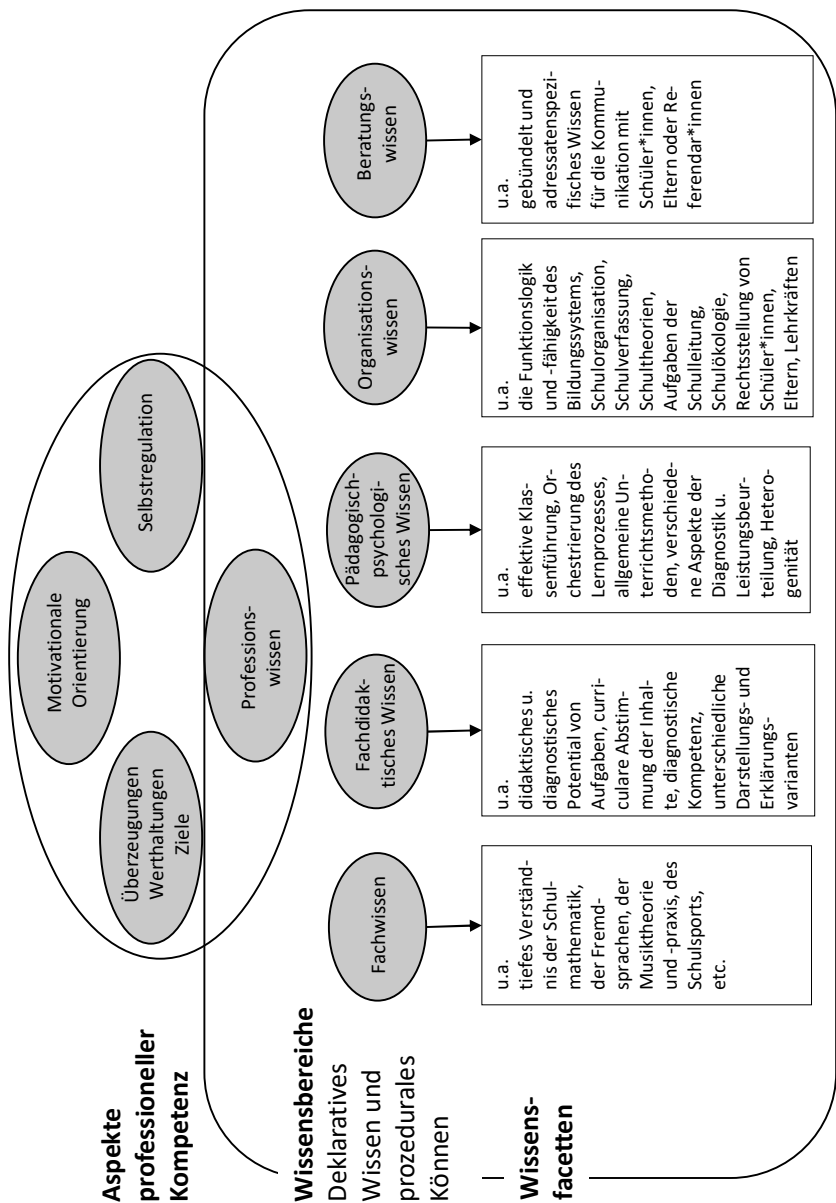


Abbildung 4: Modell professioneller Handlungskompetenz (Baumert & Kunter, 2006)

ler*innen, diagnostische Kompetenz und unterschiedliche Darstellungs- und Erklärungsvarianten. Das *pädagogisch-psychologische Wissensrepertoire* umfasst u.a. die Klassenführung, allgemeine Unterrichtsmethoden, verschiedene Aspekte der Diagnostik, Lernprozesse und Heterogenität. Das *Organisationswissen* bezieht sich auf die Funktionslogik und -fähigkeit des Bildungssystems. Wissensfacetten sind hierbei etwa das Bildungssystem und seine Rahmenbedingungen, Schultheorien und -organisation und die Rechtsstellung von Schüler*innen, Erziehungsberechtigten und Lehrkräften. Auf das fachunabhängige *Beratungswissen* sind professionelle Lehrkräfte im Umgang mit Laien angewiesen, beispielsweise wenn es um die Beratung von Schüler*innen, Eltern oder Referendar*innen geht. Beratungsanlässe können bspw. Schullaufbahneempfehlungen, Lernschwierigkeiten und Verhaltensprobleme sein (Kunter & Baumert, 2011).

Den Kern professioneller Handlungskompetenz bilden Fachwissen, fachdidaktisches Wissen und pädagogisch-psychologisches Wissen. Die Bereiche Organisations- und Beratungswissen gehen über eine individuelle Kompetenzperspektive hinaus (Kunter et al., 2011). Während das Fachwissen und das fachdidaktische Wissen an eine schulische Domäne (Schulfach) gebunden sind, wird dem pädagogisch-psychologischen Wissen fächerübergreifende Bedeutung beigemessen (Gottfried et al., 2021). Bei dem in dieser Arbeit vorzustellenden Professionalisierungskonzept steht das pädagogisch-psychologische Lehrkraftwissen im Vordergrund, da mit der Förderung der ESE von Schüler*innen ein generischer Bereich schulischen Lehrens und Lernens im Mittelpunkt steht.

Das im COACTIV-Modell eher statisch-analytische Verständnis von Kompetenz kann jedoch nicht hinreichend zur erfolgreichen Bewältigung der komplexen Anforderungssituationen im schulischen Kontext und zur Erklärung der dabei ablaufenden Prozesse herangezogen werden. So argumentiert Frey (2014), dass sich Kompetenz nicht nur aus theoretisch vorhandenem Wissen und potentiellen Fähigkeiten zusammensetzt, sondern sich insbesondere auch im Transfer der Dispositionen auf die Handlung in unterschiedlichen Situationen zeigt. Im Rahmen dieser Diskussion schlagen Blömeke et al. (2015) eine Erweiterung des Kompetenzbegriffes um situationsspezifische Prozesse vor. Anhand des *P-I-D model of competence transformation* modellieren sie professionelle Handlungskompetenz als ein Zusammenspiel aus einer *Disposition* (Kognitionen/Professionswissen und affektiv-motivationale Merkmale) und den *situationsspezifischen Skills* Wahrnehmung (perception), Interpretation (interpretation) und Entscheidungsfindung (decision-making). Professionelle Kompetenz zeige sich schließlich über die *Performanz* des Handelns (performance) anhand beobachtbaren Verhaltens (Blömeke et al., 2015). Die im Zuge inklusiver Beschulung relevanten normativen Werte wie Chancengleichheit, Teilhabe, Entwicklung, Bewahrung von

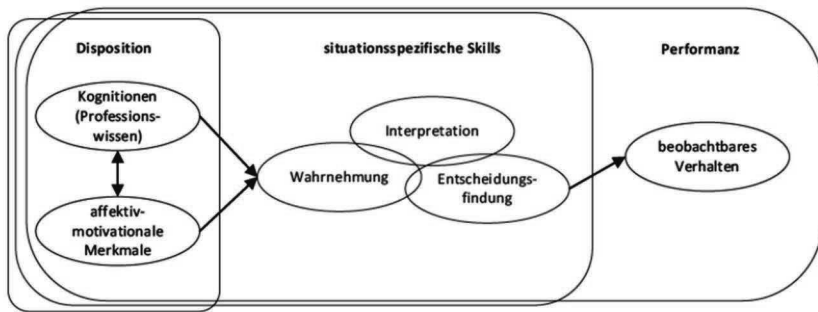


Abbildung 5: Professionelle Handlungskompetenz als Kontinuum; entnommen aus Leidig (2019; modifizierte Darstellung nach Blömeke et al., 2015)

Gemeinschaften und Achtung der Diversität (European Agency 2009, S. 30) zeigen sich, so Leidig (2019, S. 23), vor allem moderiert durch situationsspezifische Skills. Abbildung 5 veranschaulicht den Prozess der professionellen Handlungskompetenz als Kontinuum nach Blömeke et al. (2015) und den von Leidig (2019) ergänzten Interaktionen zwischen den Dispositionen.

Neben der Ergänzung des COACTIV-Kompetenzmodells um die dynamisierende Dimension der situationsspezifischen Skills ist es im Zuge eines sich verändernden Professionalisierungsprofils der Lehrkräfte (Melzer & Hillenbrand, 2015) erforderlich, zusätzliche Erweiterungen des klassischen Kompetenzmodells vorzunehmen. In Anlehnung an existierende Kompetenzmodelle verstehen König et al. (2019, S. 49) professionelles Wissen zur Inklusion als Erweiterung der Wissensbereiche fachdidaktisches, pädagogisch-psychologisches sowie Organisations- und Beratungswissen. Anhand eines systematischen Reviews von 40 Publikationen aus dem deutschen Sprachraum und den darin aufgeführten aktuellen Kompetenzkatalogen zu den Anforderungen im Bereich Inklusion extrahieren König et al. (2019, S. 50) vier große Aufgabenfelder als „Basiswissen oder möglicherweise auch [...] spezifisches Wissen für inklusive Bildung und Erziehung“. Die folgenden erweiterten Anforderungen an Regelschullehrkräfte modellieren sie, wie in Abbildung 6 dargestellt, als Spezifizierung der von Baumert und Kunter (2006) eingeführten professionellen Handlungskompetenz:

1. Diagnose (u.a. kategoriale Förderbedarfe, Eingangsdiagnosen, Fortschrittskontrollen, Curriculum-based-measurements)
2. Intervention (u.a. unterrichtsimmanente Förderung, Einzel- und Gruppenförderung, Aufstellen von Förderplänen, unterrichtliche Klassenführung, Strukturierung des Unterrichts)

3. Management und Organisation (u.a. Zusammenarbeit von Regelschullehrkraft und Sonderpädagog*in, Klassenstufenteam, Jugendgerichtshilfe, Träger der Jugendhilfe)
4. Beratung und Kommunikation (u.a. multiprofessionelle bzw. interdisziplinäre Teams)

Die ersten beiden Aufgabenfelder *Diagnose* und *Intervention* sind konkret auf den Unterricht, die anderen beiden Aufgabenfelder *Management und Organisation* sowie *Beratung und Kommunikation* eher systemisch über diesen hinaus zu beziehen. Die Begriffe *Diagnose* und *Diagnostik* werden in den Ausführungen von König et al. (2019) synonym verwendet – dies spiegelt möglicherweise die Wortwahl der analysierten Publikationen wider – wobei nur Ersterer Eingang in ihre Modellierung gefunden hat. Bei ihren Ausführungen zur Erweiterung der Lehrkraftkompetenzen um inklusive Aspekte orientieren sich König et al. (2019) vorrangig an folgenden von Grosche (2015) vorgeschlagenen operationalisierten Definitionen von Inklusion: der *Förderzieldefinition* („Inklusion ist die Förderung von Schulleistungen und anerkennenden Beziehungen zwischen Klassenkameradinnen und Klassenkameraden sowie zu den Lehrkräften“) und der *Outcome-Definition* („Inklusion ist die Erreichung von Lernzielen bei allen Kindern und Jugendlichen gemäß ihrer jeweiligen Lernvoraussetzungen sowie die Erreichung von wertvollen und tragfähigen Beziehungen in der Klassengemeinschaft.“). Für ein präventives Verständnis von Inklusion, auf welches sich die vorliegende Arbeit vornehmlich fokussiert, soll insbesondere auch die *Präventionsdefinition* („Inklusion ist die gezielte Förderung zur Verhinderung eskalierender Lernverläufe bereits vor dem Auftreten von Lernrückständen“) herausgestellt werden (Grosche, 2015). Nach diesem Verständnis von Diagnose/Diagnostik geht es nicht vorrangig darum, Lern- und Entwicklungsrisiken festzustellen, wenn diese bereits auftreten und somit die Attestierung von Diagnosen. Eine präventiv ausgerichtete (Förder-)Diagnostik impliziert stattdessen die frühzeitige Bestimmung des Lern- und Entwicklungsstands aller oder einzelner Kinder einer Klasse, etwa anhand von Screenings oder anderer niederschwelliger und effizient durchzuführender diagnostischer Maßnahmen, um daraus präventive Fördermaßnahmen ab- und einzuleiten. Ferner geht es bei einer präventiven (Förder-)Diagnostik, im Folgenden *schulische Diagnostik* genannt, auch um die frühzeitige Überprüfung des Erfolgs der Fördermaßnahmen und gegebenenfalls deren Anpassung. Die schulische Diagnostik ist somit sehr eng mit einer präventiven schulischen Förderung verknüpft – mehr noch, beide bedingen einander. Da sie im Sinne eines präventiven Verständnisses von Inklusion zentral sind, werden die Begriffe *Diagnose* und *Intervention* im Modell von König et al. (2010) daher um die Spezifizierungen (*schulische*) *Diagnostik* und (*präventive*) *Förderung* ergänzt.

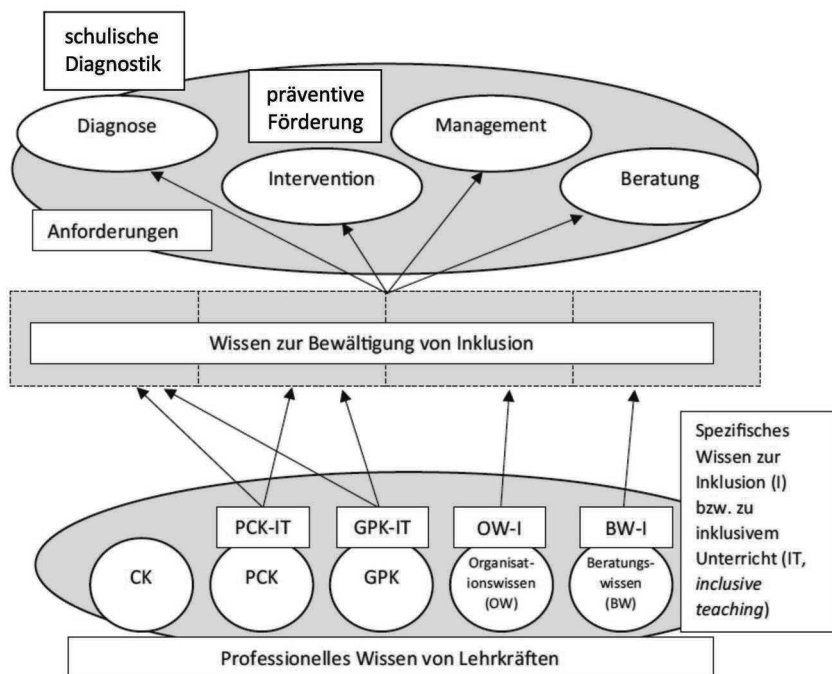


Abbildung 6: Professionelles Wissen von Lehrkräften und Anforderungen der Inklusion; Weiterentwicklung des von König et al. (2019) adaptierten Modells von Baumert & Kunter (2006)

Anmerkungen. CK – Content Knowledge (Fachwissen), PCK – Pedagogical Content Knowledge (Fachdidaktisches Wissen), GPK – General Pedagogical Knowledge (Pädagogisch-psychologisches Wissen), OW – Organisationswissen, BW – Beratungswissen, I – Inklusion (Inklusion), IT – Inclusive Teaching (inklusive Unterricht).

Für den Bereich des pädagogisch-psychologischen Wissens, das im Rahmen des entwickelten Professionalisierungskonzepts im Fokus steht, stellen die Autor*innen eine noch konkretere Erweiterung des Anforderungsbereichs für den inklusiven Unterricht vor, das sogenannte *pädagogische Wissen für inklusiven Unterricht* (*General Pedagogical Knowledge for Inclusive Teaching* [GPK-IT], S. 51). Mit dem Fokus auf den „Unterricht als Kernaufgabe“ (König et al. 2019, S. 52) werden die erweiterten Aufgaben der schulischen Diagnostik/Diagnose und präventiven Förderung/Intervention als zentral herausgestellt. Im Bereich der schulischen Diagnostik/Diagnose impliziert ein erweitertes pädagogisches Wissen, das Wissen über Dispositionen und Unterschiede (Aspekte der Statusdiagnostik), das Wissen über Lernprozesse (Aspekte der Prozessdiagnostik) und das methodische Wissen über Diagnose

(z.B. zur Bewertung von Verfahren). Neben dem Wissen um Lernprozesse der Schüler*innen, die sich vor allem auf die schulischen Lernbereiche Lesen, Schreiben und Rechnen beziehen, sollte das Modell außerdem durch das Wissen um spezifische *Entwicklungsprozesse*, z.B. für die Bereiche ESE, Sprache oder Kognition, erweitert werden.

Für den Bereich der präventiven Förderung/Intervention werden folgende zusätzliche Wissensfacetten als zentral erachtet: Klassenführung mit besonderem Schwerpunkt in inklusiven Settings (Bsp. Unterricht mit Schüler*innen mit dem FSP ESE), Strukturierung sowie Maßnahmen der Binnendifferenzierung und Individualisierung. Die Begriffe *Binnendifferenzierung*, *Individualisierung* und *Strukturierung* sind nicht erst seit der Debatte um Inklusion aktuell, sondern stellen bereits seit Langem grundlegende Merkmale „guten Unterrichts“ dar (Meyer, 2011). König et al. (2019) erläutern diese jedoch nicht weiter. Mit Blick auf spezifisch inklusive Anforderungsbereiche an Lehrkräfte müsste daher wohl eher von einer Erweiterung bzw. Spezifizierung von Maßnahmen der Binnendifferenzierung und Individualisierung ausgegangen werden. Die Wissensfacetten werden daher um diese Zusätze ergänzt.

Im Sinne eines engen Verständnisses von Inklusion¹² und dem dafür benötigten „sonderpädagogischen Spezialwissen“ ergänzen König et al. (2019, S. 58) die Wissensfacetten um Kenntnisse über Interventionskonzepte bei Schüler*innen unter Hochrisikobedingungen und mit spezifischen Behinderungen sowie Krankheiten, die Begleitung basaler Lernprozesse, Hilfsmittelversorgung und sonderpädagogische Diagnostik. Tabelle 7 stellt das (sonder-)pädagogisch-psychologische Basis- sowie Spezialwissen für den inklusiven Unterricht nach König et al. (2019) inklusive seiner erläuterten Erweiterungen dar.

Auf Basis der konzeptuellen Überlegungen zum *GPK-IT* und den darin enthaltenen Wissensbereichen Diagnose und Intervention entwickelten Gottfried et al. (2021) ein Theoriemodell für die Inklusion von Schüler*innen mit emotional-sozialen Beeinträchtigungen und Verhaltensproblemen, das *General Pedagogical Knowledge for Inclusive Teaching focussing on Social and Emotional Learning (GPK-IT-SEL)*. Das Modell ergänzt das GPK-IT um spezifischeres, sonderpädagogisches Wissen unter der Annahme, dass dieses auch für Regelschullehrkräfte zunehmend relevanter wird. Grundlegend wird angenommen, dass die Unterrichtsorganisation, das Lehrkraftfeedback sowie soziale Kompetenzen einen positiven Effekt auf das Klassenklima und den Umgang miteinander ausüben, sodass diese Faktoren durch unterschiedliche

12 Unter einem engen Verständnis von Inklusion verstehen König et al. ein Verständnis, „das sich in der Sonderpädagogik verortet und auf Menschen mit Behinderung bezieht; es zielt auf die Veränderung von Bildungsinstitutionen im Sinne einer Anpassung an die Lernvoraussetzungen und -bedürfnisse von Menschen mit Beeinträchtigungen [ab]“ (2019, S. 44).

Tabelle 7: Pädagogisch-psychologisches Basiswissen sowie sonderpädagogisches Spezialwissen für den inklusiven Unterricht; Erweiterung der herausgestellten Merkmale nach König et al. (2019)

Pädagogisch-psychologisches Basiswissen für inklusiven Unterricht / GPK-IT	
erweiterte Aufgaben der schulischen Diagnostik/Diagnose	erweiterte Aufgaben der präventiven Förderung/ Intervention
Wissen über Dispositionen und Unterschiede (Aspekte der Statusdiagnostik)	Klassenführung mit besonderem Schwerpunkt in inklusiven Settings (z.B. Unterricht mit Schüler*innen mit dem Förderschwerpunkt ESE)
Wissen über Lern- und Entwicklungsprozesse (Aspekte der Prozessdiagnostik)	Erweiterung bzw. Spezifizierung von Maßnahmen der Strukturierung
methodisches Wissen über Diagnose (z.B. zur Bewertung von Verfahren)	Erweiterung bzw. Spezifizierung von Maßnahmen der Binnendifferenzierung und Individualisierung
Sonderpädagogisch-psychologisches Spezialwissen für den inklusiven Unterricht	
Kenntnisse über Interventionskonzepte für Schüler*innen unter Hochrisikobedingungen und mit spezifischen Behinderungen sowie Krankheiten	
Begleitung basaler Lernprozesse	
Hilfsmittelversorgung	
sonderpädagogische Diagnostik	

Interventionen der Lehrkräfte gezielt gefördert werden können. Lehrkräfte sollten sich der „Bedingung von personenbezogenen Faktoren und externen Umweltfaktoren“ bewusst sein und „über Wissen in diesen Bereichen verfügen sowie Interventionen kennen, um beide Faktoren positiv zu beeinflussen“ (Gottfried et al. 2021, S. 263). Für den Umgang mit externalisierenden Verhaltensproblemen erachten die Autor*innen präventives Classroom Management (vgl. Kap. 2.1.3.1) sowie evidenzbasierte Handlungsmodelle (vgl. Kap. 2.1.3.3 und 2.1.3.4) als besonders relevant. Das *GPK-IT-SEL* kann zusätzlich um spezifisches Lehrkraftwissen zu Lern- und Entwicklungsstörungen (Leko & Roberts, 2014) sowie um Prinzipien und Maßnahmen zur gezielten Förderung akademischen und sozial-emotionalen Lernerfolgs (Osher et al., 2014) ergänzt werden. Ferner gilt es, die Zusammenhänge pädagogisch-psy-

chologischen Wissens mit übergeordneten affektiv-motivationalen Aspekten und Dispositionen, beispielsweise in Bezug auf Einstellungen zur Inklusion und Selbstwirksamkeitserwartung in inklusiven Kontexten, genauer herauszuarbeiten (Gottfried et al., 2021; Hecht et al., 2016). So weisen aktuelle Forschungsergebnisse zur inklusiven Beschulung vermehrt darauf hin, dass sehr unterschiedliche Einflussfaktoren den Erfolg inklusiven Unterrichts bedingen, wie etwa Überzeugungen, die Teilnahme an Fortbildungen oder die Erfahrung im Umgang mit Menschen mit Beeinträchtigungen (Van Mieghem et al., 2020).

Wenn man sich nach diesen Ausführungen nun wieder dem COACTIV-Modell (Baumert & Kunter, 2006) zuwendet, so ergibt sich, dass dieses vornehmlich die universelle Förderung von Schüler*innen in den Blick nimmt. Für die inklusive Beschulung sind aber – wie bereits herausgestellt wurde – insbesondere Methoden und Strukturen der Differenzierung und Individualisierung notwendig. Mehrebenenmodelle zur Realisierung organisierter schulischer Prävention stellen vielversprechende Ansätze dar, um inklusive Schule und damit die bestmögliche und individuelle Förderung aller Schüler*innen zu realisieren (Blumenthal et al. 2020, S. 37; vgl. Kap. 2.1.3.4). Im Sinne der Erweiterung des Professionsverständnisses von Lehrkräften wurde das COACTIV-Kompetenzmodell deshalb anhand der mehrstufigen präventiven Förderebenen des RTI-Modells (Gresham et al., 2005) für den Kompetenzbereich pädagogisch-psychologischen Wissens modelliert und dabei um die herausgearbeiteten inklusiven Wissensfacetten (Gottfried et al., 2021; König et al., 2019; Leko & Roberts, 2014; Osher et al., 2014) ergänzt (s. Abbildung 7). Eine kompetente und professionelle Regelschullehrkraft agiert hierbei mindestens auf den ersten zwei Ebenen, wobei sie zusätzlich und mit jeder Förderstufe zunehmend von weiteren qualifizierten Personen (z.B. Sonderpädagog*innen oder Schulpsycholog*innen) hinsichtlich spezifischer Maßnahmen beraten und unterstützt wird. Auf der Förderebene I findet die universelle Förderung aller Kinder durch hochwertigen Klassenunterricht statt, der in diesem Wissensbereich u.a. durch eine effektive Klassenführung, die Orchestrierung des Lernprozesses, die Anwendung allgemeiner Unterrichtsmethoden, die Berücksichtigung verschiedener Aspekte der Diagnostik, Leistungsbeurteilung und Heterogenität (Baumert & Kunter, 2006; Blumenthal et al. 2020, S. 37) sowie (sonder-)pädagogisches Basiswissen zum inklusiven Unterricht (Gottfried et al., 2021; König et al., 2019) gekennzeichnet ist. Auf der Förderebene II werden die Schüler*innen gefördert, die durch unterdurchschnittliche schulische Leistungen oder Entwicklungsprobleme (z.B. Probleme im emotionalen und sozialen Bereich) auffallen. Lehrkräfte benötigen auf dieser Förderebene (sonder-)pädagogisches Spezialwissen (Gottfried et al., 2021; König et al., 2019; Leko & Roberts, 2014; Osher et al., 2014), um anhand individuellerer pädagogischer Maßnahmen und engmaschiger Verhaltensverlaufsdiagnostik den Unterricht zu ergänzen. Auf der

Ebene III werden Kinder gefördert, deren Leistungen/Verhalten sich trotz erfolgter zusätzlicher pädagogischer Maßnahmen deutlich von der Altersnorm unterscheiden. Eine speziell qualifizierte Lehrkraft (z.B. Sonderpädagog*in) fördert die betreffenden Kinder in einer Einzelsituation oder in Kleinstgruppen von maximal drei Kindern durch hochindividualisierte sonderpädagogische und/oder therapeutische Einzelfallförderung und engmaschige Verhaltensverlaufsdiagnostik (mehrmals täglich). Für Schüler*innen der Ebene III sind beratende und differenzierende Maßnahmen für den Unterricht auf den Ebenen I und II notwendig (Blumenthal et al., 2020). Die von König et al. (2019) als zentral für inklusiven Unterricht herausgestellten Aufgabenfelder Diagnostik, Intervention, Management und Organisation sowie Beratung und Kommunikation finden grundsätzlich auf allen RTI-Förderebenen statt, wobei mit der Erhöhung des Individualisierungsgrades auch die Förderintensität und damit der Einsatz von Diagnostik und Interventionen zunimmt. Gleichzeitig findet eine zunehmend engere Abstimmung zwischen Lehrkraft und weiteren qualifizierten Personen (z.B. Sonderpädagog*innen, Schulpsycholog*innen, Therapeut*innen, etc.) statt und erfordert dadurch einen Mehraufwand für Management und Organisation sowie Beratung und Kooperation. Mit zunehmender Individualisierung der Förderung benötigt die Regelschullehrkraft vermehrt spezialisiertes (sonder-)pädagogisches Wissen. Förderung auf den Ebenen I und II findet unterrichtsimmanent (Blumenthal et al., 2020) und damit maßgeblich durch die Klassenlehrkraft statt. Die Übergänge zwischen allgemein-pädagogischem und sonderpädagogischem Wissen verlaufen hierbei fließend. Förderung auf Ebene III findet außerhalb des Unterrichts durch weitere qualifizierte Personen statt. Auf dieser Ebene ist spezialisiertes sonderpädagogisches oder gar therapeutisches Wissen nötig.

Die vorliegende Arbeit rekurriert in ihrem Verständnis professioneller Handlungskompetenz auf die Ergänzung und Weiterentwicklung des COACTIV-Kompetenzmodells (Baumert & Kunter, 2006) um situationsspezifische Skills (Blömeke et al., 2015) sowie (sonder-)pädagogisches Basis- und Spezialwissen zu inklusivem Unterricht (Gottfried et al., 2021; König et al., 2019; Leko & Roberts, 2014; Osher et al., 2014). Der RTI-Ansatz als mehrstufiges präventives Förderkonzept bildet den strukturellen Rahmen der neuen Modellierung inklusiver pädagogisch-psychologischer Kompetenz.

Kompetenztheoretische Professionalisierungsansätze postulieren, dass Kompetenzen sich nicht in festgeschriebenen Persönlichkeitsmerkmalen erschöpfen, sondern dass diese erlern- und veränderbar sind (Baumert & Kunter, 2006; Gräsel & Trempler, 2017; Kunter et al., 2011). Den sich daraus ergebenden Fragen, wie sich die professionelle Handlungskompetenz von Lehrkräften (weiter-)entwickelt und wie die modellierte inklusive pädagogisch-psychologische Kompetenz anhand ihrer Kompetenzmerkmale Wissen, Können und Überzeugungen durch berufsbegleitende Professionalisierungsmaßnahmen gefördert werden können, soll im Folgenden nachgegangen werden.

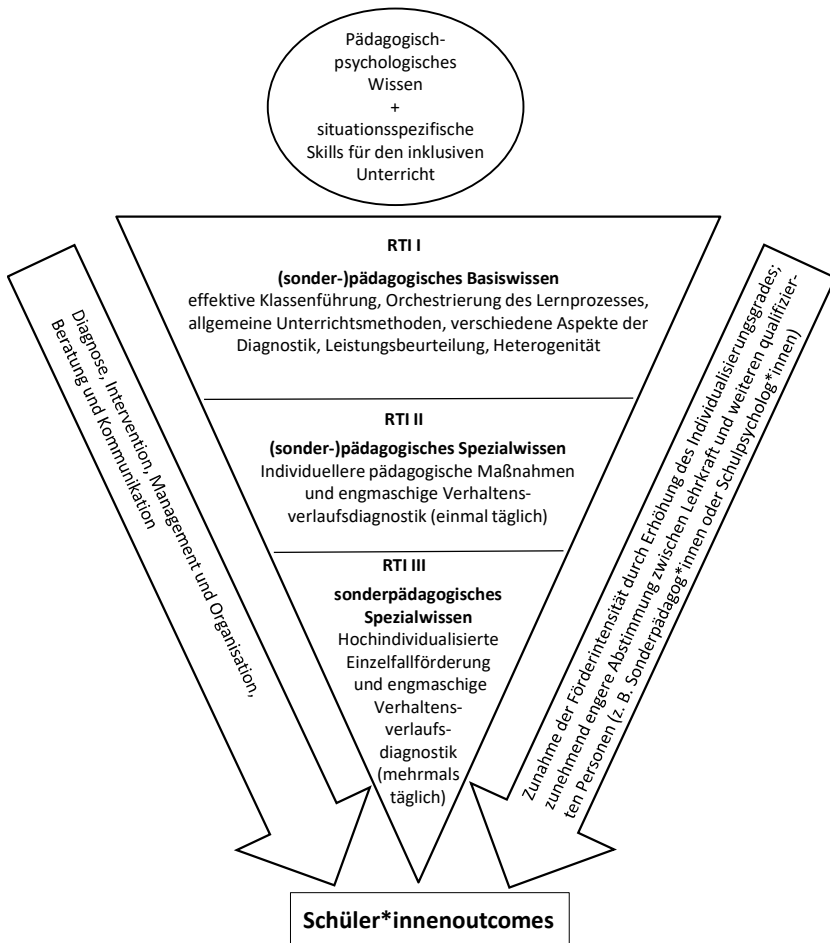


Abbildung 7: Modellierung der Lehrkraftkompetenzen im Bereich des pädagogisch-psychologischen Professionswissens (Baumert & Kunter, 2006), erweitert um situationsspezifische Skills (Blömeke et al., 2015) und (sonder-)pädagogisches Basis- und Spezialwissen für den inklusiven Unterricht (Gottfried et al., 2021; König et al., 2019; Leko & Roberts, 2014; Osher et al., 2014) anhand des mehrstufigen Förderkonzepts des Response-to-Intervention Modells (Gresham et al., 2005)

Anmerkung. RTI = Response-to-Intervention.

2.2.2.2 (Weiter-)Entwicklung professioneller Handlungskompetenz von Lehrkräften

Zur Analyse der (Weiter-)Entwicklung professioneller Handlungskompetenz über die berufliche Laufbahn hinweg erfolgt im Kontext kompetenztheoretischer Ansätze der Rückgriff auf die Expertiseforschung (König, 2010). Der *Expertiseansatz* nach Berliner (2006) fokussiert auf das Wissen als Grundlage des Handelns (Bromme, 1992) und fragt nach der erfolgreichen Lehrkraft (Expert*in) im Kontrast zu einer Person, die nicht über die vermutete Expertise verfügt (Noviz*in). Der Ansatz erklärt die Entwicklung professioneller Lehrkraftkompetenz in der Abfolge von fünf Entwicklungsstadien. Diese sind zunehmend geprägt von einer Vernetzung des Wissens, Veränderung der kategorialen Wahrnehmung von Situationen im Unterricht, Automatisierung von Prozessen sowie der damit einhergehenden Schnelligkeit und Flexibilität (König, 2010). Ausgehend von *Noviz*innen* in der ersten Phase der Lehrer*innenbildung, die weder über professionelles Wissen noch über Erfahrung als Lehrperson verfügen, entwickeln sich die zukünftigen Lehrkräfte im Referendariat zu *fortgeschrittenen Anfänger*innen*. Mit der Aneignung einer selbstbestimmten Eigenaktivität bei der Durchführung von Unterricht erreichen sie im dritten bis fünften Berufsjahr dann in der Regel das Stadium der *kompetent Handelnden*. Ab dem fünften Berufsjahr würde sich das vierte Stadium der *Profilierten* anschließen. Dieses, durch einen hohen Erfahrungsreichtum und die Automatisierung beruflicher Tätigkeiten geprägte Kompetenzniveau, erreichen aber nur noch wenige Lehrkräfte. Die bereits sehr hoch entwickelten Kompetenzen können ab dem fünften Berufsjahr auch im *Expert*innen*status münden, der sich durch eine noch ausgeprägtere Automatisierung der beruflichen Tätigkeiten und herausragende Leistungen in der Lehrtätigkeit auszeichnet (Berliner 2004, S. 207; König 2010, S. 55). Studien auf der Basis des Expert*innenparadigmas konnten den postulierten Kompetenzzuwachs zwischen Noviz*innen und berufserfahrenen Lehrkräften nachweisen (u.a. Berliner, 2004; Herzmann & König 2016, S. 82f.; Kaiser et al., 2017; König, 2010). Es liegen bislang allerdings kaum empirische Belege für die postulierte Fünfstufigkeit vor (Baumert & Kunter, 2006; König 2010, S. 55f.).

Anders als der Expertiseansatz vermuten lässt, ist nicht davon auszugehen, dass die professionelle Handlungskompetenz von Lehrkräften allein aufgrund zunehmender Berufserfahrung ansteigt (Brunner et al., 2006). Sie hängt viel eher von konkreten Lerngelegenheiten, der Qualität des Lernangebots und der Reflexion des Lernprozesses ab (Lipowsky, 2014; Richter, 2016). Bei der Nutzung von Lerngelegenheiten wie auch der Nutzbarmachung des Gelernten im Kontext der Weiterentwicklung der eigenen professionellen Handlungskompetenz spielen wiederum kognitive Fähigkeiten und die Persönlichkeit der einzelnen Lehrkraft sowie relevante Vorerfahrungen, Berufsmotivation, berufliches Selbstkonzept und individuelle berufliche

Ziele eine wichtige Rolle (Lipowsky, 2014; Kunter et al., 2011; Schellenbach-Zell, 2009). Für die (Weiter-)Entwicklung professioneller Kompetenz von Lehrkräften lassen sich grundsätzlich zwei Argumentationslinien herausstellen: die *Eignungs-* und die *Qualifikationshypothese* (Kunter et al. 2011, S. 56ff.). Die *Eignungshypothese* besagt, dass Lehrkräfte bereits vor ihrer Ausbildung gewisse Dispositionen bzw. ein Talent für den späteren Beruf mitbringen. Diese Merkmale beziehen sich u.a. auf kognitive, motivational-selbstregulierende Fähigkeiten sowie Verhaltens- und Erlebenstendenzen. Die *Qualifikationshypothese* stellt wiederum die Art und Weise der Ausbildung von Lehrkräften als richtungsweisend für deren professionelle Kompetenz heraus und zwar unabhängig davon, welche Persönlichkeitsmerkmale eine Person mitbringt. Im Gegensatz zur Eignungshypothese, die sich auf berufunspezifische Eigenschaften bezieht, stellt die Qualifikationshypothese professionsspezifische Kenntnisse und Strategien in den Vordergrund, die durch die Vermittlung grundlegender Konzepte und Prinzipien erlernbar sind (Kunter et al., 2011). Das COACTIV-Kompetenzmodell, das in der vorliegenden Arbeit das theoretische Basisverständnis für professionelle Kompetenz bildet, verbindet beide konkurrierenden Ansätze. Neben berufunspezifischen Eingangsvoraussetzungen, wie motivational-selbstregulativen Merkmalen, sind für die (Weiter-)Entwicklung professioneller Kompetenz von Lehrkräften demnach vor allem Lerngelegenheiten zentral, bei denen kognitive Merkmale wie Wissen, Können und Überzeugungen erworben oder entwickelt werden. Kunter et al. (2011, S. 60) gliedern Lerngelegenheiten in formale, non-formale und informelle Situationen; damit umfassen diese sowohl das Konzept des Professional Learning als auch des Professional Development (s. Kap. 2.2.1). Studienergebnisse aus Übersichtsarbeiten (u.a. Darling-Hammond et al., 2017; Lipowsky & Rzejak, 2012; Timperley et al., 2007) unterstreichen insbesondere die Relevanz von systematisch geplanten Fortbildungen im Rahmen der (Weiter-)Entwicklung professioneller Handlungskompetenz.

Zusammenfassend ist festzuhalten: Nach dem Expertiseansatz von Berliner (2004) können im Schuldienst befindliche Lehrkräfte, wie sie auch im Rahmen dieser Arbeit im Fokus stehen, als kompetent Handelnde bezeichnet werden, wobei deren Expertise-Status maßgeblich von den eigenen individuellen Persönlichkeitsmerkmalen und dem Lernpotential von Lerngelegenheiten abhängt. Wie herausgestellt wurde (Kap. 2.2.2.1), verändert sich jedoch das Expertiseprofil von Lehrkräften im Zuge der inklusiven Ausrichtung des Schulsystems. Um den Schüler*innen und dem eigenen Anspruch an die Lehrtätigkeit gerecht werden zu können, gilt es somit, die professionelle Handlungskompetenz/Expertise von im Schuldienst befindlichen Lehrkräften anhand von Lerngelegenheiten weiterzuentwickeln. Als eine effektive Möglichkeit, Lehrkräfte im Beruf (weiter) zu professionalisieren, haben sich *Fortbildungsmaßnahmen* erwiesen (u.a. Darling-Hammond et al., 2017; Kun-

ter, 2011; Lipowsky & Rzejak, 2012; Timperley et al., 2007). Im folgenden Kapitel soll untersucht werden, wie der Erwerb und die Erweiterung von Lehrkraftkompetenzen bei formalen Lerngelegenheiten in Form von Fortbildungen gelingt.

2.2.3 *Lehrkraftprofessionalisierung durch Fortbildung*

Fortbildungen in der Lehrer*innenbildung zielen auf die Unterstützung der Lehrkräfte in ihrer aktuellen Berufspraxis ab und sollen die Sicherstellung und Weiterentwicklung professionsbezogener Lernprozesse erreichen (Fussangel et al., 2016). Sie können außerdem als formale Lerngelegenheiten und im Sinne des Professional Development als koordinierte, umfassende und intensive Aktivitäten bezeichnet werden, die darauf abzielen, das Wissen, die Überzeugungen, die Fähigkeiten und die Praktiken der Lehrkräfte zu verändern bzw. zu verbessern, um die Ergebnisse der Schüler*innen zu steigern (Kunter et al. 2011, S. 58f.; Leko & Roberts, 2014; Lipowsky 2014, S. 512). Bei dem Transfer von Fortbildungsinhalten in die Unterrichtspraxis handelt es sich allerdings um einen komplexen und anspruchsvollen Prozess (Lipowsky, 2010).

Betrachtet man zunächst Lerngelegenheiten in der dritten Phase der Lehrer*innenbildung im Allgemeinen, so können zur Erklärung der Kompetenzentwicklung auf Seiten der Lehrkräfte sehr unterschiedlich ausgerichtete Modelle herangezogen werden (u.a. Clarke & Hollingsworth, 2002; Desimone, 2009; Evans, 2014; Guskey, 2002; Opfer & Pedder, 2011). In einer vergleichenden Analyse fünf verschiedener Modelle subsumieren Boylan et al. (2018), dass diese auf sehr unterschiedlichen Zielen und Schwerpunkten beruhen, sogar bis zu einem gewissen Grad „aneinander vorbei sprechen“ (S. 37) und dass keines der Modelle die Komplexität der beruflichen Professionalisierungsprozesse umfassend abbilden kann. Dennoch ist es möglich, gewisse Parallelen festzuhalten: In allen Modellen wird angenommen, dass Wechselwirkungen zwischen dem Wissen und den affektiv-motivationalen Merkmalen der Lehrkräfte, der Realisierung praktischen Handelns, den formalen, non-formalen und informellen Lerngelegenheiten als auch zwischen den Lehrkräften und dem jeweiligen schulischen System bestehen und diese somit die konkreten beruflichen Veränderungsprozesse beeinflussen (Durksen et al., 2017; Kunter et al., 2011; Opfer & Pedder, 2011). Im deutschsprachigen Raum hat vor allem Lipowsky zum „Lernen im Beruf“ von Lehrkräften (2004, 2010, 2014) ein entscheidendes Grundmodell etabliert. Im *erweiterten Angebots- und Nutzungsmodell* zur Erklärung der Wirksamkeit von Fortbildungs- und Professionalisierungsmaßnahmen werden die herausgestellten komplexen Zusammenhänge zwischen individuellen Determinanten und kontextuellen Bedingungen dargestellt und um strukturelle und didakti-

sche Merkmale der Fortbildungen sowie deren Interaktionen erweitert. Neben den Merkmalen auf Seiten der lernenden Lehrpersonen, wie individuelle kognitive, motivationale, volitionale und persönlichkeitsbezogene Voraussetzungen und ihre (berufs-)biographische Situation, wirken die Komponenten der Fortbildung sowie eine Vielzahl schulischer Kontextbedingungen auf diesen Prozess ein, etwa Merkmale der Schule, der Klasse und der Lernenden (Lipowsky, 2010). Im Zentrum des Modells steht die Wahrnehmung und Nutzung des Angebots innerhalb einer Fortbildung, die anhand der Intensität aktiven und engagierten Lernens und der Zufriedenheit sowie der wahrgenommenen Relevanz der Fortbildung sichtbar wird. Die Wahrnehmung und Nutzung des Fortbildungsangebots ist wiederum abhängig von den Voraussetzungen der Lehrkräfte – etwa den Zielen, der Persönlichkeit, den Überzeugungen und/oder dem Wissen – sowie der strukturell-inhaltlichen und didaktischen Konzeption der Fortbildungsmaßnahme selbst. Kontextbedingungen, wie schulische Anforderungen, Schulprogramm, Klima und Konsens im Kollegium oder der finanzielle Rahmen können Einfluss auf die Durchführung der Fortbildungsmaßnahme haben, entfalten ihre volle Wirkung aber vor allem erst nach deren Durchführung. So kann die Unterstützung der Schulleitung, die Gelegenheit zur Anwendung des Gelernten oder das Kooperationsklima im Kollegium den Transferprozess mit beeinflussen und zum Fortbildungserfolg beitragen oder diesen eher behindern. Letzterer wird außerdem durch die individuellen Dispositionen der Lehrkraft wie der Erfolgserwartung, dem beigemessenen subjektiven Wert der Veränderungen, der Förderung und dem Erhalt der Teilnahmemotivation sowie der Transfermotivation mit beeinflusst. Das komplexe Zusammenwirken der genannten Faktoren ist in Abbildung 8 dargestellt.

Trotz des „Variablenbündels“ (Lipowsky 2010, S. 50) an individuellen Determinanten, kontextuellen Bedingungen sowie strukturellen und didaktischen Merkmalen der Fortbildungen, die Einfluss auf den Fortbildungserfolg nehmen können, wird in den letzten Jahren eine wachsende Tendenz der empirischen Überprüfung von Lehrkraftfortbildungen sichtbar. Die Ergebnisse weisen auf eine deutlich erkennbare Wirksamkeit von Fortbildungs- und Professionalisierungsmaßnahmen hin, sofern diese unter günstigen Voraussetzungen stattfinden (u.a. Lipowsky, 2010; Rzejak et al., 2020). Bevor genauer auf die Wirksamkeitsforschung zu Lehrkraftfortbildungen im Kontext von Inklusion unter besonderer Berücksichtigung des FSP ESE eingegangen wird, soll zunächst ein allgemeiner Überblick über die Wirkungsweise und Wirksamkeit von Lehrkraftfortbildungen gegeben werden, der auch für den inklusiven Unterricht handlungsleitend ist. Hieraus lassen sich bereits wertvolle allgemeingültige Anhaltspunkte für die Konzeption und Evaluation von Professionalisierungsmaßnahmen zur (Weiter-)Entwicklung der inklusiven pädagogisch-psychologischen Kompetenz herleiten.

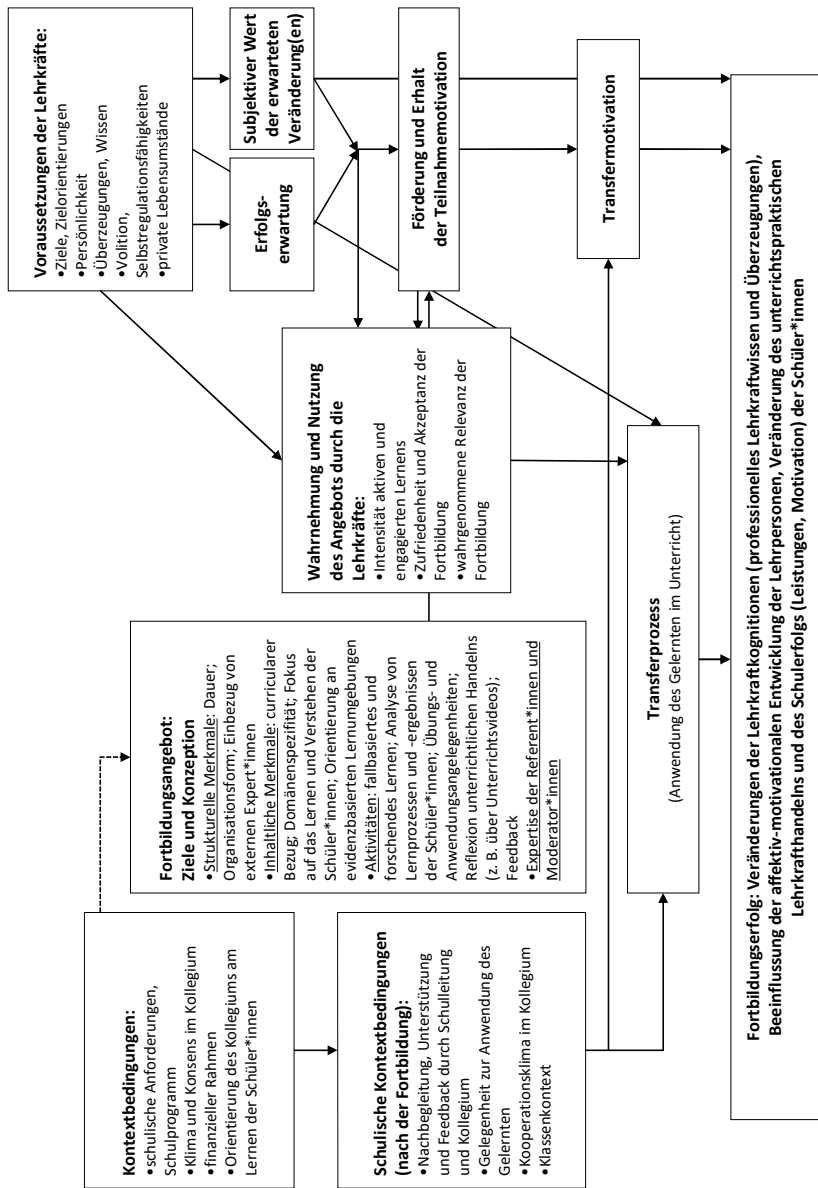


Abbildung 8: Erweitertes Angebots- und Nutzungsmodell zur Erklärung der Wirksamkeit von Lehrkraftfortbildungen; erstellt nach Lipowsky (2010, S. 51)

2.2.3.1 Wirkungsweise und Wirksamkeit von Lehrkraftfortbildungen

Die Wirkungsweise von Lehrkraftfortbildungen lässt sich anhand der von Lipowsky (2010, 2014), Guskey (2014) und King (2014) eingeführten Klassifikation modellieren, welche die Reichweite von Lehrkraftfortbildungen anhand von verschiedenen nicht-hierarchischen Wirkungsebenen erklärt (s. Tabelle 8). Die erste Ebene stellt die Reaktionen und Einschätzungen der Lehrkräfte in Bezug auf die Fortbildungsmaßnahme dar, bei der insbesondere deren Zufriedenheit und Nützlichkeitseinschätzung entscheidend sind. Auf Ebene zwei findet die *Erweiterung des deklarativ-prozeduralen Lehrkraftwissens* und der *affektiv-emotionalen Lehrkraftkognitionen* statt. Auf der dritten Ebene erfolgt die *Veränderung des unterrichtspraktischen und pädagogischen Lehrkrafthandelns* und auf der vierten Ebene wird die Wirkung auf Seiten der Schüler*innen sichtbar, etwa durch ein *verändertes akademisches* (Schulleistungen) *und/oder emotional-soziales Verhalten*, im Folgenden *Schüler*innenoutcomes* genannt. Eine fünfte Ebene, auf der *schulstrukturelle Unterstützung und Veränderung* stattfindet, soll zwar erwähnt, aber aufgrund der Studienkonzeption in der vorliegenden Arbeit nicht eingehender berücksichtigt werden.

Die Wirksamkeit von Lehrkraftfortbildungen wurde in vielfältigen und unterschiedlich ausgerichteten Studien untersucht. Im Folgenden sollen zentrale Ergebnisse anhand der ersten vier Wirkungsebenen vorgestellt werden.

Hinsichtlich der ersten Ebene, der Lehrkraftreaktionen, stellte sich heraus, dass Zufriedenheit und Akzeptanz teilnehmender Lehrpersonen besonders hoch sind, wenn die Professionalisierungsmaßnahmen möglichst alltagsnah erfolgen und somit einen Bezug zu Unterricht und Curriculum haben, wenn Partizipationsmöglichkeiten bestehen, eine angenehme Atmosphäre herrscht, der Austausch mit anderen Teilnehmenden möglich ist und die Veranstaltung professionell und kompetent geleitet wird (Guskey, 2002; Jäger & Bodensohn, 2007; Kwakman, 2003; Nir & Bogler, 2008; Schellenbach-Zell & Gräsel, 2007; Smith & Gillespie, 2007). Die Relevanz eines konkreten Bezugs der Fortbildungsinhalte zum schulischen Alltag konnte auch bei Fortbildungen im Kontext von Inklusion festgestellt werden. So zeigen Kurniawati et al. (2014) anhand ihres systematischen Reviews von 13 im Peer-Review veröffentlichten empirischen Studien auf, dass sich die berufliche Fortbildung von Lehrkräften im Kontext von Inklusion als effektiver erweist, wenn sie sich auf spezifische Bedürfnisse oder sonderpädagogische Förderbedarfe von Schüler*innen konzentriert und nicht auf inklusive Bildung im Allgemeinen. Die Ergebnisse für die Lehrkraftreaktionen lassen sich gut unter der Selbstbestimmungstheorie von Deci und Ryan (1985) subsumieren, die die Relevanz eines Lerngegenstands als eine wichtige Determinante für den Aufbau von Motivation ansehen. Ein hoher Grad an Mitbestimmung und Selbstverantwortung kann andererseits aber auch überfordern, wenn beispielsweise

Tabelle 8: Wirkungsebenen von Lehrkraftfortbildungen nach Lipowsky (2010, 2014), Guskey (2014) und King (2014)

Wirkungsebene		Konkretisierung
1	Reaktionen und Einschätzungen der Lehrkräfte	Zufriedenheit und Nützlichkeits-einschätzung
2	Erweiterung des Lehrkraftwissens und der -kognitionen	Erweiterung des fachlichen, fachdi-daktischen, pädagogisch-psychologi-schen und diagnostischen Wissens (deklarativ-prozedural); Veränderung von Überzeugungen und subjektiven Theorien (affektiv-emo-tional)
3	Veränderung des unterrichtsprakti-schen und pädagogischen Lehrkraft-handelns	Veränderte Unterrichtsmethoden, Instruktionsformen, Kommunikation; Qualitätssteigerung des Unterrichts
4	Beeinflussung des Schulerfolgs von Schüler*innen/Schüler*innenoutcomes	Verändertes akademisches und/oder emotional-soziales Verhalten
5	schulstrukturelle Unterstützung und Veränderung	Angebotswahrnehmung und Unter-stützung im System; Transferklima; erlebte Unterstützung durch Schullei-tung und Kollegium

nicht genügend Rückmeldungen und Impulse von außen erfolgen und/oder die Lehrkräfte nicht die nötigen selbstregulativen Kompetenzen mitbringen (Richardson, 1990). Bisher konnten kaum einschlägige Hinweise dafür gene-riert werden, dass ein hohes Maß an Akzeptanz und Zufriedenheit der Lehr-personen sich auch positiv auf die anderen drei Ebenen Lehrkraftwissen und -kognitionen, Lehrkrafthandeln sowie Schüler*innenoutcomes auswirken (Copur-Gencturk & Thacker, 2020; Goldschmidt & Phelps, 2007; Hochhol-dinger & Schaper, 2007; Salas & Cannon-Bowers, 2001). Angenommen werden kann jedoch, dass ein Minimum an Akzeptanz eine notwendige Voraussetzung für einen erfolgreichen Lern- und Transferprozess darstellt (Lipowsky, 2010).

In Bezug auf die Wirkungsebene zwei, die Erweiterung des Lehrkraft-wissens und der Lehrkraftkognitionen, gibt es eine Reihe ermutigender em-pirischer Nachweise zur Effektivität von Professionalisierungsmaßnahmen und deren Transfer auf andere Ebenen. So können positive Ergebnisse bezüg-lich der Erweiterung des Lehrkraftwissens selbst (u.a. Heran-Dörr, 2006; Lipowsky et al., 2006) als auch dessen Auswirkung auf das Lehrkrafthandeln

(Gärtner, 2007) und die Schüler*innenoutcomes (Möller et al., 2006; Timperley et al., 2007) festgestellt werden. Besonders enge Zusammenhänge können zwischen der Erweiterung des Lehrkraftwissens und einem veränderten Lehrkrafthandeln konstatiert werden (Garet et al., 2001; Stipek et al., 2001; Vehmeyer et al., 2007). Grundsätzlich ist anzunehmen, dass verändertes, erfolgreiches Unterrichtshandeln wiederum auch die affektiv-motivationalen Merkmale, Einstellungen und Überzeugungen von Lehrkräften nachhaltig verändern kann (Lipowsky, 2010). In der Inklusionsforschung werden die Einstellung und das Selbstwirksamkeitsempfinden von Lehrkräften für das konkrete unterrichtliche Handeln, insbesondere im FSP ESE, als besonders relevant erachtet (Van Mieghem et al. 2020, S. 6). Studien zeigen, dass die Erweiterung deklarativen Wissens über sonderpädagogische Förderschwerpunkte und die Erfahrung mit inklusiver Pädagogik die Einstellungen von Lehrkräften positiv beeinflussen können (de Boer et al., 2011; Qi & Ha, 2012).

Studien zur Veränderung des unterrichtspraktischen und pädagogischen Lehrkrafthandelns sind methodisch sehr aufwendig, da hierbei das konkrete Unterrichtshandeln der Lehrkräfte nach einer Fortbildungsmaßnahme mehrmals unter vergleichbaren Kontextbedingungen untersucht werden muss. Die Datenerhebung kann etwa durch Videoaufnahmen (Wackermann, 2008), strukturierte Tagebucheinträge, durch die Angabe der Häufigkeit eingesetzter Unterrichtsmethoden und/oder Instruktionsformen (Correnti, 2007; Desimone et al., 2002) oder Schüler*innenbefragungen (Kunter & Baumert, 2006) erfolgen. Oft bleibt die Ebene unterrichtspraktischen Handelns bei Wirksamkeitsstudien unberücksichtigt. Stattdessen wird öfter von den Veränderungen auf Schüler*innenebene auf potentiell stattgefundene Veränderungen im Unterricht rückgeschlossen (Lipowsky, 2010). Es ist aber davon auszugehen, dass Fortbildungsmaßnahmen, die die spezifischen Anliegen der Lehrkräfte und ihren Unterrichtskontext berücksichtigen, am ehesten dazu führen, auch Veränderungen in der schulischen Praxis der Lehrkräfte anzustoßen (Kurniawati et al., 2014; Roberts & Simpson 2016; Qi & Ha, 2014).

Studien, die sich mit den Effekten von Professionalisierungsmaßnahmen auf Schüler*innenleistungen und -verhalten beschäftigen, bewegen sich größtenteils in den kognitiven Lernbereichen Mathematik (Saxe et al., 2001; Shayer & Adhami, 2007), Naturwissenschaften (Adey, 2004; Fishman et al., 2003; Kleickmann & Möller, 2007;) und Sprachen (Guthrie et al., 2004; McDowall et al., 2007; Parr et al., 2007; Taylor et al., 2005) und haben teilweise eine hohe Wirksamkeit nachweisen können. Uneinheitliche Ergebnisse zeigen sich bezüglich der affektiv-motivationalen Entwicklung der Lernenden. Viele Studien können entweder keine überzeugenden und nachhaltigen Effekte feststellen (Fischer, 2006; Otto, 2007) oder betonen, dass es offenbar erheblicher und wahrscheinlich auch langfristiger Anstrengungen bedarf, mit Lehrkraftfortbildungen die affektiv-motivationale Entwicklung der Schü-

ler*innen nachhaltig zu beeinflussen (Rheinberg & Krug, 1999). Effekte auf das lernbezogene Verhalten zeigten sich bisher erst in wenigen Studien (Kurniawati et al., 2014).

Die vorliegenden Befunde verdeutlichen die Wirkungsweise und das Potential von Fortbildungen sowohl für Veränderungen bei Lehrkräften als auch Schüler*innen, zeigen aber auch Forschungsdesiderata auf. So gilt es insbesondere, die Schüler*innenoutcomes stärker in den Fokus von Effekstudien zu nehmen (Kurniawati et al., 2014). Da die Wirksamkeit von Lehrkraftfortbildungen und damit die (Weiter-)Entwicklung professioneller Handlungskompetenz maßgeblich von der Qualität der Professionalisierungsmaßnahme abhängen, sollen im Folgenden anhand der internationalen Befundlage zur Wirksamkeitsforschung im Bereich der Lehrkraftprofessionalisierung (u.a. Darling-Hammond et al., 2017; Leidig 2019, S. 44ff.; Rzejak et al. 2020, S. 21ff.; Rzejak & Lipowsky 2019, S. 20ff) konkrete strukturelle und didaktisch-methodische Merkmale wirksamer Fortbildungen herausgestellt werden, die auch für den inklusiven Kontext handlungsleitend sind (Leko & Roberts, 2014).

Verschränkung von Input, Erprobung, Feedback und Reflexion

Wirksame Fortbildungsmaßnahmen erschöpfen sich nicht in der Vermittlung von Wissen, sondern kombinieren die Abfolge von Input-, Erprobungs-, Feedback- und Reflexionsphasen (u.a. Darling-Hammond et al., 2017; Leidig, 2019; Rzejak et al., 2020; Rzejak & Lipowsky, 2019). Um neu erworbenes deklaratives Wissen in prozedurales Wissen zu überführen und die professionelle Handlungskompetenz weiterzuentwickeln, benötigen die Lehrkräfte die Möglichkeit, das erworbene Wissen selbstständig anzuwenden bzw. die vermittelten Handlungsstrategien, ausgehend von konkreten Beispielen zu erproben, um im Anschluss die Wirkung der gezeigten Performanz, mögliche Bedingungsfaktoren sowie die sich daraus ergebenden Konsequenzen für ihr professionelles Handeln zu reflektieren. Nach einer ersten Inputphase erfolgt etwa die Anwendung und Erprobung der neu erworbenen Fähig- und Fertigkeiten im schulischen Alltag. Dieser Prozess sollte wiederum durch unterschiedliche Feedbackformen (z.B. videobasiertes oder kollegiales Feedback, Rückmeldung und Analyse von Lernprozessdaten der Schüler*innen) begleitet werden, um dadurch Reflexionsprozesse bei den Lehrkräften auslösen zu können. Eine solche Verschränkung von Input-, Erprobungs-, Feedback- und Reflexionsphasen benötigt Zeit, insbesondere dann, wenn sie auf längerfristig angelegte Prozesse mit wiederholten Erprobungsphasen im Unterricht abzielt (Leidig, 2019; Rzejak et al., 2020; Rzejak & Lipowsky, 2019).

Längere Angebotsdauer

Auch wenn die Ergebnisse stark differieren, so zeigen die Ergebnisse von mehreren Metaanalysen, dass Kurzzeitfortbildungen nicht zu substanziellen Veränderungen des unterrichtlichen Handelns (Darling-Hammond et al., 2017) bzw. zu positiven Wirkungen auf Schüler*innenebene führen (Timperley et al., 2007; Yoon et al., 2007). Es kann hingegen konstatiert werden, dass wirksame Professionalisierungsangebote mindestens 14 Stunden betragen (Yoon et al., 2007) und sich in der Länge über mindestens ein halbes Jahr erstrecken (Timperley et al., 2007). Die genannten Werte sind aber eher als Schwellenwerte zu interpretieren und die Teilnahme an zeitintensiveren Maßnahmen bietet keine Garantie für größere Fortbildungseffekte (Rzejak & Lipowsky, 2019). Außerdem hängt die Wirkung der Fortbildungsmaßnahme maßgeblich mit dem anvisierten Kompetenzerwerb zusammen. Bei der Erweiterung von deklarativem Wissen etwa können bereits kürzere Fortbildungsformate positive Wirkungen zeigen, wie Lipowsky und Rzejak (2015) themenübergreifend und Kurniawati et al. (2014) für Fortbildungen von Grundschullehrkräften im Kontext Inklusion in ihren Überblicksarbeiten herausstellen. Geht es stattdessen um die Veränderung von Einstellungen anhand tiefergehender Reflexionen, so sind längere Formate vonnöten, die sich mitunter über mehrere Schuljahre erstrecken können (Leidig, 2019; Rzejak & Lipowsky, 2019).

Kohärenz

Unter Kohärenz kann die Abstimmung und Verzahnung zwischen dem Angebot der Fortbildung und den Bedarfen der Teilnehmer*innen, den curricularen und rechtlichen Vorgaben, den Vorerfahrungen und vorhandenen Kenntnissen und Überzeugungen der Teilnehmer*innen verstanden werden. Fortbildungen sind erfolgreicher, wenn sie an der Unterrichtspraxis und den schulischen Problemlagen der Lehrkräfte ansetzen (Leidig 2019, S. 44).

Orientierung an Befunden der Lehr-Lernforschung

Effektive Professionalisierungsmaßnahmen orientieren sich an den Befunden und Erkenntnissen aus der Lehr-Lernforschung. Nachweislich können insbesondere „Tiefenstrukturmerkmale von Unterricht“ (Rzejak et al. 2020, S. 22) – wie inhaltliche Klarheit und Strukturierung des Unterrichts, kognitive Aktivierung und metakognitive Förderung der Lernenden, konstruktives Feedback an die Schüler*innen und formatives Assessment, eine effektive Klassenführung, Formen intelligenten Übens sowie eine konstruktive Lernunterstützung (Hattie, 2018) – Lernprozesse nachweislich positiv beeinflussen.

Damit korrespondierend haben sich Fortbildungen als wirksam erwiesen, die ein oder gar mehrere Merkmale der Tiefenstruktur zum Gegenstand hatten und in denen es gelang, die unterrichtspraktische Relevanz dieser Merkmale zu verdeutlichen (effektive Klassenführung: Piwowar et al., 2013; Feedback/formatives Assessment: Andersson & Palm, 2017; Kiemer et al., 2015).

Fokussierung auf erlernbare Kernpraktiken

Die Orientierung an Merkmalen der Tiefenstruktur von Unterricht geht einher mit der Fokussierung auf erlernbare Kernpraktiken, die das Lernen von Schüler*innen nachweislich positiv beeinflussen. Die Herausforderung im Rahmen von Fortbildungsmaßnahmen besteht darin, einerseits die Komplexität von Lehr-Lern-Prozessen zu berücksichtigen und andererseits letztere zu dekonstruieren, sodass spezifische unterrichtliche Verhaltensweisen oder Strategien für Lehrkräfte erkenn- und erlernbar werden (Rzejak et al., 2020; Rzejak & Lipowsky, 2019).

*Analyse von Lernprozessen der Schüler*innen*

Viele Fortbildungen, die positive Wirkungen auf der Ebene des Unterrichts und des Lernens der Schüler*innen erzielen, zeichnen sich durch eine Analyse von Lernprozessen der Schüler*innen aus (u.a. Darling-Hammond et al., 2017; Timperley et al., 2007). In diesen Fortbildungen werden die Lern- und Verstehensprozesse von Schüler*innen, etwa anhand von Schüler*innendokumenten oder videografierten Unterrichtssequenzen, systematisch in den Fokus genommen. Hierbei erweitern Lehrkräfte vor allem ihre diagnostischen und fachdidaktischen Kompetenzen (Rzejak et al., 2020; Rzejak & Lipowsky, 2019).

Verdeutlichung der Wirkungen des Lehrkrafthandelns

Neben der Analyse von Lernprozessen der Schüler*innen kann sich die gezielte Untersuchung des eigenen unterrichtlichen Handelns auf das Lernen von Schüler*innen erfolgreich auf die professionelle Kompetenz von Lehrkräften und damit auf den Erfolg der Fortbildungsmaßnahme auswirken. Anhand der erlebten Erfahrung und Erkenntnis, dass teilweise schon kleine Veränderungen im unterrichtlichen Handeln sichtbare Auswirkungen auf das Lernen von Schüler*innen haben, kann die Motivation und die Selbstwirksamkeitserwartung der Lehrkräfte gefördert werden (Rzejak et al., 2020; Rzejak & Lipowsky, 2019).

Förderung einer unterrichtsbezogenen Kooperation

In vielen wirkungsvollen Fortbildungen werden die Lehrkräfte zu einer konstruktiven unterrichtsbezogenen Zusammenarbeit angeregt (u.a. Darling-Hammond et al., 2017). Diese zeichnet sich u.a. dadurch aus, dass die Lehrkräfte gemeinsame Werte und Überzeugungen zum Lehren und Lernen besitzen, sich in ihrer Zusammenarbeit auf die Förderung des Lernens von Schüler*innen fokussieren, kollegiale Unterrichtshospitationen und reziproke Feedbackformate zum Unterricht realisieren und sich intensiv über die Qualität und die Wirkung von Unterricht austauschen (Rzejak et al., 2020; Rzejak & Lipowsky, 2019).

Zusammenfassend gilt: Die berufliche Weiterbildung von Lehrkräften ist zentral, um die professionelle Handlungskompetenz im Sinne von deklarativem und prozeduralem Wissen (Baumert & Kunter, 2006) sowie situationspezifische Skills (Blömeke et al., 2015) weiterzuentwickeln. Dazu gehört, Lehrkraftwissen zu generieren, Einstellungen zu ändern und damit die Fähigkeit und Bereitschaft zu erhöhen, Schüler*innen mit spezifischen Beeinträchtigungen in der Regelschule zu unterrichten (Kurniawati et al., 2014). Fortbildungsmaßnahmen sollten evidenzbasierte pädagogische Strategien vermitteln und sich auf die spezifischen Bedürfnisse der Schüler*innen, die konkreten Anliegen der Lehrkräfte und ihren Unterrichtskontext konzentrieren (Van Mieghem et al., 2020). Den wissenschaftlichen Erkenntnissen zur generellen Gestaltung von Lehrkraftfortbildungen folgend und die Spezifität sowie den konkreten Alltagsbezug der Fortbildungsinhalte beachtend, sollen im folgenden Kapitel anhand von Erkenntnissen zu Lehrkraftfortbildungen im Themenfeld ESE konkrete Qualitätsmerkmale von Fortbildungsangeboten für dieses Professionalisierungsfeld herausgestellt werden.

2.2.3.2 Erkenntnisse zu Lehrkraftfortbildungen zur Förderung der ESE bei Schüler*innen

Im deutschsprachigen Raum existieren nicht viele Studien, die explizit die Wirksamkeit von Lehrkraftfortbildungen zur inklusiven Beschulung im FSP ESE in den Blick nehmen. Leidig et al. (2016) liefern erstmalig ein systematisches Review über die nationale und internationale Befundlage und leiten anhand der Ergebnisse Konsequenzen für wirksame Professionalisierungsmaßnahmen im FSP ESE ab. In ihrer Überblicksarbeit finden Evaluationsstudien Berücksichtigung, die jeweils an konkrete Fortbildungsmaßnahmen geknüpft waren, in Fachzeitschriften nach 1994 veröffentlicht wurden (Peer-Review) und bei denen die Studienteilnehmer*innen Regelschullehrkräfte und/oder Sonderpädagog*innen waren. Unter den auf diese Weise ermittelten 15 englischsprachigen Studien befindet sich bezeichnenderweise keine Wirksamkeitsstudie, die im deutschsprachigen Raum durchgeführt wurde. Einschränkung muss erwähnt werden, dass neben den Studien aus dem Peer-

Review-Verfahren keine weiteren Forschungsarbeiten, wie etwa Doktorarbeiten oder Forschungsberichte, berücksichtigt worden sind.

Das Review stellt heraus, dass die meisten Fortbildungsmaßnahmen auf die Erweiterung des Lehrkraftwissens abzielen, wobei sowohl die Fortbildungsdauer (zwei Stunden bis zwei Jahre) als auch die Teilnehmer*innenanzahl ($n = 1$ bis $n = 145$) eine hohe Varianz aufweisen. Anhand der Vermittlung evidenzbasierter Handlungsstrategien wird auf die gezielte Förderung von sozial-emotionalen Kompetenzen und/oder die Reduktion von (Unterrichts-)Störungen fokussiert. Die Angebote arbeiten überwiegend mit Präsenzterminen in einer Kombination aus verschiedenen methodischen Formen wie Input, Diskussion, Übung und Selbstlernphasen. Die Mehrheit der Fortbildungsmaßnahmen setzt an der konkreten schulischen Praxis der Lehrkräfte an, sodass die evidenzbasierten Maßnahmen an den Einzelfall und den jeweiligen Kontext angepasst und die erarbeiteten Strategien im Unterrichtsalldag erprobt werden können. Die Maßnahmen werden teilweise mit Beratung, Coaching und/oder Feedback kombiniert.

Die Analyse der Evaluationen erfolgte auf der Basis einer Kombination des evidenzbasierten Rahmenmodells zur Evaluation von King (2014) und des erweiterten Angebots-Nutzungsmodells im Kontext beruflichen Lernens von Lipowsky (2014). Im Folgenden sollen die Ergebnisse der ersten vier Wirkungsebenen berichtet werden, da diese im eigenen entwickelten Professionalisierungskonzept Berücksichtigung finden. Die Ergebnisse zeigen auf Ebene eins (Lehrkraftreaktionen) durchgängig eine hohe Zufriedenheit und die positive Einschätzung der Anwendbar- und Machbarkeit der vermittelten Strategien. Auf Ebene zwei (Erweiterung Lehrkraftwissen/-kognitionen) werden Effekte im Bereich deklarativen und prozeduralen Wissens (sieben Fortbildungsmaßnahmen) und eine positive Entwicklung der Einstellungen zur Inklusion von Schüler*innen im FSP ESE berichtet – letztere wird allerdings auch weiterhin skeptisch betrachtet (zwei Fortbildungsmaßnahmen). Außerdem werden in zwei Studien Zuwächse im Selbstwirksamkeitserleben genannt. Auf der Ebene drei (Veränderung unterrichtlichen und pädagogischen Handelns) zeigen sich fast durchweg positive Effekte (neun Fortbildungsmaßnahmen). Die einbezogenen kontrollierten Einzelfallstudien zeigen eine hohe Implementationsqualität der vermittelten Strategien. Ihre Umsetzung führt auf der vierten Ebene (Schüler*innenoutcomes) zu positiven Verhaltensänderungen, sodass Lehrkräfte sich in der Realisierung als erfolgreich erleben und den Nutzen der jeweiligen Handlungsstrategie wahrnehmen konnten. Im Folgenden werden die wichtigsten strukturellen und inhaltlichen Merkmale der Fortbildungsmaßnahmen sowie die zentralen Befunde zur deren Wirksamkeit im Überblick dargestellt.

Tabelle 9: Befunde zu strukturellen und inhaltlichen Merkmalen von Fortbildungsmaßnahmen im FSP ESE und zu deren Wirksamkeit nach Leidig et al. (2016)

Strukturelle und inhaltliche Merkmale von Fortbildungsmaßnahmen im FSP ESE	
Zielgruppe	Regelschullehrkräfte – Primarstufe: 7 Studien – Sekundarstufe I: 3 Studien – mehrere Schulstufen umfassend: 5 Studien
Format/Länge	– Zeitspanne: zwei Stunden bis zwei Jahre – Umfang: mind. zwei Arbeitseinheiten
Schwerpunkte/ Inhalte	– Verhaltens- und Aufmerksamkeitsprobleme (3 Studien), Gefühls- und Verhaltensstörungen (7 Studien), Lern- und Entwicklungsstörungen unter besonderer Berücksichtigung von Gefühls- und Verhaltensstörungen (4 Studien) – Vermittlung von Grundlagenwissen im FSP ESE; konkrete Handlungsstrategien im Unterricht unter Berücksichtigung der Kriterien effektiven Classroom-Managements und evidenzbasierter Praktiken wie funktionale/strukturelle Verhaltensbeurteilung und -unterstützung einschließlich der Durchführung und Evaluation des eigenen Unterrichts (5 Studien), Classroom-Management, Strategien zur Verstärkung positiven Verhaltens und Training sozialer Kompetenzen (8 Studien), Arbeiten in multiprofessionellen Teams (2 Studien)
Didaktisch-methodische Konzeption	– E-learning: 1 Studie – Präsenztermine (14 Studien) mit Kombination verschiedener Vermittlungsformen wie Input, Diskussion, Übung, Selbstlernphasen (4 Studien); Erprobung der erarbeiteten Strategien im Unterrichtsalltag (12 Studien); Elemente der Beratung, des Coachings und Feedbacks (8 Studien); Hospitationen i. S. von Demonstrationen im Best-Practice-Feld (1 Studie); Analyse von Videobeispielen (2 Studien)

Befunde zur Wirksamkeit der Fortbildungsmaßnahmen	
Fokussierte Evaluations-ebenen	– mehrere Ebenen (12 Studien) und eine Ebene (3 Studien) fokussierend – Ebene 1 (7 Studien) – Ebene 2-4 (10 Studien)
Stichproben-größe	– Lehrkräfte: $n = 1$ bis $n = 145$; in 7 Studien $n < 10$ – Schüler*innen: $n = 1$ bis $n = 338$
Forschungs-design/-methodik	– Quantitative Designs (Kontrollierte Einzelfallstudien im MBD, alternierende Versuchspläne, ABA- und ABAB-Pläne, vorexperimentelle und/oder quasiexperimentelle Designs): 13 Studien – Mixed-Method-Design: 1 Studie – Aktionsforschungsansatz: 1 Studie – Längsschnittstudie mit drei MZP: 1 Studie – Follow-up: 2 Studien
Evaluations-ergebnisse	Ebene 1 – Hohe Zufriedenheit, (sehr) positive Einschätzungen hinsichtlich der Anwendbarkeit/Machbarkeit der vermittelten Strategien Ebene 2 – Effekte im Bereich deklaratives und prozedurales Wissen (7 Studien) – Positive Entwicklung der Einstellungen zur Inklusion von SuS im FSP ESE, aber Schwerpunkt wird weiterhin skeptisch betrachtet (2 Studien) – Zuwächse im Selbstwirksamkeitserleben (2 Studien) Ebene 3 – positive Effekte (9 Studien), keine Veränderung bei E-Learning-Konzeption (1 Studie) Ebene 4 – Abbau störender/Aufbau angemessener Verhaltensweisen durch den Einsatz von evidenzbasierten Strategien (8 Studien); Ausweitung der Beschulung im Kerncurriculum der Regelschule (1 Studie)

Anmerkungen. Evaluationsebenen (Lipowsky, 2014): Ebene 1 = Lehrkraftreaktionen, Ebene 2 = Erweiterung Lehrkraftwissen/-kognition, Ebene 3 = Veränderung unterrichtlichen und pädagogischen Handelns, Ebene 4 = Veränderung Schüler*innenoutcomes; n = Probandenanzahl; MBD = Multiple Baseline Design; ESE = Emotional-soziale Entwicklung; SuS = Schüler*innen; FSP = Förderschwerpunkt.

Anhand der Ergebnisse leiten Leidig et al. (2016) zwei zentrale Konsequenzen für die Konzeption von Lehrkraftfortbildungen im Feld ESE ab. Zum einen ist die Vermittlung von im beruflichen Alltag umsetzbaren und evidenzbasierten Strategien zum Umgang mit herausfordernden Verhaltensweisen sowie die prozessbezogene Begleitung und Reflexion bei deren Umsetzung ein wichtiges Qualitätsmerkmal. Dies ist nicht durch die im deutschsprachigen Raum üblichen Kurzzeitfortbildungen erreichbar, sondern sollte durch langfristig angelegte Formate realisiert werden, in deren Rahmen insbesondere evidenzbasierte Maßnahmen (Informationen, Übungen, Coaching) genutzt werden. Zum anderen können die besten Maßnahmen für einzelne Schüler*innen bzw. -gruppen durch fallbezogenes Arbeiten ausgewählt werden. Mithilfe „datengestützter Rückkopplungsprozesse“ (Leidig et al. 2016, S. 66), etwa durch den Einsatz von Verhaltensbeobachtung oder -beurteilung, kann zusätzlich die Implementationsqualität erhöht, erlernte Strategien langfristig in das Handlungsrepertoire der Lehrkräfte überführt sowie das Selbstwirksamkeitserleben gestärkt werden. Weiterführend stellen die Autor*innen heraus, dass wirksame Lehrkraftfortbildungen auf die konsequente Unterstützung durch Schulleitung und Schulaufsicht angewiesen sind. Ferner sind Studien vonnöten, die die längerfristige Wirkung von Interventionen untersuchen – insbesondere die Erfassung der Outcomes auf Schüler*innenebene gilt es mehr in den Blick zu nehmen. Das systematische Review zeigt außerdem auf, dass Lehrkräfte eine hohe Zufriedenheit mit universitär angebundenen Fortbildungsmaßnahmen aufweisen, nicht zuletzt, weil hierbei wissenschaftliche Ergebnisse einbezogen wurden und eine hohe Umsetzungstreue gewährleistet werden konnte.

Die Ergebnisse dieses Reviews berücksichtigen in hohem Maße die aus der allgemeinen Wirksamkeitsforschung zu Lehrkraftprofessionalisierung abgeleiteten Merkmale wirksamer Fortbildungen und spezifizieren diese. Im Folgenden werden die zentralen Befunde zusammenfassend dargestellt und Schlussfolgerungen für die Konzipierung von Professionalisierungsmaßnahmen abgeleitet.

2.2.4 Zusammenfassung und Schlussfolgerungen zur Professionalisierung von Lehrkräften für die schulische Förderung der ESE von Kindern

Kompetenzorientierte Ansätze gehen von der Annahme aus, dass die Lehrkraftprofessionalität und die dafür nötigen Kompetenzen grundsätzlich erlernbar sind (Baumert & Kunter, 2006). Demnach kann bei der Entwicklung von Professionalität von verschiedenen Entwicklungsstadien ausgegangen werden, in denen die Lehrkraft zunehmend routinierter Professionswissen, spezialisierte Fertigkeiten und Reflexionsfähigkeiten aufbaut und anwendet (Berliner, 2006; König et al., 2010). Anhand des Modells zur professionellen Handlungskompetenz (Baumert & Kunter, 2006) und dessen Erweiterung um situationsspezifische Skills (Blömeke et al., 2015) sowie inklusive Kompetenzaspekte (Blömeke et al., 2015; Gottfried et al., 2021; König et al., 2019; Leko & Roberts, 2014; Osher et al., 2014), lässt sich ein inklusives Professionalisierungsprofil von Lehrkräften beschreiben. Der darin enthaltene Bereich des inklusiven pädagogisch-psychologischen Wissens (*GPK-IT*, König et al., 2019) und dessen spezifischere Ausprägung emotional-sozialen Lernens (*GPK-IT-SEL*; Gottfried et al., 2021) zeigen einen noch konkreteren Professionalisierungsbedarf auf. Die (Weiter-)Entwicklung professioneller Handlungskompetenz kann in der dritten Phase der Lehrer*innenbildung am effektivsten im Rahmen von Lehrkraftfortbildungen realisiert werden. Wie wirksam eine Fortbildungsmaßnahme ist, lässt sich anhand der Reaktionen, des Wissens/der Kognitionen und des Handelns der Lehrkräfte sowie der Schüler*innenoutcomes untersuchen (Lipowsky, 2010). Fortbildungsmaßnahmen, die auf einer oder mehreren Wirkungsebene/n positive Effekte zeigen, weisen spezifische strukturelle und didaktisch-methodische Kriterien auf (Darling-Hammond et al., 2017; Kurniawati et al., 2014; Leidig 2019, S. 44ff.; Rzejak et al. 2020, S. 21ff.; Rzejak & Lipowsky 2019, S. 20ff; Van Mieghem et al., 2020), die auch für den inklusiven Unterricht handlungsleitend sind. Für Fortbildungsmaßnahmen im Kontext Inklusion im FSP ESE können ebenfalls spezifische erfolgversprechende Kriterien identifiziert werden (Leidig et al., 2016). In Abbildung 9 werden die allgemeinen und spezifischen Erfolgskriterien zusammenfassend dargestellt, die bei der Konzeption von Professionalisierungsmaßnahmen zur Förderung der ESE von Schüler*innen handlungsleitend sein sollten.



Abbildung 9: Allgemeine und spezifische, auf den FSP ESE bezogene Kriterien erfolgreicher Lehrkraftfortbildungen, die als Grundlage für die Konzipierung von Professionalisierungsmaßnahmen im Kontext Inklusion dienen können/sollten (Darling-Hammond et al., 2017; Kurniawati et al., 2014; Leidig 2019, S. 44ff.; Leidig et al., 2016; Rzejak et al. 2020, S. 21ff.; Rzejak & Lipowsky 2019, S. 20ff; Van Mieghem et al., 2020) (eigene Darstellung)

Anmerkung. FSP ESE = Förderschwerpunkt Emotionale-soziale Entwicklung.

2.3 Gesamtzusammenfassung zur schulischen Förderung der ESE durch Lehrkraftprofessionalisierung

Anhand des Forschungsstands zu Kindern mit Verhaltensauffälligkeiten und emotional-sozialen Entwicklungsstörungen (vgl. Kap. 2.1.2) und ihrer schulischen Förderung (vgl. Kap. 2.1.3), ergeben sich konkrete Desiderate bzw. Entwicklungsaufgaben für die Professionalisierung der Lehrkräfte (vgl. Kap. 2.2). Im Folgenden werden die zentralen Argumentationslinien dargestellt (Abbildung 10) und zusammengefasst. Sie leiten über zu dem in Kapitel 3 ausführlich dargestellten Professionalisierungskonzept für Grundschullehrkräfte zur Erweiterung ihrer Lehrkraftkompetenzen für die wirksame Förderung von Schüler*innen im Bereich ESE.

Bei knapp einem Fünftel der Grundschulkinder in Deutschland können Auffälligkeiten im emotionalen Erleben und sozialen Handeln beobachtet werden (Klasen et al., 2017; Otto et al., 2021). Haben sich Verhaltensauffälligkeiten erst einmal als emotional-soziale Entwicklungsstörungen manifestiert, so zeigen sich hierbei oftmals anhaltende bzw. wiederkehrende und kumulierende Risikokonstellationen (Beelmann & Raabe, 2007; Ghandour et al., 2019; Salmanian et al., 2019; Visser et al., 2019) mit ungünstigen Entwicklungsprognosen (Ihle & Esser, 2008). Eine Diagnostik erfolgt in aller Regel aber erst dann, wenn eine erhebliche Problematik erkannt wurde (Wait-to-fail-Problem und Etikettierungs-Ressourcen-Dilemma). Dies führt zur Forderung eines frühzeitigen und niederschweligen Eingreifens, das je nach Ausprägung der Auffälligkeit bzw. Störung anhand universeller, selektiver oder indizierter Maßnahmen realisiert werden sollte (Beelmann & Raabe, 2007; Cheney et al., 2008; Gresham et al., 2005; Heinrichs et al., 2013; Hillenbrand, 2015). Die Schule gilt dabei als essentielles Präventionssetting für die Förderung von Kindern (Beelmann, 2008; Heinrichs et al., 2013; Ihle & Esser, 2008; Reicher & Jauk, 2012). Schulbasierte Präventionsmaßnahmen sollten möglichst theoretisch fundiert und evidenzbasiert sein (Blumenthal et al., 2020; Hillenbrand, 2015). Die im schulischen Kontext geforderte evidenzbasierte Praxis (Blumenthal & Mahlau, 2015; Grosche, 2017) erfordert neben der externen und sozialen Evidenz auch die praktische Expertise und Professionalität der Pädagog*innen (interne Evidenz). Hier formiert sich ein inklusives Professionalisierungsprofil, das im Bereich ESE insbesondere inklusives pädagogisch-psychologisches Wissen (GPK-IT; König et al., 2019) und dessen spezifischere Ausprägung emotional-sozialen Lernens (GPK-IT-SEL; Gottfried et al., 2021) impliziert. Ausgehend von kompetenzorientierten Professionalisierungsansätzen, die die grundsätzliche Erlernbarkeit professioneller Handlungskompetenz postulieren (Baumert & Kunter,

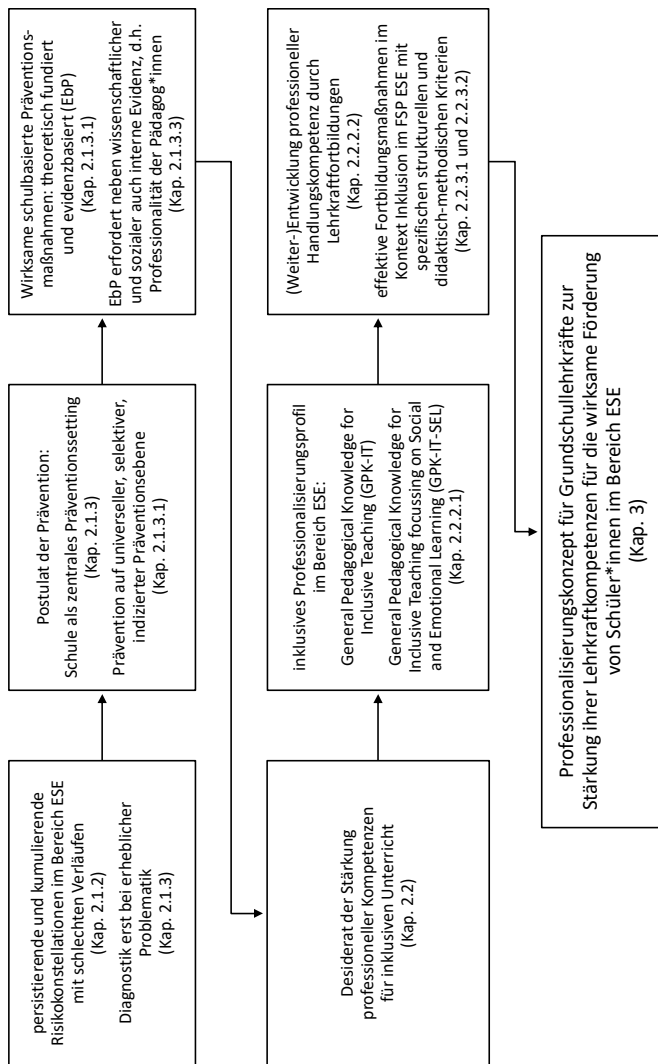


Abbildung 10: Gesamtzusammenfassung des Forschungsstands zur schulischen Förderung der ESE von Kindern durch Lehrkraftprofessionalisierung; Zentrale Argumentationslinien und Ableitung des entwickelten Professionalisierungskonzepts für Grundschullehrkräfte zur Erweiterung spezifischer Lehrkraftkompetenzen (eigene Darstellung)

Anmerkungen. ESE = emotional-soziale Entwicklung, EbP = Evidenzbasierte Praxis, FSP ESE = Förderschwerpunkt Emotional-soziale Entwicklung.

2006), können für die (Weiter-)Entwicklung der professionellen Kompetenzen in der dritten Phase der Lehrer*innenbildung vor allem Fortbildungen als zentrale Lerngelegenheiten für Lehrkräfte identifiziert werden (u.a. Darling-Hammond et al., 2017; Kunter, 2011; Lipowsky & Rzejak, 2012; Timperley et al., 2007). Effektive Fortbildungsmaßnahmen weisen spezifische strukturelle und didaktisch-methodische Kriterien auf (Darling-Hammond et al., 2017; Kurniawati et al., 2014; Leidig 2019, S. 44ff.; Rzejak et al. 2020, S. 21ff.; Rzejak & Lipowsky 2019, S. 20ff; Van Mieghem et al., 2020), die für den Kontext Inklusion im FSP ESE noch weiter konkretisiert werden können (Leidig et al., 2016). Das Konglomerat der Erkenntnisse aus dem Forschungsstand fließt in die konkrete Ausgestaltung eines Professionalisierungskonzepts für Grundschullehrkräfte ein, welches im folgenden Kapitel vorgestellt wird.

3 Entwurf eines Professionalisierungskonzepts zur Förderung der ESE von Grundschüler*innen durch Erweiterung spezifischer Lehrkraftkompetenzen

Auf Basis der Leitfrage „Kann die emotionale und soziale Entwicklung von Kindern durch die Erweiterung spezifischer Lehrkraftkompetenzen gefördert werden?“ und unter Berücksichtigung des Forschungsstandes zur schulischen Förderung von Kindern durch Lehrkraftprofessionalisierung im Kontext der Inklusion im FSP ESE wurde ein prozessbegleitendes Professionalisierungskonzept für Grundschullehrkräfte entworfen. Bevor die Konzeption detailliert beschrieben wird, werden zunächst die Ziele und konzeptionellen Überlegungen dargelegt. Überlegungen zu den Stärken und Schwächen der Konzeption schließen das Kapitel ab und leiten über zur Evaluationsstudie.

3.1 Ziele und konzeptionelle Überlegungen

Das Hauptziel des Professionalisierungskonzepts ist die Veränderung der Schüler*innenoutcomes durch Stärkung der professionellen Handlungskompetenzen ihrer Lehrkräfte. Im Rahmen der Professionalisierungsmaßnahme sollen diejenigen Schüler*innen in den Blick genommen werden, die in der Schule herausfordernde Verhaltensweisen zeigen und die mit der universellen, unspezifischen Förderung im regulären Unterricht, beispielsweise durch hochwertiges Classroom Management, nicht genügend angesprochen werden können (vgl. Kap. 2.1.3). Bezogen auf die Modellierung professioneller Handlungskompetenz als Kontinuum und deren Erweiterung um Wissens- und Kompetenzmerkmale für den inklusiven Unterricht, erfordert dies die Professionalisierung der Lehrkräfte im Bereich inklusiven pädagogisch-psychologischen Wissens mit dem Fokus auf die ESE der Schüler*innen. Das Professionalisierungskonzept baut somit auf Facetten pädagogisch-psychologischen Wissens wie effektive Klassenführung, Umgang mit Heterogenität, Anwendung vielfältiger Unterrichtsmethoden und pädagogischer Maßnahmen sowie Diagnostik auf und zielt auf deren Erweiterung um inklusive Aspekte (sonder-)pädagogischen Wissens, wie insbesondere die schulische Diagnostik/Diagnose und die präventive Förderung/Intervention ab (vgl. Kap.

2.2.2.1). Die Konzeption fokussiert außerdem die in der Forschung zu Lehrkraftfortbildungen im Kontext der Inklusion im FSP ESE identifizierten Professionalisierungsbedarfe (Leidig 2016, S. 65):

- Entwicklung einer positiven Haltung gegenüber Kindern mit dem FSP ESE
- Aufbau von Wissen über Gefühls- und Verhaltensstörungen
- Vermittlung von Handlungsstrategien zum Umgang mit herausfordernden Verhaltensweisen
- Ermöglichung der erfolgreichen Erprobung der Handlungsstrategien im Alltag
- Stärkung des Selbstwirksamkeitserlebens der Lehrkräfte

Die Modellierung dieses anvisierten Professionalisierungs- und Förderprofils erfolgt anhand des mehrstufigen präventiven Differenzierungsansatzes des RTI-Modells, wie in Kap. 2.2.2.1 beschrieben. Dies erlaubt eine genauere Darstellung des auszubauenden Professionswissens der Lehrkräfte sowie des differenzierteren Förderverfahrens der Schüler*innen. Das Professionalisierungskonzept setzt auf der RTI-Förderebene I (Primäre Prävention) mit dem Wissen um und die Reflektion über das eigene Classroom Management an und zielt darauf ab, dieses anhand von individuelleren pädagogischen Maßnahmen und engmaschiger Verhaltensverlaufsdiagnostik (RTI-Ebene II; Sekundäre Prävention) zu erweitern. Die Wirkung der Professionalisierungsmaßnahme und die damit einhergehenden erweiterten und neu erworbenen Lehrkraftkompetenzen sollten dann anhand positiver Schüler*innenoutcomes (emotional-soziales und Regelverhalten) sichtbar werden. Abbildung 11 veranschaulicht das im Rahmen des entwickelten Professionalisierungskonzepts zu erweiternde Professionswissen der Lehrkräfte, gegliedert nach Wissensbereichen, Wissensfacetten (vgl. COACTIV-Kompetenzmodell; Baumert & Kunter, 2006) und Schüler*innenoutcomes sowie gestaffelt nach den RTI-Förderebenen Primäre Prävention, Sekundäre Prävention und Tertiäre Prävention (Gresham et al., 2005).

Von diesem Forschungsstand ausgehend und das Hauptziel fokussierend – die Veränderung der Schüler*innenoutcomes durch Stärkung spezifischer professioneller Kompetenzen ihrer Lehrkräfte – wurden drei Handlungsfelder benannt und dazugehörige Professionalisierungselemente entwickelt:

- A) Erweiterung des Lehrkraftwissens
- B) Veränderung des unterrichtspraktischen Handelns
- C) Reflexion durch Beratung

Handlungsfeld A zielt auf die Erweiterung des Lehrkraftwissens ab. Die dazu eingesetzten Professionalisierungselemente werden im Rahmen zweier mehrstündiger Fortbildungsveranstaltungen eingeführt. Zum einen werden hierbei

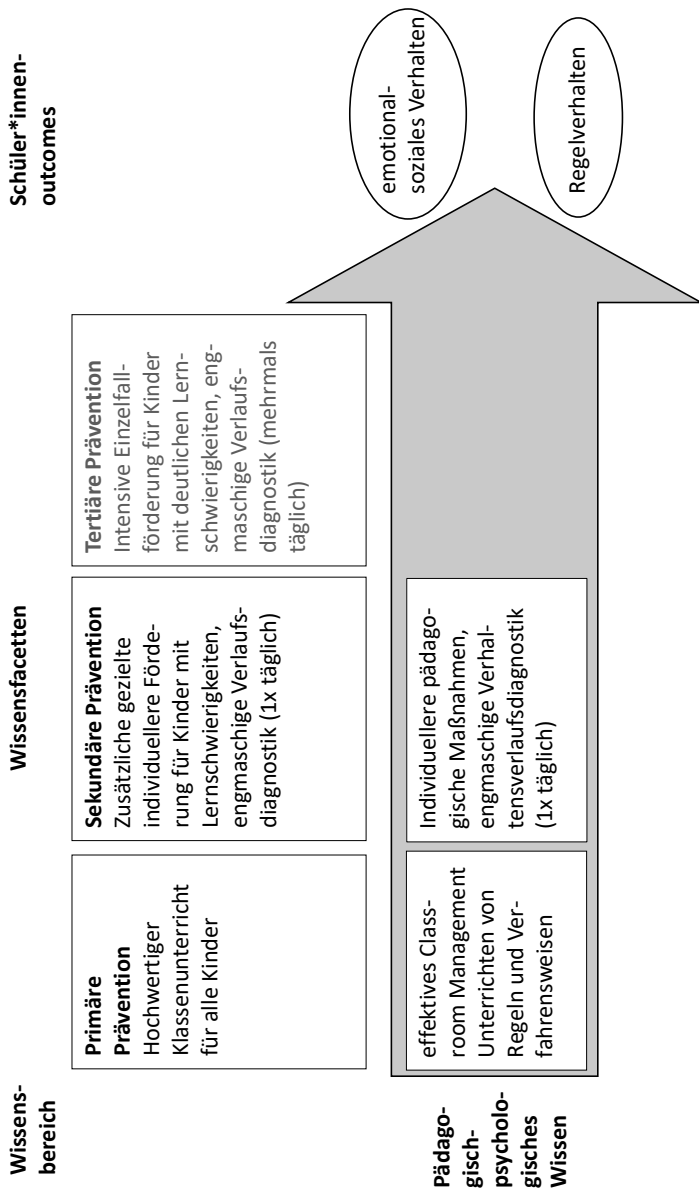


Abbildung 11: Das im Rahmen des entwickelten Professionalisierungskonzepts zu erweiternde Professionswissen der Lehrkräfte gegliedert nach Wissensbereichen, Wissensfacetten und Schüler*innenoutcomes (eigene Darstellung)

Grundlagen der ESE von Kindern und etwaiger Störungsformen behandelt, zum anderen erfolgt die Reflexion über und die Wissenserweiterung zu Handlungsstrategien der Primären und Sekundären Prävention. Da die Zielgruppe der Kinder auf der Ebene der Primären Prävention nicht genügend gefördert werden kann, liegt der Fokus in den Fortbildungsveranstaltungen vor allem auf pädagogischen Maßnahmen im Bereich der Sekundären Prävention. So erarbeiten die Lehrkräfte theoriegeleitet Maßnahmen für die individuellere unterrichtsintegrierte Förderung einzelner Schüler*innen und werden in das änderungssensitive verhaltensverlaufdiagnostische Verfahren *Direct Behavior Rating* (DBR; vgl. Kap. 2.1.3.4) eingeführt.

Handlungsfeld B strebt die Veränderung des unterrichtlichen Lehrkrafthandelns an. So sind die Lehrkräfte angehalten, die in den Fortbildungen erworbenen Kenntnisse und Fertigkeiten direkt im Anschluss an die Veranstaltungen im Unterricht anzuwenden. Hierbei ist zum einen der unterrichtliche Einsatz von Handlungsstrategien der Sekundären vor dem Hintergrund Primärer Prävention anhand des Manuals *84 Handlungsmöglichkeiten* (vgl. Kap. 2.1.3.4) vorgesehen. Zum anderen kommt das prozessbegleitende Verhaltensmonitoring mittels DBR zum Einsatz.

Handlungsfeld C umfasst die psychosoziale Beratung der Lehrkräfte. Durch eine kooperativ-systemische Fallberatung werden die Lehrkräfte bei der Umsetzung der *84 Handlungsmöglichkeiten* begleitet und es wird ihnen die Möglichkeit der Reflexion der angewandten unterrichtlichen Maßnahmen angeboten. Die Handlungsfelder des Professionalisierungskonzepts und die dazugehörigen Professionalisierungselemente werden in Abbildung 12 dargestellt.

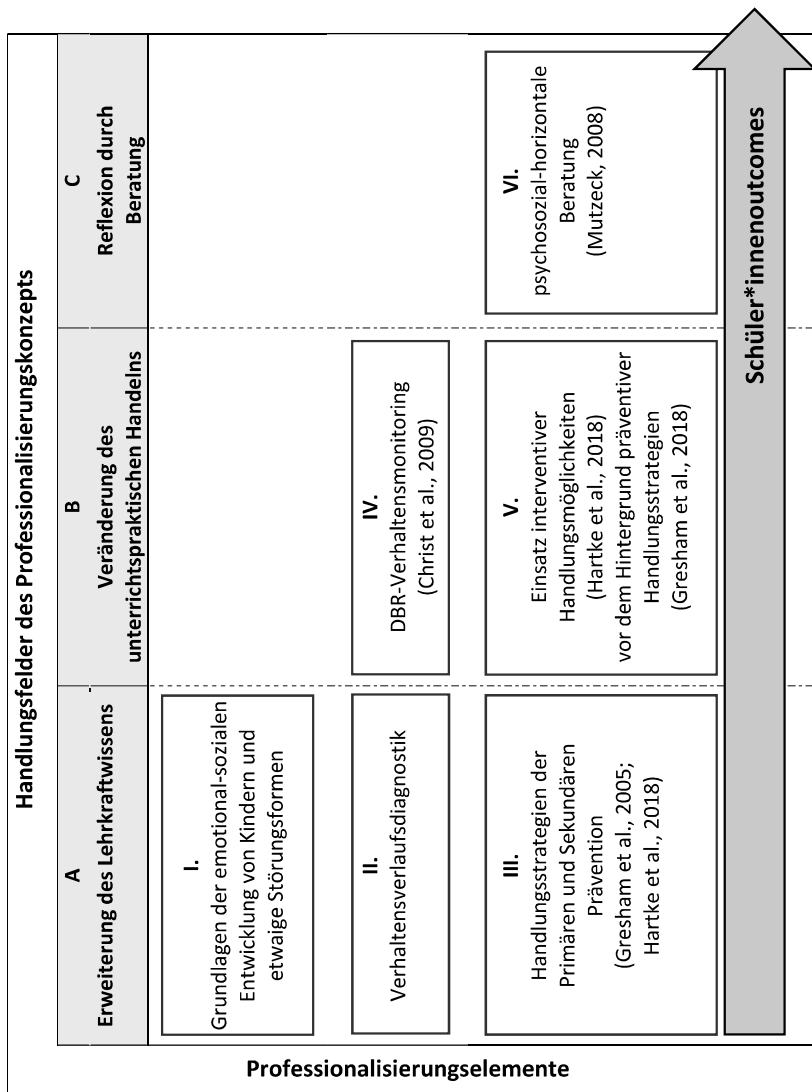


Abbildung 12: Handlungsfelder im Professionalisierungskonzept mit den dazugehörigen Professionalisierungselementen (eigene Darstellung)

Anmerkung. DBR = Direct Behavior Rating.

3.2 Handlungsfelder und Professionalisierungselemente

3.2.1 Handlungsfeld A: Erweiterung des Lehrkraftwissens

Handlungsfeld A stellt die Erweiterung des Lehrkraftwissens dar. Die Vermittlung der Inhalte sowie die Einführung und Erprobung der Professionalisierungselemente werden im Rahmen zweier mehrstündiger Fortbildungssitzungen (Dauer insgesamt: 7 Std.) realisiert. Im Folgenden sollen die Professionalisierungselemente anhand der didaktischen Planung der entsprechenden Fortbildungseinheiten einzeln erläutert werden.

3.2.1.1 Professionalisierungselement I: Grundlagen der ESE von Kindern und etwaige Störungsformen

Das erste Professionalisierungselement stellt die Vermittlung von Grundlagenwissen im FSP ESE dar. Zu Beginn der Fortbildungseinheit findet eine Sensibilisierung bezüglich schulischen Regelverhaltens statt. Die Teilnehmenden überlegen, wodurch ihnen „schwierige“ Schüler*innen im Schulalltag auffallen. Die Nennungen werden durch den wissenschaftlichen Stand und um die Konzepte der internalisierenden und externalisierenden Verhaltensweisen (vgl. Kap. 2.1.2.2) ergänzt. Anschließend werden einige der genannten Verhaltensweisen mit Symptomen klinischer Diagnosen in den Bereichen Motivation, Aufmerksamkeit, Aggression und Angst verglichen und deren Prävalenzen, Ursachen, Symptomatik und spezifische Verlaufsformen besprochen. Viele der genannten Verhaltensweisen und Symptome lassen sich auch unter dem Aspekt schulischen Regelverhaltens betrachten. Zusammen mit den Teilnehmenden werden gesellschaftliche und damit auch schulische Regeln reflektiert und als notwendig für ein erfolgreiches Zusammenleben und -arbeiten in der Gemeinschaft herausgestellt. Die Teilnehmenden werden angeleitet, verschiedene Kriterien für schulische Regeln zu definieren (z.B. logische Nachvollziehbarkeit, Transparenz, Allgemeingültigkeit). Abschließend werden unterschiedliche Gründe thematisiert, warum Schüler*innen Regeln in der Schule nicht einhalten. Neben schülerbedingten Faktoren werden hierbei vor allem auch systemische Faktoren, wie Intransparenz, Bezugslosigkeit, fehlende Konsequenzen, zu hoher Abstraktionsgrad oder fehlende Logik besprochen und diskutiert.

3.2.1.2 Professionalisierungselement II: Verhaltensverlaufsdiagnostik

Das zweite Professionalisierungselement stellt die Einführung in die Verhaltensverlaufsdiagnostik dar. Im Fortbildungsmodul werden die Teilnehmen-

den in die verhaltensverlaufsdiagnostische Methode DBR (vgl. Kap. 2.1.3.4) eingeführt. Neben der Entwicklung und Formulierung konkreter und gut zu beobachtender Verhaltensziele (Operationalisierung) lernen die Teilnehmenden, welche Art von Ratingskala sich für die Messung eignet, wie sie die Verhaltensbeurteilung planen und diese im Schulalltag ökonomisch einsetzen können (Huber, 2015a; Mahlau, 2019). Abschließend werden die Teilnehmenden in der Beobachtung und Beurteilung von Verhalten geschult und üben dies anhand von medial vermittelten Beispielen. In diesem Zuge werden auch Aspekte der Interraterreliabilität und somit die Notwendigkeit der konstanten Verhaltensbeurteilung durch dieselbe Person besprochen (Conley et al., 2014; Huber, 2015a). Ziel des Fortbildungsmoduls ist, dass die teilnehmenden Lehrkräfte für ein bis drei konkrete, gut zu beurteilende Verhaltensbereiche für ein bis zwei Schüler*innen ihrer Klasse operationalisiert haben. Im Nachgang der Fortbildung sollen die Lehrkräfte beginnen, eine Verhaltensverlaufsdiagnostik bzw. ein Verhaltensmonitoring für jedes ausgewählte Kind zu erstellen.

3.2.1.3 Professionalisierungselement III: Handlungsstrategien der Primären und Sekundären Prävention

Im dritten Professionalisierungselement erfolgt die Reflexion über präventive Handlungsstrategien sowie die Wissenserweiterung zu pädagogischen Handlungsmöglichkeiten im Unterricht. Als Grundlage für das entsprechende Fortbildungsmodul dient die Handreichung *84 Handlungsmöglichkeiten*. Die Systematik des Fördermaterials beruht auf einer datenbasierten Planung, Durchführung und Evaluation unterrichtsintegrierter Förderung (vgl. Kap. 2.1.3.4).

Zu Beginn wird das in den vorangegangenen Fortbildungsmodulen erarbeitete Wissen zu regelkonformem Verhalten aufgegriffen und im Dialog erweitert (Auftragsklärung). Hierbei sollen die Teilnehmenden Fragen und Beobachtungen aus ihrem pädagogischen Alltag einbringen können. Folgende Themen können bearbeitet werden: Wie kann man Regeln und Verhaltensziele mit der Klasse erarbeiten? Was ist beim Formulieren der Regeln oder Verhaltensziele zu beachten? Welche Möglichkeiten der Visualisierung haben sich bewährt? Anschließend soll zunächst die Förderung von Schüler*innen auf der Ebene der Primären Prävention in den Fokus genommen werden. Von den Erfahrungen der Teilnehmenden ausgehend wird diskutiert, wie soziale Kompetenzen gezielt gefördert werden können, wenn mit Klassenzielen, anstatt begrenzenden Regeln und Sanktionen gearbeitet und deren Erreichung unterrichtsimmanent positiv verstärkt wird (Hartke et al. 2018, S. 131). Anhand einer Checkliste zur Selbstüberprüfung des eigenen Classroom Managements (ebd., S. 134 ff.) beschäftigen sich die Teilnehmenden anschließend mit weiteren Möglichkeiten der Unterrichts- und Beziehungsgestaltung und aktualisieren ihr Wissen zu pro-aktiven pädagogischen Inter-

ventionen, die die gesamte Klasse beim Erlernen unterrichtsbezogenen Verhaltens unterstützt.

Anschließend rücken re-aktive, interventive Maßnahmen der Ebene der Sekundären Prävention in den Vordergrund, die Schüler*innen mit Verhaltensauffälligkeiten und emotional-sozialen Entwicklungsstörungen in besonderer Weise fördern. In dieser Phase erhalten die Teilnehmenden die Gelegenheit, ihr Wissen zu aktualisieren und ggf. zu erweitern sowie insbesondere auch ein Grundverständnis zum Lernen von Verhalten aufzubauen, welches den Menschenbildannahmen der Humanistischen Psychologie folgt (Mutzeck 2008, S. 48 ff.).

Anhand der Planungshilfe „Förderungsorientierte Auswertung der schulischen Entwicklung des Verhaltens und der Entwicklung (1)“ (Hartke et al. 2018, S. 15) konkretisieren die Teilnehmenden die Förderziele für ihre Schüler*innen, für die sie seit der ersten Fortbildung bereits täglich ein Verhaltensmonitoring per DBR erstellen. Anhand dieser Förderziele lernen die Teilnehmenden vordergründige Problemfelder zu benennen, für die es in dem genutzten Werk jeweils spezifische Planungshilfen gibt. Entsprechend der für die eigenen Schüler*innen identifizierten Problembereiche erarbeiten die Teilnehmenden in Gruppen bis zu drei verhaltenstheoretische, kognitionspsychologische und/oder schüler*innenzentrierte Handlungsmöglichkeiten (vgl. Kap. 2.1.3.4) zur unterrichtsintegrierten Förderung in dem jeweiligen Bereich. In einer sich anschließenden Diskussionsphase stellen die Teilnehmenden die individuell gewählten pädagogischen Interventionen im Plenum vor (Förderplanung).

3.2.2 Handlungsfeld B: Veränderung des unterrichtlichen Lehrkrafthandelns

Handlungsfeld B im prozessbegleitenden Professionalisierungskonzept stellt die Veränderung des Lehrkrafthandelns dar. Die dazugehörigen Professionalisierungselemente werden im Folgenden einzeln erläutert.

3.2.2.1 Professionalisierungselement IV: Verhaltensmonitoring

Das vierte Professionalisierungselement sieht den unterrichtlichen Einsatz von Verhaltensmonitoring anhand der Erhebung verhaltensverlaufsdiagnostischer DBR-Daten vor. Im Rahmen des ersten Fortbildungsmoduls operationalisieren die Lehrkräfte das auffällige Verhalten von ein bis zwei Schüler*innen ihrer Klasse, um dieses in den nachfolgenden Wochen anhand eines Verhaltensmonitorings täglich beobachten und einschätzen zu können. Die auf diese Weise erstellte Verhaltensverlaufsdiagnostik hat zum Ziel, die Lehrkräfte für das konkrete Verhalten der zu fördernden Schüler*innen noch stärker zu sensibilisieren, dieses vor und während des Einsatzes der interven-

tiven Handlungsmöglichkeiten valide abzubilden und somit etwaige kleinschrittige Fördererfolge sichtbar machen zu können (vgl. Kap. 2.1.3.4).

3.2.2.2 Professionalisierungselement V: Einsatz interventiver Handlungsmöglichkeiten vor dem Hintergrund präventiver Handlungsstrategien

Das fünfte Professionalisierungselement zielt auf unterrichtlichen Einsatz interventiver Handlungsmöglichkeiten (*84 Handlungsmöglichkeiten*) ab. Nach dem Absolvieren aller Fortbildungsmodule wenden die Lehrkräfte die erarbeiteten Handlungsmöglichkeiten zur unterrichtsintegrierten Förderung für ein bis zwei zu fördernde Schüler*innen in ihrem Unterricht an. Parallel dazu führen sie weiterhin das Verhaltensmonitoring durch und dokumentieren außerdem, an welchen Tagen sie die Handlungsmöglichkeiten einsetzen und an welchen nicht. So können im Nachgang Aussagen über die Umsetzungstreue und Implementationsqualität getroffen werden. Für die Einsatzfrequenz der *84 Handlungsmöglichkeiten* gibt es keine Empfehlungen oder Richtwerte (Hartke et al., 2018). Wie oft die jeweiligen Handlungsmöglichkeiten von den Lehrkräften eingesetzt werden, hängt unter anderem von deren methodischer Ausrichtung, der Möglichkeit ihrer praktischen Durchführung und dem konkreten Schüler*innenverhalten ab. Die Lehrkraft, als der*die verantwortliche, die einzelne Situation mit ihren speziellen Bedingungen erfassende Pädagog*in, entscheidet, in welcher Frequenz und Häufigkeit die Handlungsmöglichkeiten zur Anwendung kommt (Hartke et al., 2018). Um die ausgewählten Schüler*innen jedoch wirksam und nachhaltig in ihrem Verhalten zu fördern, wird im Rahmen der Durchführung des Professionalisierungskonzepts der häufige und regelmäßige Einsatz angestrebt. Anhand dieser Erwägungen wurde zur Beschreibung der Einsatzhäufigkeit der *84 Handlungsmöglichkeiten* folgende quantitative Abstufung festgelegt: sehr oft (= 80-100 %), oft (= 60-80 %), gelegentlich (= 40-60 %), selten (= 20-40 %) und kaum (< 20 %). In Vorbereitung auf die Interventionsphase werden die Lehrkräfte aufgefordert, die Handlungsmöglichkeiten an mindestens jedem zweiten Tag einzusetzen. Dies entspricht einer Einsatzhäufigkeit von 50 % und damit einem gelegentlichen Einsatz. Der Wert von 40 % sollte idealerweise nicht unterschritten werden.

3.2.3 Handlungsfeld C: Reflexion durch Beratung

Im Handlungsfeld C findet Reflexion durch Beratung statt. Während des unterrichtlichen Einsatzes der *84 Handlungsmöglichkeiten* haben nach dem Zufallsprinzip ausgewählte Lehrkräfte die Möglichkeit, eine 90-minütige psychosozial-horizontale Beratung in Anspruch zu nehmen. Die Gespräche finden jeweils in den Räumlichkeiten der Schule statt. Auch nicht ausgewähl-

te Lehrkräfte können nach Ende des Projekts kostenfrei an einem Beratungsgespräch teilnehmen.

3.2.3.1 Professionalisierungselement VI: Psychosozial-horizontale Beratung

Die psychosozial-horizontale Beratung der Lehrkräfte basiert weitestgehend auf der Struktur individueller kooperativer Beratungssitzungen zu problematischen Schulsituationen nach Mutzeck (2008). Die kooperative Beratung ist eine personenzentrierte und ressourcenorientierte Gesprächsführung, welche die berufliche Entwicklung von Lehrpersonen unterstützt und fördert. Die systematische Struktur des Gesprächsverlaufs – vorgesehen sind klar definierte Elemente und eine spezifische Abfolge an Beratungsschritten – ermöglicht und unterstützt ein logisches und flexibles Arbeiten. Die Gespräche werden von einer Diplom-Psychologin durchgeführt, die als ausgebildete systemische Therapeutin weitere Gesprächsmethoden einbringen kann. Die systemische Therapie ist ein psychotherapeutisches Verfahren, dessen Schwerpunkt auf dem sozialen Kontext psychischer Störungen liegt. Probleme werden nicht als Störung eines einzelnen Menschen begriffen, sondern als Folge einer Störung im sozialen Umfeld des Individuums – also des Systems (Schlippe & Schweitzer, 2012). Folgende Grundhaltungen können sowohl für die systemische als auch die kooperative Beratung als Basis definiert werden: Neutralität, Ressourcen- und Lösungsorientierung, Verzicht auf instruktive Interaktion, Kontextsensibilität, Auftragssensibilität/Klient*innenorientierung, Symptommutzung statt Symptombekämpfung, Berater*innen sind keine Expert*innen für die Lösung und positive Konnotationen. Letzteres meint die wertschätzende Beschreibung von Verhaltensweisen, Symptomen und Interaktionen, die üblicherweise als problematisch definiert werden (Mutzeck, 2008). Dies wird auch als „Reframing“ bezeichnet. Die freie Wählbarkeit der Inhalte als Grundgedanke der psychosozial-horizontalen Beratung ist eine zwingende Voraussetzung für das gute Gelingen der Gespräche (Mutzeck, 2008; Schlippe & Schweitzer, 2012).

3.2.4 Zusammenfassung

Bei der entwickelten Konzeption steht der Gedanke der Prozessbegleitung im Vordergrund. So sind die drei Handlungsfelder und die darin integrierten Professionalisierungselemente nicht als separate, für sich stehende, sondern vielmehr als aufeinander aufbauende und miteinander verzahnte Einheiten zu sehen. Bei dem vorliegenden Professionalisierungskonzept handelt es sich außerdem um ein langfristig angelegtes Format, welches die Lehrkräfte von Handlungsfeld A bis C über ein halbes Schuljahr hinweg begleitet. Die Erprobung des prozessbegleitenden Professionalisierungskonzepts wurde an-

hand einer Evaluationsstudie wissenschaftlich untersucht, die sich über das gesamte Schuljahr 2018/19 erstreckte. In Abbildung 13 (auf der nächsten Seite) sind der Aufbau, die Inhalte und die praktische Durchführung des Konzepts sowie die empirisch untersuchten Wirkungsebenen zusammenfassend dargestellt.

3.3 Überlegungen zu den Stärken und Schwächen der Konzeption

Die vorliegende Konzeption zur Professionalisierung von Grundschullehrkräften für die pädagogische Arbeit mit Schüler*innen mit Verhaltensschwierigkeiten wurde unter der Berücksichtigung der Erkenntnisse aus der allgemeinen Wirksamkeitsforschung zu Lehrkraftfortbildungen (vgl. Kap. 2.2.3.1) und anhand der Merkmale wirksamer Lehrkraftfortbildungen zur inklusiven Beschulung im FSP ESE (vgl. Kap. 2.2.3.2) entworfen. Planung und Umsetzung des Professionalisierungskonzepts unterlagen dabei organisatorischen, personellen sowie ökonomischen Einschränkungen. So war die Durchführung des Konzepts zum einen aufgrund der Projektlaufzeit auf ein Schuljahr begrenzt, zum anderen wurde die Durchführung und Evaluation der Maßnahmen im Wesentlichen durch eine halbe wissenschaftliche Mitarbeiterstelle geleistet, deren finanzielle Ausstattung begrenzt war. Diese Rahmenbedingungen eröffneten Konzeptionsspielräume, sie setzten dem theoretisch Möglichen und wissenschaftlich Geforderten aber auch Grenzen. Im Folgenden sollen die wichtigsten Qualitätsmerkmale des Professionalisierungskonzepts, aber auch dessen vermutete Schwächen herausgestellt werden.

Die Konzeption weist folgende vermutete Stärken auf:

- Vermittlung von im beruflichen Alltag umsetzbaren und evidenzbasierten Strategien: Fortbildungsmaßnahmen im Kontext der Inklusion im FSP ESE sollten die Vermittlung von Strategien im Umgang mit herausforderndem Schüler*innenverhalten ermöglichen, die evidenzbasiert und im beruflichen Alltag umsetzbar sind (Hillenbrand, 2015; Leidig et al. 2016, S. 66). Mit dem Einsatz des Manuals *84 Handlungsmöglichkeiten* kommen pädagogische Strategien und Interventionen zum Einsatz, die explizit für den praktischen Einsatz im Unterricht konzipiert wurden und für die erste Wirksamkeitsbelege vorliegen (Hartke & Vrbán, 2011). Der Einsatz der *84 Handlungsmöglichkeiten* wird außerdem anhand des verhaltensverlaufdiagnostischen Messinstruments DBR begleitet, das in der Forschung bereits umfassend evaluiert wurde. Regelmäßige, kurze Messungen erlauben es, Verhaltensänderungen bzw. Fördererfolge auch in einem bewegten Schulalltag valide, zeitnah und effizient abbilden zu können (Casale et al., 2017; Christ et al., 2009; Huber & Rietz, 2015).

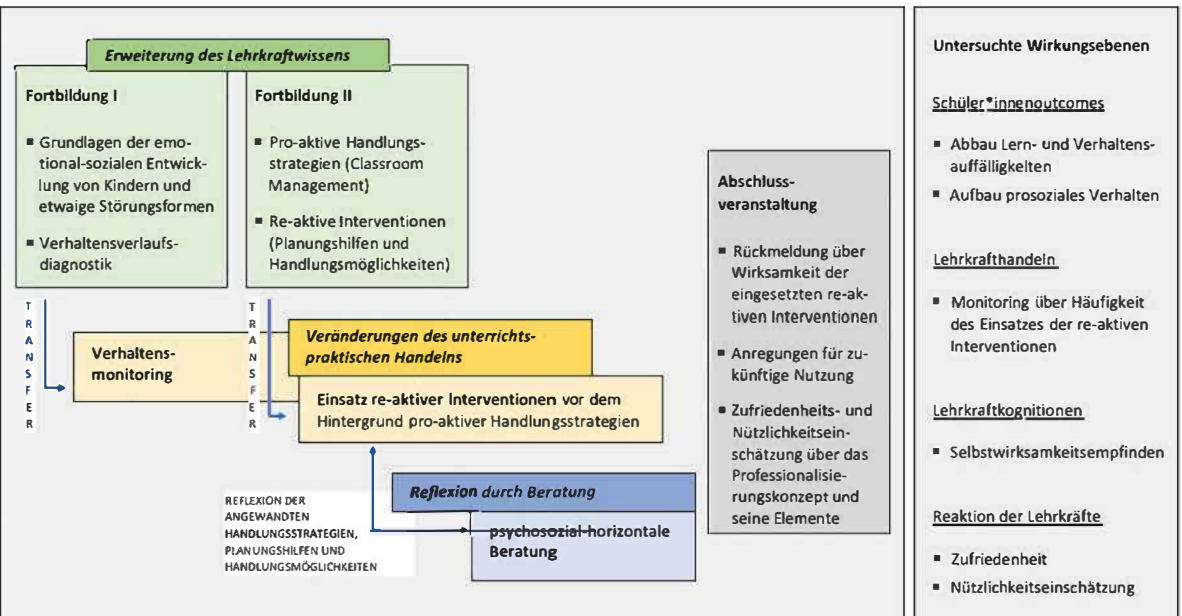


Abbildung 13: Aufbau und Ablauf des Professionalisierungskonzepts für Grundschullehrkräfte zur Förderung der ESE ihrer Schüler*innen (Schuljahr 2018/19) sowie empirisch untersuchte Wirkungsebenen (eigene Darstellung)

- Langfristig angelegtes Format sowie prozessbezogene Begleitung und Reflexion bei der Umsetzung der Konzeption: Die in Deutschland weit verbreiteten Kurzzeitfortbildungen sind angesichts der komplexen Anforderungen im inklusiven Setting kaum zur Initiierung nachhaltiger Veränderungsprozesse geeignet (Kurniawati et al. 2014, S. 321; Lipowsky, 2014). Das vorliegende Professionalisierungskonzept und dessen Evaluation ist daher auf die Dauer eines ganzen Schuljahres angelegt. Auf diese Weise werden die Lehrkräfte von der Erweiterung ihres deklarativen und prozeduralen Wissens, über die Veränderung ihres unterrichtspraktischen Handelns bis hin zur Evaluation des Fördererfolges über einen längeren Zeitraum begleitet. Durch zeitlich gestreckte Input-Phasen und die Unterstützung bei der Durchführung des Verhaltensmonitorings kann eine stete prozessbezogene Begleitung bei der Umsetzung des Professionalisierungskonzepts gewährleistet werden. Im Rahmen der psychosozial-horizontalen Beratung besteht darüber hinaus die Möglichkeit, diesen Prozess und verändertes unterrichtspraktisches Handeln zu reflektieren. Auf diese Weise werden Input-, Erprobungs- und Reflexionsphasen miteinander verbunden, um so die erlernten Strategien in das Handlungsrepertoire der Lehrkräfte zu überführen (Lipowsky, 2014).
- Fallbezogenes Arbeiten: Die effektivsten Maßnahmen für einzelne Schüler*innen zeichnen sich durch fallbezogenes Arbeiten aus (Leidig 2019, S. 55). Wichtig dabei ist zunächst, die Lernausgangslage im verhaltensbezogenen Bereich zu erfassen – mit dem Ergebnis kann anschließend die erzieherische Zielfindung sowie die Handlungsplanung erfolgen (Hartke et al., 2016). Im vorliegenden Konzept wird die Lernausgangslage der Schüler*innen durch das *Kurzscreening schulbezogenen Verhaltens* (Blumenthal & Blumenthal, in Vorb.) erfasst. Im Rahmen der Fortbildungsmodule identifizieren die Lehrkräfte für die einzelnen Schüler*innen dann theoriegeleitet individuelle Fördermaßnahmen aus dem Pool der *84 Handlungsmöglichkeiten*. Um den Fördererfolg zu überprüfen, operationalisieren die Pädagog*innen unter wissenschaftlicher Anleitung jeweils drei individuelle Verhaltensbereiche der Schüler*innen, die sie anhand von DBR kontinuierlich einschätzen. Auf diese Weise erfolgt die Auswahl der evidenzbasierten Maßnahmen auf den konkreten Einzelfall und die jeweilige schulische Situation bezogen. Rezeptartige Vorgehensweisen werden vermieden (Hillenbrand 2015, S. 175). Stattdessen soll für jedes Kind die beste Maßnahme identifiziert und der Fördererfolg zusätzlich durch Verhaltensverlaufsdiagnostik überprüft werden (evidenzbasierte Praxis, vgl. Kap. 2.1.3.3).
- Verhaltensmonitoring: Durch den Einsatz einer Verhaltensverlaufsdiagnostik können die Implementationsqualität erhöht, erlernte Strategien langfristig in das Handlungsrepertoire der Lehrkräfte überführt sowie das Selbstwirksamkeitserleben gestärkt werden (Leidig et al., 2016). Die

Durchführung der Professionalisierungsmaßnahmen sieht vor, dass die Lehrkräfte das Verhalten der Schüler*innen vor und während der Durchführung der Fördermaßnahme(n)/Intervention(en) dokumentieren und somit die Wirksamkeit des eigenen Handelns überprüfen. Die Rückmeldung über die Ergebnisse der datengestützten Erhebungen erfolgt nach Beendigung der Intervention.

- Einbeziehung der Schulleitung: Zwischen der Ausprägung der inklusiven Einstellungen der Schulleitung und der der Lehrkräfte besteht ein bedeutender Zusammenhang. Wirksame Lehrkraftfortbildungen sind daher auf die konsequente Unterstützung durch die Schulleitung angewiesen (Urton et al., 2014). Im vorliegenden Professionalisierungskonzept wird die professionelle Weiterentwicklung und Fortbildung der gesamten Schule angestrebt. Neben der aktiven Einbindung der Schulleitungen soll das ganze Schulkollegium samt pädagogischer Fachkräfte von den Fortbildungsmaßnahmen profitieren. Im Kontrast zur Qualifikation Einzelner sollen auf diese Weise die Implementationsqualität des neuen professionellen Knowhows erhöht, die Kollaboration im Kollegium zur gemeinschaftlichen Umsetzung von Vorhaben initiiert und dadurch die nachhaltige Veränderung der Unterrichtskultur angestoßen werden.
- Untersuchung der (längerfristigen) Wirkung der Fortbildungen und Interventionen: Die vermehrte Anzahl an Fortbildungstätigkeiten im inklusiven Feld geht nicht mit einem Anstieg von Evaluationstätigkeiten einher (Amrhein & Badstieber 2013, S. 15). Dies ist problematisch, da so nicht gewährleistet werden kann, dass Ressourcen für Fortbildungen effektiv und nachhaltig eingesetzt werden. Es gilt insbesondere, die nachhaltige Wirkung von Professionalisierungsmaßnahmen und eingesetzten Interventionen in den Blick zu nehmen. Das vorliegende Professionalisierungskonzept wird über ein Schuljahr wissenschaftlich begleitet, um auch noch vier Monate nach Beendigung der Professionalisierungsmaßnahmen deren Wirkung sowohl auf der Ebene der Lehrkräfte als auch der der Schüler*innen abbilden zu können.
- Erfassung der Outcomes auf Schüler*innenebene: Mehrere Studien haben hervorgehoben, dass die Outcomes der Schüler*innen, insbesondere im affektiv-motivationalen Bereich, bei der Wirksamkeitsforschung von Lehrkraftfortbildungen mehr in den Blick genommen werden müssen (Leidig, 2016; Lipowsky, 2010; Timperley et al., 2007). Aus diesem Grund liegt in der vorliegenden Konzeption – neben den Wirkungsebenen Lehrkraftzufriedenheit, Lehrkraftkognitionen und Lehrkrafthandeln – ein besonderes Augenmerk auf den Schüler*innenoutcomes. Auf diese Weise sollen Rückschlüsse über die Wirkungsintensität und die Reichweite des Professionalisierungskonzepts und dessen Maßnahmen möglich werden.
- Universitäre Anbindung: Durch die universitäre Anbindung des Professionalisierungskonzepts soll die wissenschaftliche Begleitung und eine

hohe Umsetzungstreue gewährleistet werden. Ferner konnte gezeigt werden, dass Lehrkräfte mit universitär getragenen Angeboten sehr zufrieden sind und dies mit einem positiven Einfluss auf deren affektiv-motivationales Befinden einhergehen kann (Avramidis et al., 2000; Leidig et al., 2016).

Neben Stärken weist die Konzeption eine Reihe von vermuteten Schwächen auf. Nachfolgend werden diese anhand von Modifikationsvorschlägen erläutert.

- Umfangreichere Inputphasen: Das deklarative und prozedurale Wissen wird maßgeblich im Rahmen zweier, insgesamt sieben Stunden umfassender, Fortbildungen erweitert. Im Vergleich zu anderen Fortbildungsmaßnahmen im Kontext Inklusion im FSP ESE, deren Inputphasen zwischen 45 Minuten und 30 Stunden umfassten (Leidig et al. 2016, S. 58f.), ist der Inputumfang im vorliegenden Professionalisierungskonzept als mittelmäßig groß einzuschätzen. Durch eine Ausweitung der Inputphasen könnten die vermittelten Inhalte vertieft und wesentlich umfassender behandelt werden.
- Engere Begleitung der Erprobungsphasen: Die Erprobungsphasen werden durch den Einsatz von Verhaltensverlaufsdiagnostik, die Durchführung psychosozial-horizontaler Beratung sowie die koordinative universitäre Unterstützung begleitet. Eine noch intensivere und methodisch aufwendigere Begleitung der Erprobungsphasen, etwa durch Unterrichtsbeobachtungen und -hospitationen oder Videographie, könnte die nachhaltige Überführung der erlernten Strategien in das Handlungsrepertoire der Lehrkräfte weiter unterstützen, das Wirksamkeitsempfinden der Lehrkräfte stärken und deren positive Einstellungen gegenüber Kindern mit dem FSP ESE fördern (Leidig, 2019; Van Mieghem et al., 2020).
- Erhöhung der Reflexionstätigkeiten: Die konsequente fortbildungsdidaktische Einbindung von Beratungs-, Reflexions-, Coaching- oder Feedbackphasen in den Prozess der Erprobung neuer Handlungsstrategien wirken sich positiv auf die Implementationsqualität und die soziale Validität aus (Leidig et al., 2016). Im vorliegenden Professionalisierungskonzept sind mindestens zwei prozessbegleitende Einzelberatungssitzungen geplant. Eine noch engmaschigere Begleitung durch Beratung und damit die Erhöhung potentieller Reflexionstätigkeiten könnten sich positiv auf die Einsatzqualität und Akzeptanz der eingeführten Maßnahmen auswirken. Ferner könnten umfassendere fortbildungsintegrierte Beratungselemente, etwa mittels des Aufbaus von Problemlöseteams, Co-Teaching- oder Supervisionsstrukturen, die Kooperation und Zusammenarbeit im Schulkollegium nachhaltig voranbringen.
- Denkbar wäre, eine fortbildungsintegrierte Kooperation oder Zusammenarbeit der Schulakteur*innen durch die Initiierung von Problemlöseteams

oder Co-Teaching-Strukturen anzustoßen. Auf diese Weise würden die im FSP ESE besonders geforderten pädagogischen Strukturen der Beratung und Kooperation (Huber, 2015b), etwa anhand der Arbeit in multiprofessionellen Teams, und damit die gegenseitige Professionalisierung gefördert und weiterentwickelt werden können.

- Engere Verzahnung von Input-, Erprobungs- und Reflexionsphasen: Neben der quantitativen Erhöhung der Anteile von Input-, Erprobungs- und Reflexionsphasen, wäre eine noch engere Verzahnung dieser Elemente förderlich für die nachhaltige Modifikation von Einstellungen, Überzeugungen und handlungspraktischen Kompetenzen der Lehrkräfte. Ein kleinschrittigeres Vorgehen bei gleichzeitigem Ausbau der Phasen würde vermutlich zu noch vielfältigeren und anspruchsvolleren Lerngelegenheiten führen, die es der Lehrkraft wiederum ermöglichen, „die Wirksamkeit eigenen Handelns und die Reichweite eigener Überzeugungen zu erkennen [...], erworbene Kompetenzen in vielfältiger Weise zu trainieren [...], anzuwenden und auf ihre Wirkungen zu überprüfen“ (Lipowsky 2014, S. 522).
- Einsatz der Verhaltensverlaufsdiagnostik als Förderinstrument: Im vorliegenden Professionalisierungskonzept ist vorgesehen, dass die Lehrkräfte das verhaltensverlaufsdiagnostische Instrument DBR anwenden, um das Verhalten einzelner Schüler*innen zu monitoren und so die Wirksamkeit ihrer pädagogischen Förderung anhand der *84 Handlungsmöglichkeiten* überprüfen zu können. Die Auswertung der erhobenen Daten erfolgt nach Abschluss der gesamten Professionalisierungsmaßnahme, um den Ablauf der pädagogischen Intervention und damit die Effektstudie nicht zu beeinflussen. Eine Nutzung der Verhaltensdiagnostik als Förderinstrument würde implizieren, dass die Lehrkräfte anhand dieser „datengestützten Rückkopplungsprozesse“ (Leidig et al. 2016, S. 66) eine unmittelbare Rückmeldung zu ihrem Fördererfolg bekommen und so ihre Förderentscheidungen im Prozess anpassen können. Mit der Nutzung von Verhaltensverlaufsdiagnostik als Förderinstrument könnten Prozesse des Erlebens der eigenen Wirksamkeit wahrscheinlich noch stärker in den Fokus gerückt werden.

4 Evaluation des entwickelten Professionalisierungskonzepts

4.1 Ziele der Evaluationsstudie und Vorbemerkungen

Ziel der Studie ist es, das entwickelte Professionalisierungskonzept zur Förderung der ESE von Grundschüler*innen durch die Erweiterung spezifischer Lehrkraftkompetenzen möglichst umfassend zu untersuchen. Ausgehend von den Ergebnissen soll eruiert werden, inwiefern sich dieses sowie die darin integrierten Professionalisierungselemente *84 Handlungsmöglichkeiten* und *psychosozial-horizontale Beratung* als geeignete Professionalisierungsmaßnahmen im Feld der inklusiven Beschulung im FSP ESE erweisen. Das Professionalisierungskonzept inklusive seiner Handlungsfelder und Professionalisierungselemente wird im Folgenden als *Treatment*, der Einsatz des Professionalisierungselements *84 Handlungsmöglichkeiten* als *Intervention* und die *psychosozial-horizontale Beratung* als *Beratung* bezeichnet.

Das Professionalisierungskonzept wird auf den vier Wirkungsebenen von Lehrkraftfortbildungen nach Lipowsky (2010) – Lehrkraftreaktionen, Lehrkraftkognitionen, Lehrkrafthandeln und Schüler*innenoutcomes (vgl. Kap. 2.2.3.1) – evaluiert. Da der Hauptuntersuchungsgegenstand der Studie die Schüler*innenoutcomes sind, soll mit dieser Evaluationsebene begonnen werden, um anschließend schrittweise die anderen Ebenen zu betrachten. Auf der Ebene der Schüler*innenoutcomes wurden die Verhaltensveränderungen der Schüler*innen sowohl während der *Intervention* mithilfe einer Verlaufsdagnostik als auch vor und nach der *Intervention* und des gesamten *Treatments* anhand von Statusdiagnostiken (Fremdbeurteilung) erfasst. Auf der Ebene des Lehrkrafthandelns fand anhand der Dokumentation der eingesetzten Handlungsmöglichkeiten während der *Intervention* ein Monitoring über verändertes unterrichtliches Lehrkrafthandeln statt. Auf Ebene der Lehrkraftkognitionen wurde zu Anfang sowie nach der *Intervention* und dem gesamten *Treatment* das Selbstwirksamkeitserleben der Lehrkräfte erhoben. Schließlich wurden auf der Ebene der Lehrkraftreaktionen nach Abschluss der Professionalisierungsmaßnahme Zufriedenheits- und Nützlichkeitseinschätzungen erfragt.

Die Evaluationsstudie wurde an sieben Schulen im Schulamtsbereich Greifswald durchgeführt und umfasste von der Prä-Erhebung über die evaluierende Begleitung der Durchführung bis hin zur Post- und Follow-up-Erhebung das gesamte Schuljahr 2018/19.

4.2 Fragestellungen und Hypothesen

4.2.1 Fragestellungen

In der vorliegenden Evaluationsstudie wird untersucht, ob sich während und nach dem *Treatment* und seinen Elementen *Intervention* und *Beratung* Merkmalsveränderungen auf den vier Evaluationsebenen Schüler*innen-outcomes, Lehrkrafthandeln, Lehrkraftkognitionen und Lehrkraftreaktionen zeigen. Außerdem wird untersucht, ob zwischen den Schüler*innenoutcomes und den unterschiedlich ausgerichteten Handlungsmöglichkeiten (*Intervention*), dem Lehrkrafthandeln und den Lehrkraftkognitionen Interaktionseffekte festgestellt werden können. Folgende Forschungsfragen sollen auf den einzelnen Evaluationsebenen untersucht werden:

*Ebene der Schüler*innenoutcomes*

Fragestellung 1a

Verringern sich die spezifischen Verhaltensauffälligkeiten der geförderten Schüler*innen während der *Intervention*?

Fragestellung 1b

Verringert sich das nichtregelkonforme Verhalten der geförderten Schüler*innen während der *Intervention*?

Fragestellung 1c

Zeigen sich unterschiedliche Interaktionseffekte zwischen den Schüler*innenoutcomes (spezifische Verhaltensauffälligkeiten und nichtregelkonformes Verhalten) und den 84 Handlungsmöglichkeiten in Abhängigkeit zu deren Erklärungs- und Handlungsansatz?

Fragestellung 1d

Reduzieren sich die Verhaltensauffälligkeiten der geförderten Schüler*innen nach dem *Treatment* und ist der Effekt vier Monate später noch konstant?

Fragestellung 1e

Verbessert sich das prosoziale Verhalten der geförderten Schüler*innen nach dem *Treatment* und ist der Effekt vier Monate später noch konstant?

Ebene des Lehrkrafthandelns

Fragestellung 2a

Setzen die Lehrkräfte die *84 Handlungsmöglichkeiten* angemessen oft (> 40 % der Ratingtage) ein?

Fragestellung 2b

Zeigt sich ein positiver Interaktionseffekt zwischen der Einsatzhäufigkeit der Handlungsmöglichkeiten und den Schüler*innenoutcomes?

Ebene der Lehrkraftkognitionen

Fragestellung 3a

Verbessert sich das Selbstwirksamkeitsempfinden der Lehrkräfte nach dem *Treatment* und ist der Effekt vier Monate später noch konstant?

Fragestellung 3b

Empfinden sich Lehrkräfte nach Erhalt einer *Beratung* selbstwirksamer als Lehrkräfte, die keine *Beratung* erhalten haben und ist der Effekt vier Monate später noch konstant?

Fragestellung 3c

Zeigt sich ein positiver Interaktionseffekt zwischen dem Selbstwirksamkeitsempfinden der Lehrkräfte und den Schüler*innenoutcomes?

Ebene der Lehrkraftreaktionen

Fragestellung 4a

Wie zufrieden sind die Lehrkräfte mit der gesamten Professionalisierungsmaßnahme und wie nützlich empfinden sie diese für ihre tägliche Arbeit?

Fragestellung 4b

Wie zufrieden sind die Lehrkräfte mit den Professionalisierungselementen *84 Handlungsmöglichkeiten*, *Beratung* und *Direct Behavior Rating* und wie nützlich empfinden sie diese für ihre tägliche Arbeit?

4.2.2 Hypothesen

Folgende gerichtete Hypothesen sollen auf den einzelnen Evaluationsebenen geprüft werden:

*Ebene der Schüler*innenoutcomes*

Hypothese 1a

Während der *Intervention* (inter) verringern sich die spezifischen Verhaltensauffälligkeiten (spezV) der Schüler*innen signifikant im Vergleich zur Grundrate (grund). Die Verringerung zeigt sich durch einen sprunghaften oder kontinuierlichen Abfall des Verhaltensverlaufes.

$$H_0: \text{spezV}_{\text{grund}} = \text{spezV}_{\text{inter}} \text{ und } H_1: \text{spezV}_{\text{grund}} > \text{spezV}_{\text{inter}}$$

Hypothese 1b

Während der *Intervention* verringert sich das nichtregelkonforme Verhalten (nrV). Die Verringerung zeigt sich durch einen sprunghaften oder kontinuierlichen Abfall des Verhaltensverlaufes.

$$H_0: \text{nrV}_{\text{grund}} = \text{nrV}_{\text{inter}} \text{ und } H_1: \text{nrV}_{\text{grund}} > \text{nrV}_{\text{inter}}$$

Hypothese 1c

Es zeigen sich unterschiedliche Interaktionseffekte zwischen den Schüler*innenoutcomes (spezV und nrV) und den 84 Handlungsmöglichkeiten in Abhängigkeit zu deren Erklärungs- und Handlungsansatz:

Bei lerntheoretisch ausgerichteter Förderung zeigt sich vor allem ein Level-Effekt (level), d.h. die auffälligen Verhaltenssymptome und das nichtregelkonforme Verhalten verringern sich sprunghaft nach dem Einsetzen der *Intervention*.

$$H_0: \text{spezV}_{\text{grund-level}} = \text{spezV}_{\text{inter-level}} \text{ und } H_1: \text{spezV}_{\text{grund-level}} > \text{spezV}_{\text{inter-level}}$$

$$H_0: \text{nrV}_{\text{grund-level}} = \text{nrV}_{\text{inter-level}} \text{ und } H_1: \text{nrV}_{\text{grund-level}} > \text{nrV}_{\text{inter-level}}$$

Bei einer kombinierten Förderung (lerntheoretisch und kognitionspsychologisch/schüler*innenzentriert) zeigt sich vor allem ein Slope-Effekt (slope), d.h. die spezifischen Verhaltensauffälligkeiten und das nichtregelkonforme Verhalten verringern sich kontinuierlich nach dem Einsetzen der *Intervention*.

$$H_0: \text{spezV}_{\text{grund-slope}} = \text{spezV}_{\text{inter-slope}} \text{ und } H_1: \text{spezV}_{\text{grund-slope}} > \text{spezV}_{\text{inter-slope}}$$

$$H_0: \text{nrV}_{\text{grund-slope}} = \text{nrV}_{\text{inter-slope}} \text{ und } H_1: \text{nrV}_{\text{grund-slope}} > \text{nrV}_{\text{inter-slope}}$$

Hypothese 1d

Im Vergleich zur Prätestmessung (prä) reduzieren sich die Verhaltensauffälligkeiten (V) nach dem *Treatment* (post). Der positive Effekt ist auch nach vier Monaten (follow-up) noch feststellbar.

$$H_0: V_{\text{prä}} = V_{\text{post}} = V_{\text{follow-up}} \quad \text{und} \quad H_1: V_{\text{prä}} > V_{\text{post}} \geq V_{\text{follow-up}}$$

Hypothese 1e

Das prosoziale Verhalten (proV) verbessert sich nach dem *Treatment* und der Effekt ist vier Monate später noch feststellbar.

$$H_0: \text{proV}_{\text{prä}} = \text{proV}_{\text{post}} = \text{proV}_{\text{follow-up}} \quad \text{und} \quad H_1: \text{proV}_{\text{prä}} < \text{proV}_{\text{post}} \leq \text{proV}_{\text{follow-up}}$$

Ebene des Lehrkrafthandelns

Hypothese 2a

Die Lehrkräfte setzen die *Intervention* im Mittel wenigstens an 40 % der Ratingtage ein.

Hypothese 2b

Es zeigt sich ein positiver Interaktionseffekt zwischen der Einsatzhäufigkeit der 84 *Handlungsmöglichkeiten* und den Schüler*innenoutcomes: Werden die Handlungsmöglichkeiten wenigstens an 40 % der Ratingtage (öfter) eingesetzt, so wirkt sich dies positiv auf die Schüler*innenoutcomes (spezV und nrV) aus. Finden die Handlungsmöglichkeiten an weniger als 40 % der Ratingtage (seltener) Anwendung, wird keine positive Wirkung auf das Schüler*innenverhalten vermutet.

$$H_0: \text{spezV}_{\text{öfter}} = \text{spezV}_{\text{seltener}} \quad \text{und} \quad H_1: \text{spezV}_{\text{öfter}} < \text{spezV}_{\text{seltener}}$$

$$H_0: \text{nrV}_{\text{öfter}} = \text{nrV}_{\text{seltener}} \quad \text{und} \quad H_1: \text{nrV}_{\text{öfter}} < \text{nrV}_{\text{seltener}}$$

Ebene der Lehrkraftkognitionen

Hypothese 3a

Das Selbstwirksamkeitsempfinden (SW) der Lehrkräfte verbessert sich nach dem *Treatment*. Der positive Effekt ist auch nach vier Monaten noch feststellbar.

$$H_0: \text{SW}_{\text{prä}} = \text{SW}_{\text{post}} = \text{SW}_{\text{follow-up}} \quad \text{und} \quad H_1: \text{SW}_{\text{prä}} < \text{SW}_{\text{post}} \leq \text{SW}_{\text{follow-up}}$$

Hypothese 3b

Die Lehrkräfte, die eine kooperativ-systemische Fallberatung erhalten (mit B), zeigen unmittelbar nach dem *Treatment* eine höhere Selbstwirksamkeit

als die Lehrkräfte, die keine *Beratung* erhalten haben (ohne B). Der positive Effekt ist auch nach vier Monaten noch feststellbar.

$$H_0: SW_{\text{mit B}} = SW_{\text{ohne B}} \text{ und } H_1: SW_{\text{mit B}} > SW_{\text{ohne B}}$$

Hypothese 3c

Es zeigt sich ein positiver Interaktionseffekt zwischen dem Selbstwirksamkeitsempfinden der Lehrkräfte und den Schüler*innenoutcomes. Eine hohe Selbstwirksamkeit (SWhoch) geht mit der Verbesserung der Schüler*innenoutcomes (spezV/nrV) einher und vice versa, wohingegen sich bei vergleichsweise niedrigem Selbstwirksamkeitsempfinden (SWniedrig) die Schüler*innenoutcomes nicht bzw. kaum verbessert haben.

$$H_0: \text{spezV}_{\text{SWhoch}} = \text{spezV}_{\text{SWniedrig}} \quad \text{und} \quad H_1: \text{spezV}_{\text{SWniedrig}} > \text{spezV}_{\text{SWhoch}}$$

$$H_0: \text{nrV}_{\text{SWhoch}} = \text{nrV}_{\text{SWniedrig}} \quad \text{und} \quad H_1: \text{nrV}_{\text{SWniedrig}} > \text{nrV}_{\text{SWhoch}}$$

Ebene der Lehrkraftreaktionen

Hypothese 4a

Die Lehrkräfte sind mehrheitlich zufrieden mit der gesamten Professionalisierungsmaßnahme und erachten diese als nützlich für ihre tägliche Arbeit.

Hypothese 4b

Die Lehrkräfte sind mit der Durchführung der Professionalisierungselemente *84 Handlungsmöglichkeiten*, *Beratung* und *Direct Behavior Rating* mehrheitlich zufrieden und erachten diese als nützlich für ihre tägliche Arbeit.

4.3 Methode

4.3.1 Studiendesign und Untersuchungsplan

Bei der vorliegenden Primärstudie – die Daten wurden selbst erhoben und erstmalig ausgewertet (vgl. Bortz & Döring, 2016, S. 191) – handelt es sich um eine explanative, quasi-experimentelle Feldstudie im Ein-Gruppen-Design mit Messwiederholungen (Prätest-, Posttest- und Follow-up-Test-Messungen), bei der sowohl kontrollierte Einzelfalluntersuchungen, standardisierte Verfahren, als auch tentative, vertiefende Analysen zur Erhebung von Schüler*innen- und Lehrkraftdaten eingesetzt wurden. Somit handelt es sich um eine Studie im Mixed-Method-Design. Im Folgenden sollen die methodischen Grundlagen des Studiendesigns genauer erläutert werden.

Explanative Studien überprüfen, anders als etwa explorative oder deskriptive Studien, vorher aufgestellte und aus Theorien abgeleitete Hypothesen auf ihre Gültigkeit. Hierbei sollen insbesondere Ursache-Wirkungs-Relationen bestätigt/verworfen bzw. diese grundlegend diskutiert werden können (Bortz & Döring 2016, S. 192). In der vorliegenden Studie werden die aus den dargestellten Theorien (vgl. Kap. 2) abgeleiteten, gerichteten Hypothesen über Merkmalsänderungen auf den vier Wirkungsebenen (s. Abschnitt 4.2.2) demnach einseitig geprüft.

Bei einem quasi-experimentellen Studiendesign findet, anders als bei einem experimentellen Design, keine Randomisierung statt. Das bedeutet, dass die Versuchspersonen den Untersuchungsgruppen nicht zufällig, sondern aufgrund einer Eigenschaft zugeordnet werden. Quasi-experimentelle Designs werden eingesetzt, wenn eine zufällige Zuweisung von Personen in die Untersuchungsgruppe nicht möglich ist. Die Evaluation eines Treatments findet stattdessen unter Feldbedingungen statt, die nur eingeschränkt oder gar nicht manipulier- bzw. kontrollierbar sind (Röhrig et al., 2009). In der vorliegenden Studie bildeten sich die Untersuchungsgruppen in Abhängigkeit davon, ob sich die jeweiligen Lehrkräfte im Zuge der Studienakquise dazu bereit erklärten, an den Professionalisierungsmaßnahmen und der begleitenden Studie teilzunehmen. Die Schüler*innen wurden der Untersuchungsgruppe dann anhand der Eigenschaft kritischer Werte im *Kurzscreening schulbezogenen Verhaltens* (Blumenthal & Blumenthal, in Vorb.) zugeordnet.

Eine Feldstudie findet, im Kontrast zur Laborstudie, im natürlichen Umfeld statt, so dass die Untersuchungsbedingungen den Alltagsbedingungen ähneln und auf diese möglichst gut übertragbar sind. Feldstudien sind zwar in der Regel durch multiple Effekte erschwert, sodass prägnante Aussagen oft kaum möglich sind. Dennoch liegt es für viele Anwendungsfelder nahe, die Probleme dort zu erforschen, wo sie entstehen, d.h. unter Alltagsbedingungen in Beruf und Freizeit (Röhrig et al., 2009). Bei der vorliegenden Studie wurden die Merkmale an der Schule, also dort, wo sie tagtäglich auftreten, erhoben.

Da die vorliegende Studie ein Ein-Gruppen-Prätest-Posttest-Follow-up-Test-Design hat, stellt das Treatment (T) keine unabhängige Variable dar, da sie nicht variiert wird und damit kein Vergleich mit einem Alternativtreatment oder einer Kontrollgruppe stattfindet. Stattdessen fungiert der Messzeitpunkt (MZP), also die Vorher- (Prä-) versus Nachher- (Post-/Follow-up-) Messung, als unabhängige Variable. Nach einer Prätestmessung (MZP 1) erfolgt das Treatment und danach eine erneute Messung, die Posttestmessung (MZP 2). Ein Vergleich dieser beiden Messungen liefert Hinweise über mögliche zwischenzeitlich eingetretene Veränderungen (Bortz & Göring, 2016, 202). Eine dritte Messung (MZP 3) kann mittel- oder langfristige Wirkungen erfassen. Somit ergibt sich folgende Messreihung:

MZP 1 → T → MZP 2 → MZP 3

Die vorliegende Studie ist als Gruppenstudie konzipiert. Gruppenstudien haben den Vorteil, dass sie mehrere Fälle zusammenfassend betrachten und somit von den Eigenheiten eines Einzelfalles abstrahieren. „Auf diese Weise werden generalisierbare Erkenntnisse über die Grundgesamtheit bzw. das Untersuchungsfeld gewonnen“ (Bortz & Göring 2016, S. 215). Im Rahmen der Gruppenuntersuchung wurden Screenings, voll- und teilstandardisierte Fragebögen sowie eine kontrollierte Einzelfallstudie durchgeführt.

Screenings werden eingesetzt, um eine erste Aussage darüber zu treffen, ob ein Kind in einem bestimmten Bereich von der Altersnorm abweichende Leistungen/Merkmale hat oder nicht. Die Kinder, deren Ergebnisse im Screening eine weiterführende Diagnostik als sinnvoll erscheinen lassen, müssen nachgehend differenzierter diagnostiziert werden und „bei entsprechendem Befund den erforderlichen Förder- und Therapiemaßnahmen zugewiesen werden“ (Kany & Schöler 2009, S. 103). In der vorliegenden Studie wurde noch vor MZP 1 ein Screening eingesetzt, um Schüler*innen zu identifizieren, die Verhaltensauffälligkeiten in der Schule zeigten und damit eine intensivere Förderung benötigten (Blumenthal et al. 2020, S. 39). Aus diesem Proband*innenpool wurde die Stichprobe gezogen, die anhand eines standardisierten Verfahrens im Nachgang differenzierter untersucht wurde.

Vollstandardisierte oder quantitative Fragebögen bestehen aus geschlossenen Fragen (d.h. Fragen oder Items mit Antwortvorgaben), die in der Regel anzukreuzen oder mit Zahlenwerten zu beantworten sind und aus denen quantitative bzw. numerische Daten (Messwerte) resultieren, die statistischen Methoden der Datenanalyse unterzogen werden können. Bei vollstandardisierten quantitativen Fragebögen sind die Antworten der Befragten vergleichbar und erlauben bei ausreichender Stichprobengröße auf Basis statistischer Verfahren einen Schluss auf die Gesamtpopulation (Bortz & Döring 2016, S. 32ff.). In der vorliegenden Studie wurden vollstandardisierte Fragebögen eingesetzt, um Veränderungen auf der Ebene der Schüler*innenauscomes und der Lehrkraftkognitionen über die Zeit (Prä-, Post- und Follow-up-Messungen) zu untersuchen.

In teilstandardisierten oder qualitativen Fragebögen wird eine Liste von offenen Fragen schriftlich vorgelegt, bei der keinerlei Antwortvorgaben gemacht werden, damit die Befragten ihre Antworten in eigenen Worten formulieren und sich frei zu Themen äußern können, etwa wenn es um das Einholen von Feedback, Kritik oder Anregungen geht. Grundlage für die qualitative Fragebogengenerierung, d. h. die Auswahl und Formulierung der offenen Fragen und die Zusammenstellung des Fragebogens (Fragenanzahl, Fragenreihenfolge, Layout), können das eigene Forschungsanliegen, relevante Theorien, frühere Studien zum Themenfeld sowie eigene Vorstudien sein (Bortz & Döring 2016, S. 32ff.). Teilstandardisierte Fragebögen wurden in der vorliegenden Studie eingesetzt, um die Ebene der Lehrkraftkognitionen abzubilden.

Dabei wurden die Lehrkräfte zum Ende des Professionalisierungskonzepts zu Zufriedenheits- und Nützlichkeitsaspekten befragt.

Mit kontrollierten Einzelfalluntersuchungen untersucht man traditionell Ursache-Wirkungs-Relationen bei jeweils einer einzelnen Untersuchungsperson bzw. einer einzelnen Untersuchungseinheit (z.B. eine Schulklassse), indem man eine abhängige Variable wiederholt unter verschiedenen Manipulationen mindestens einer unabhängigen Variable misst (Barlow et al., 2009; Kazdin, 2011). Die Untersuchungsperson oder -einheit fungiert gleichzeitig als ihre eigene Kontrollgruppe bzw. Kontrollperson (Barker et al., 2015). Die Verhaltensstichproben werden in verschiedenen Situationen, zu unterschiedlichen Zeitpunkten und/oder unter variierenden Aufgabenstellungen gezogen. Kontrollierte Einzelfallstudien können neben der Überprüfung der Wirksamkeit von Maßnahmen in Einzelfällen auch als Evidenzbeleg für die Wirksamkeit einer Intervention über den Einzelfall hinaus und zur differentiellen Analyse von Interventionseffekten eingesetzt werden (Jain & Spieß 2012, S. 243). In der vorliegenden Studie wurde die kontrollierte Einzelfallstudie genutzt, um Veränderungen auf der Ebene der Schüler*innenoutcomes als eine Untersuchungseinheit abzubilden. Ziel war es zum einen, die Wirksamkeit der *Intervention* über den Einzelfall hinaus zu überprüfen und zum anderen die Interventionseffekte genauer analysieren zu können. Mithilfe eines AB-Versuchsplans wurde die Ausprägung der abhängigen Variablen (Schüler*innenverhalten) vor dem Einsetzen der Intervention – der sogenannten A-Phase (Baseline/Grundrate) – und unter Interventionsbedingungen – der B-Phase (Interventionsrate) – erfasst und dokumentiert. AB-Versuchspläne haben den Nachteil, dass ihre interne Validität eingeschränkt ist, denn auch wenn sich eine deutliche Veränderung der Daten in der Interventionsrate im Vergleich zur Grundrate zeigt, kann nicht ausgeschlossen werden, dass diese Veränderungen möglicherweise auf Störvariablen zurückzuführen sind. Um die interne Validität in der vorliegenden Studie zu erhöhen, wurde der AB-Plan um ein zweifachgestuftes Multiple-Baseline-Design (Multipler Grundratenversuchsplan) über Personen ergänzt. Bei einem Multiple-Baseline-Design über Personen wird die gleiche Intervention bei unterschiedlichen Proband*innen zeitversetzt begonnen. Durch den versetzten Interventionsbeginn sollen Störvariablen wie zwischenzeitliches Geschehen und Reifung als mögliche Alternativerklärungen ausgeschlossen und somit Rückschlüsse auf die Wahrscheinlichkeit gezogen werden, ob das gemessene Merkmal für eine größere Population verallgemeinert werden kann. Je mehr Baselines oder Grundraten in einer Studie vorhanden sind, desto höher ist die interne Validität, wobei grundsätzlich zwei Baselines genügen, um von einem Multiple Baseline Design zu sprechen und um die interne Validität eines AB-Plans deutlich erhöhen zu können (Jain & Spieß 2012, S. 227).

Evaluations-ebene	abhängige Variablen	MZP 1	Treatment-phase	MZP 2	MZP 3
Schüler*innen-outcomes	spezifische Verhaltensauffälligkeiten		Kontrollierte Einzelfallstudien im AB-Multiple Baseline Design über Personen		
	nichtregelkonformes Verhalten				
	Verhaltensauffälligkeiten	SDQ		SDQ	SDQ
	prosoziales Verhalten				
Lehrkraft handeln	Einsatzhäufigkeit der HM		Monitoring über Einsatzhäufigkeit der Interventionen		
Lehrkraft kognitionen	Selbstwirksamkeitsempfinden	Wirk-Lehr		Wirk-Lehr	Wirk-Lehr
Lehrkraftreaktionen	Zufriedenheits- und Nützlichkeitseinschätzung			teilstandardisierte Fragebögen	

Abbildung 14: Studiendesign und Untersuchungsplan zur evaluativen Begleitung der Professionalisierungsmaßnahmen (Treatment) für Grundschullehrkräfte zur Förderung der ESE ihrer Schüler*innen (eigene Darstellung)

Anmerkungen. Screening = Kurzscreening schulbezogenen Verhaltens (Blumenthal & Blumenthal, in Vorb.), HM = 84 Handlungsmöglichkeiten (Hartke et al., 2018), SDQ = Strengths and Difficulties Questionnaire – Lehrkraftversion (Goodman, 2005), WirkLehr = Skala Lehrer-Selbstwirksamkeitserwartung (Schwarzer & Schmitz, 1999).

Zusammenfassend lässt sich festhalten, dass im Rahmen der Evaluationsstudie zum Professionalisierungskonzept zu drei MZP sowohl Schüler*innen- als auch Lehrkraftdaten anhand standardisierter Erhebungsinstrumente gewonnen wurden. Außerdem wurden anhand einer kontrollierten Einzelfallstudie Verhaltensverlaufsdaten gesammelt sowie teilstandardisierte Fragebögen zur Einholung von Feedback genutzt. Auf allen Evaluationsebenen wurden Gruppenuntersuchungen durchgeführt. Das Studiendesign und der Untersuchungsplan werden in Abbildung 14 visualisiert.

4.3.2 Konkrete Durchführung der Evaluationsstudie

Die begleitende Evaluationsstudie zum Professionalisierungskonzept begann im August 2018 mit der Akquise von Schulen und Lehrkräften und endete im Juni 2019 mit der Follow-up-Erhebung. Somit dauerte die Studie ein ganzes Schuljahr. Die Rekrutierung erfolgte zunächst über postalische Anschreiben an 20 Grundschulen im Schulamtsbereich Greifswald im Zeitraum Juni/Juli 2018, nachdem das Ministerium für Bildung, Wissenschaft und Kultur des Landes Mecklenburg-Vorpommern die Durchführung der Studie schriftlich genehmigt hatte. Nach positiver Rückmeldung durch die Schulleitungen wurde im August 2018/19, zu Beginn des neuen Schuljahres, das Professionalisierungskonzept in zehn Schulen vorgestellt und die freiwillige Teilnahme angeboten. Durch diese Verfahrensweise konnten insgesamt sieben Grundschulen und 56 Lehrkräfte akquiriert werden, welche die Ausgangsstichprobe der Lehrkräfte bildeten. Anschließend wurde das Einverständnis der Erziehungsberechtigten der Schüler*innen zur Teilnahme an der Untersuchung eingeholt.

Zur Identifizierung der Schüler*innen mit Verhaltensauffälligkeiten wurden im September 2018 alle Klassen der Lehrkräfte, die sich für die Studie angemeldet hatten, mithilfe des *Kurzscreening schulbezogenen Verhaltens* (Blumenthal & Blumenthal, in Vorb.) getestet. Die Beurteilungen nahmen die Lehrkräfte selbst vor, da sie als Klassenleitung die Kinder am besten kannten und einschätzen konnten. Voraussetzung für die Testung war, dass das von den Erziehungsberechtigten im Vorfeld eingeholte Einverständnis vorlag. Per Screening wurden auf diese Weise Daten von 841 Grundschüler*innen der Klassenstufen eins bis vier erhoben. Auf Grundlage der Screening-Ergebnisse – das Auswahlkriterium waren Ergebnisse über dem kritischen Cut-Off-Wert¹³ – wählten die Lehrkräfte ein bis zwei Schüler*innen ihrer Klasse für die Studie aus, die die Stichprobe der Schüler*innen bildete. Im Rahmen des Prätests im September 2018 wurden die ausgewählten Schüler*innen mit dem *Strengths and Difficulties Questionnaire* (SDQ; Goodman, 2005) von den Lehrkräften spezifischer eingeschätzt. Außerdem beurteilten die Lehrkräfte ihre eigene Selbstwirksamkeit mit der *Skala Lehrer-Selbstwirksamkeitserwartung* (WirkLehr; Schwarzer & Schmitz, 1999).

Nach Erhebung des Prätests (MZP 1) erfolgte im September 2018 die erste Fortbildungsveranstaltung (s. Kap. 3.2.1.1 und 3.2.1.2), die zum ersten Handlungsfeld des Professionalisierungskonzepts gehörte und bei der die Lehrkräfte u.a. für den Umgang mit dem verlaufsdiagnostischen Messverfahren *Direct Behavior Rating* (Christ et al., 2009; s. Kap. 2.1.3.4 und 3.2.1.2) geschult wurden. Im Anschluss daran begannen die Lehrkräfte im Oktober

13 Der Cut-Off-Wert markiert den Übergang von unauffälligem zu auffälligem Verhalten (vgl. S. 125).

2018 damit, die Grundraten der Verhaltensverlaufsdiagnostik (kontrollierte Einzelfallstudie) der Schüler*innen zu erheben. Nach der zweiten Fortbildungsveranstaltung (s. Kap. 3.2.1.3) im November 2018 setzte die Interventionsphase ein, während der die Lehrkräfte die in der zurückliegenden Fortbildungsveranstaltung gewählten Handlungsmöglichkeiten in ihren Unterricht integrierten. Zeitgleich führten sie weiterhin nahezu täglich Verhaltensbeurteilungen durch, die nun die Interventionsrate bildeten. Die Interventionsphase war für 9-10 Schulwochen geplant und dauerte bis zu den Winterferien Ende Januar 2019. Da die kontrollierte Einzelfallstudie im zweifachgestuften Multiple-Baseline-Design über Personen durchgeführt wurde, begann die Interventionsrate bei etwa der Hälfte der Lehrkräfte eine Woche zeitversetzt. Somit ergeben sich für Grund- (17 bis 22 Schultage) und Interventionsraten (45 bis 50 Schultage) unterschiedliche Längen. Insgesamt dauerte die Erhebung der Verhaltensverlaufsdiagnostik 67 Schultage. Die Daten wurden mit einer eigens dafür programmierten App bzw. anhand von Paper-Pencil-Wochenbögen dokumentiert. Die App konnten die Lehrkräfte auf ihrem eigenen Handy oder einem Leihgerät nutzen, welches von der Universität zur Verfügung gestellt wurde. Der Vorteil der Nutzung der App bestand darin, dass die Daten bei Internetempfang sofort auf einen gesicherten Server der Universität übertragen werden konnten. Außerdem machte eine programmierte Erinnerungsfunktion die Lehrkräfte darauf aufmerksam, dass sie spätestens am Ende des Schultages ihr Rating eingaben. Einige Lehrkräfte wollten jedoch kein Smartphone nutzen und bekamen deshalb vorbereitete Paper-Pencil-Wochenbögen, die sie am Ende jeder Woche im Sekretariat hinterlegten, damit sie dort abholt werden konnten.

Nach den ersten drei Wochen der Erhebung der Interventionsrate hatten 13 nach dem Zufallsprinzip ausgewählte Lehrkräfte die Möglichkeit eine *psychozial-horizontale Beratung* (s. Kap. 3.2.3) wahrzunehmen, die von einer ausgebildeten systemischen Therapeutin geleitet wurde. Ursprünglich war geplant, wenigstens zwei Beratungsgespräche pro Person durchzuführen. Dies war jedoch aus Zeitgründen nicht möglich, da sich die eingeplante 6-wöchige Schulphase wegen Zeugniskonferenzen und Weihnachtsferien deutlich verkürzte. Die Gespräche fanden jeweils in den Räumlichkeiten der betreffenden Schule statt. Die Dauer der Gespräche betrug zwischen 45 und 90 Minuten, je nach der zur Verfügung stehenden Zeit der Lehrkraft sowie der Qualität der thematisierten Herausforderung.

Der Posttest (MZP 2) fand am ersten Tag nach den Winterferien im Februar 2019 statt. Hierbei schätzten die Lehrkräfte die Schüler*innen erneut mit dem SDQ (Goodman, 2005) ein und beurteilten die eigene Selbstwirksamkeit (WirkLehr; Schwarzer & Schmitz, 1999). Die Fragebögen waren vorher auf dem Postweg in die Schulen geschickt worden.

Tabelle 10: Durchführung der Evaluationsstudie anhand des inhaltlichen Ablaufs des Professionalisierungskonzepts, der Professionalisierungselemente (kursiv), der evaluativen Begleitung (hervorgehoben) und der eingesetzten Messverfahren

Zeitstruktur		inhaltlicher Ablauf Professionalisierungskonzept	evaluative Begleitung	eingesetzte Verfahren
August	2018	Lehrkraftakquise (Bestimmung der Proband*innengruppe der Lehrkräfte)		
September			Verhaltensscreening Schüler*innen (Bestimmung der Proband*innengruppe der Schüler*innen)	Kurzscreening
		<i>Fortbildung I</i>	Prätest	SDQ WirkLehr
Oktober			Grundratenerhebung (im zweifachgestuften MBD)	DBR
November		<i>Fortbildung II Einsatz der HM als Intervention kooperativ-systemische Beratungen</i>	Interventionsratenerhebung (im zweifachgestuften MBD)	DBR
Dezember				
Januar	2019			
Februar			Posttest	SDQ WirkLehr
März		Abschlussveranstaltung	Abschlussevaluation	Evaluationsfragebogen
April				
Mai				
Juni			Follow-Up-Test	SDQ WirkLehr

Anmerkungen. SDQ = Strengths and Difficulties Questionnaire – Lehrkraftversion (Goodman, 2005), WirkLehr = Skala Lehrer-Selbstwirksamkeitserwartung (Schwarzer & Schmitz, 1999), DBR = *Direct Behavior Rating* (Christ et al., 2009), MBD = Multiple-Baseline-Design.

Im März 2019, etwa einen Monat nach dem Posttest, fand die Abschlussveranstaltung statt, bei der den Lehrkräften die ersten Ergebnisse der Studie präsentiert und sie gebeten wurden, das Professionalisierungskonzept und seine Elemente anhand vorbereiteter teilstandardisierter Fragebögen zu evaluieren. Die sowohl offenen als auch geschlossenen Fragenformate wurden schriftlich und anonym beantwortet. Einige der geschlossenen Fragen wurden außerdem – für alle sichtbar – auf einem Smiley-Barometer per Punktevergabe beantwortet.

Vier Monate nach dem Posttest, im Juni 2019, wurden die Lehrkräfte noch einmal zu ihrem Selbstwirksamkeitsempfinden befragt und gebeten, den SDQ-Lehrkraftfragebogen für die Proband*innengruppe der Schüler*innen auszufüllen. Die Durchführung der Evaluationsstudie ist in Tabelle 10 dargestellt.

4.3.3 Bildung und Beschreibung der Untersuchungsgruppen

Die Probanden*innengruppen der Lehrkräfte und Schüler*innen wurden den Gesamtstichproben entnommen, die zu Beginn des Schuljahres 2018/2019 akquiriert bzw. erhoben worden waren. Da das Forschungsprojekt auf die freiwillige Teilnahme von interessierten Schulen und Lehrkräften angewiesen war, mussten diese vor Beginn der Studie angeworben werden. Wie in Kap. 4.3.2 zur Durchführung der Studie erläutert, wurden deshalb vor Beginn der Studie 20 Grundschulen im Schulamtsbereich Greifswald postalisch angeschrieben. Um eine möglichst repräsentative Stichprobe zu generieren, waren bei der Auswahl der Schulen folgende Kriterien ausschlaggebend:

1. ausgewogenes Verhältnis zwischen inner- und außerstädtischen/ländlichen Schulstandorten,
2. Berücksichtigung öffentlicher Schulen und Schulen in privater Trägerschaft,
3. Mitberücksichtigung von Schulen mit reformpädagogischem Profil.

Aus organisatorischen Gründen war außerdem eine gute Erreichbarkeit vom Forschungsstandort Universität Greifswald per Auto oder Zug innerhalb eines Radius von etwa 40 km zu gewährleisten.

Zur Aufnahme in die Untersuchungsgruppen wurden nach Abschluss des MZP 3 neben möglichst vollständigen Datensätzen auch aussagekräftige Verlaufsdagnostiken zum Verhalten der Schüler*innen als ein Kriterium zugrunde gelegt. Die Verfahren der Stichprobenziehungen und Untersuchungsgruppenbildung sollen im Folgenden erklärt sowie der im Zuge der zehnmonatigen Studie entstandene Dropout von Studienteilnehmer*innen detailliert erläutert werden. In Abbildung 15 ist die Rekrutierung und der Ablauf der vorgestellten Studie mit Stichprobenumfang visualisiert.

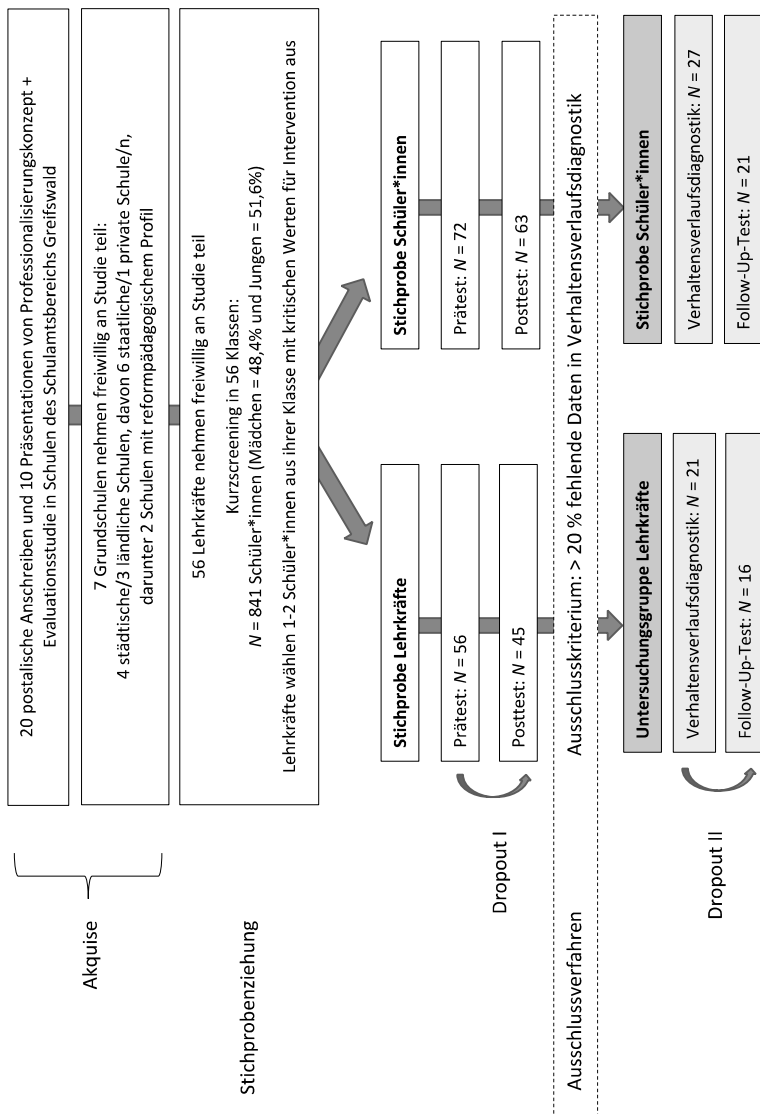


Abbildung 15: Bildung der Untersuchungsgruppen der Lehrkräfte und Schüler*innen anhand von Stichprobenziehungen, Dropouts und Ausschlussverfahren (eigene Darstellung)

Anmerkungen. Kurzscreening = Kurzscreening schulbezogenen Verhaltens (Blumenthal & Blumenthal, in Vorb.), N = Anzahl.

4.3.3.1 Lehrkräfte

Von den bei der Studienakquise kontaktierten Schulen waren zehn Schulen an einer Präsentation des Vorhabens interessiert und 56 Lehrkräfte/Klassenleiter*innen an sieben Schulen konnten schließlich für das Professionalisierungskonzept gewonnen werden. Diese bildeten die Anfangsstichprobe der Lehrkräfte, die durchweg weiblich war ($N = 56$). Die Lehrkräfte waren beim ersten MZP im Mittel 47 Jahre alt ($SD = 9.7$) und 22 Jahre im Schuldienst ($SD = 19.09$).

Zwischen MZP 1 und 2 (September 2018 – Februar 2019) kam es zu einem Dropout von 11 Lehrkräften. Zwei Lehrkräfte haben kurz vor Beginn der Professionalisierungsmaßnahmen ohne Nennung von Gründen die Evaluationsstudie wieder verlassen (Sept. 2018), zwei Lehrkräfte schieden aufgrund einer Schwangerschaft aus (Oktober 2018/Januar 2019), drei Lehrkräfte mussten wegen einer Langzeiterkrankung aussetzen (September 2018/November 2018) und vier Lehrkräften wurde die Teilnahme an dem Forschungsprojekt zu zeitaufwendig (alle November 2018).

Da die Häufigkeit der Ratings während der Erhebung der Verhaltensverlaufsdiagnostik auch Auskunft über die Umsetzungstreue der *Intervention* gibt und dies ein wesentlicher Faktor für die Interpretation der Ergebnisse der Evaluationsstudie ist, wurden bei den finalen Analysen nur diejenigen Fälle berücksichtigt, bei denen mindestens 80 % der in der Grund- und Interventionsrate möglichen Menge an Datenpunkten (min. = 54; max. = 67) vorhanden waren. Durch dieses Ausschlussverfahren konnten bei der Auswertung der Verhaltensverlaufsdiagnostik die Erhebungen von $N = 21$ Lehrkräften ($N = 27$ Schüler*innendaten) berücksichtigt werden. Diese 21 Lehrkräfte bildeten die Untersuchungsgruppe Lehrkräfte, deren Daten in die finalen Auswertungen zur Überprüfung der in der Evaluationsstudie zugrunde gelegten Hypothesen eingingen. Die Untersuchungsgruppe Lehrkräfte war im Mittel 46 Jahre alt ($SD = 9.7$) und seit 21 Jahren ($SD = 13.1$) im Schuldienst tätig.

4.3.3.2 Schüler*innen

Zur Identifikation der Schüler*innen mit Verhaltensauffälligkeiten schätzten die 56 Klassenleiter*innen, die sich für die Teilnahme an der Studie bereit erklärt hatten, in einem ersten Schritt mit dem *Kurzscreening schulbezogenen Verhaltens* die Schüler*innen ihrer Klasse ein, für die das von den Erziehungsberechtigten im Vorfeld eingeholte Einverständnis vorlag. Die so erhobene Gesamtstichprobe der Schüler*innen umfasste insgesamt 841 Schüler*innen und bestand aus 403 Mädchen (48.4 %) und 430 Jungen (51.6 %). In acht Fällen wurde kein Geschlecht angegeben. Das Geschlechterverhältnis ist damit ausgewogen. 719 Schüler*innen lernten an staatlichen Grundschulen, 168 davon an einer Schule mit reformpädagogischem Profil. Die restlichen 122 Schüler*innen der Gesamtstichprobe lernten an einer Privatschule

mit reformpädagogischem Profil. Innerhalb der Gesamtstichprobe der Schüler*innen waren 191 Erstklässler*innen (22.7 %), 227 Zweitklässler*innen (27.0 %), 250 Drittklässler*innen (29.7 %) und 173 Viertklässler*innen (20.6 %). Im Mittel waren die Schüler*innen 7.8 Jahre alt (min. = 5.4, max. = 13.2, $SD = 1.23$). Bei 103 Kindern (12.8 %) bestand zum Zeitpunkt der Erhebung eine förmlich festgestellte Teilleistungsstörung oder ein festgestellter sonderpädagogischer Förderbedarf. Bei sieben Kindern (0.83 %) war im Vorfeld der SF ESE festgestellt worden.

Anhand der Ergebnisse des Kurzscreenings – das Kriterium waren auffällige Werte in den Subskalen Regelverhalten, emotional-soziales Verhalten oder Arbeitsverhalten – wählten die Lehrkräfte ein bis zwei Schüler*innen aus ihrer Klasse für die *Intervention* aus. So kam eine Anfangsstichprobe von $N = 72$ Schüler*innen zustande ($M = 7.5$ Jahre, $SD = 1.17$), davon waren 21 Mädchen (29.2 %) und 51 Jungen (70.8 %).

Durch den Dropout von 11 Lehrkräften zwischen MZP 1 und 2 waren zum Posttest noch 63 Schüler*innendatensätze zur Auswertung vorhanden. Von ihnen blieben nach dem Ausschlussverfahren $N = 27$ übrig. Diese 27 Schüler*innen bildeten die Untersuchungsgruppe Schüler*innen, deren Daten in die finalen Auswertungen zur Überprüfung der Hypothesen eingingen.

Die Untersuchungsgruppe Schüler*innen ($N = 27$) setzt sich aus fünf Mädchen (18.5 %) und 22 Jungen (81.5 %) zusammen. Das Geschlechterverhältnis ist damit nicht ausgewogen.¹⁴ Die Schüler*innen waren im Durchschnitt $M = 7.52$ ($SD = 1.16$) Jahre alt und recht ausgewogen auf die Klassen eins bis vier verteilt: Es gab acht Erstklässler*innen (29.6 %), acht Zweitklässler*innen (29.6 %), fünf Drittklässler*innen (18.5 %) und sechs Viertklässler*innen (22.2 %). Sieben Schüler*innen (25.9 %) hatten zu Beginn der Studie bereits eine förmlich festgestellte Teilleistungsstörung (ADHS: $N = 1$) oder einen festgestellten sonderpädagogischen Förderbedarf, hier vor allem im Bereich ESE ($N = 5$, Sprache: $N = 1$).

Im Vergleich zur Gesamtstichprobe überschritt die Untersuchungsgruppe die Cut-Off-Werte¹⁵ des Kurzscreenings und befand sich durchschnittlich im kritischen Bereich. Die Mittelwerte der einzelnen Subskalen betrugen: Regelverhalten $M = 13.33$ ($SD = 3.87$), Arbeitsverhalten $M = 14.59$ ($SD = 4.37$), Emotional-soziales Verhalten $M = 11.30$ ($SD = 3.29$). Nur die Ergebnisse im Bereich Allgemeine Schuleinstellung waren unauffällig ($M = 1.93$, $SD = 1.96$). In der Tabelle 11 wird die Untersuchungsgruppe anhand ihrer Gruppenmerkmale mit der Gesamtstichprobe verglichen.

14 Dies spiegelt auch die Einschätzung des regelkonformen Verhaltens in der Gesamtstichprobe wider: Die Mädchen wurden signifikant weniger auffällig eingeschätzt als die Jungen, $F(1, 831) = 66.706$, $p < .001$.

15 Cut-Off-Werte Kurzscreening schulbezogenen Verhaltens (Blumenthal & Blumenthal, in Vorb.): Regelverhalten > 12 , Arbeitsverhalten > 12 , Emotional-soziales Verhalten > 10 , Allgemeine Schuleinstellung > 4 .

Tabelle 11: Untersuchungsgruppe ($N = 27$) im Vergleich zur Gesamtstichprobe ($N = 841$) anhand deskriptiver Daten und der Rohwerte der Subskalen des Kurzscreenings schulbezogenen Verhaltens

Gruppenmerkmal	Untersuchungsgruppe ($N = 27$)	Gesamtstichprobe ($N = 841$)
weiblich	$N = 5$ (18.5 %)	$N = 403$ (48 %)
männlich	$N = 22$ (81.5 %)	$N = 429$ (51 %)
förmlich festgestellte Teilleistungsstörung oder sonderpädagogischer Förderbedarf	$N = 7$ (25.9 %)	$N = 103$ (12 %)
Regelverhalten	$M = 13.33$ ($SD = 3.87$)	$M = 5.27$ ($SD = 4.85$)
Emotional-soziales Verhalten	$M = 11.30$ ($SD = 3.29$)	$M = 4.61$ ($SD = 3.89$)
Arbeitsverhalten	$M = 14.59$ ($SD = 4.37$)	$M = 7.23$ ($SD = 5.53$)
Allgemeine Schuleinstellung	$M = 1.93$ ($SD = 1.96$)	$M = 0.83$ ($SD = 1.16$)

Anmerkungen. N = Anzahl, M = Mittelwert, SD = Standardabweichung. *hervorgehoben: Rohwerte, die über den kritischen Cut-Off-Werten (Regelverhalten >12 , emotional-soziales Verhalten >10 , Arbeitsverhalten >12 und allgemeine Schuleinstellung >4) liegen.

4.3.4 Messverfahren

Bei der Auswahl der Erhebungsinstrumente standen neben ökonomischen vor allem inhaltliche Aspekte im Vordergrund. Standardisierte Messinstrumente mit hoher Güte und kurzer Durchführdauer (*Strenghts and Difficulties Questionnaire* – Lehrkräfte, Goodman, 2005; *Skala Lehrer-Selbstwirksamkeitserwartung*, Schwarzer & Schmitz, 1999) wurden genutzt, um Veränderungen in den Gruppen Lehrkräfte und Schüler*innen über die Zeit zu untersuchen. Die kontrollierte Einzelfallstudie anhand von Verhaltensverlaufsdagnostik (*Direct Behavior Rating*; Christ et al., 2009) wurde angewendet, um Veränderungen der Merkmale der Untersuchungseinheit Schüler*innen im Verlauf abzubilden. Um die Reaktionen, Meinungen und Einschätzungen zum Professionalisierungskonzept zu erfassen, wurden bei den Lehrkräften teilstandardisierte Befragungen eingesetzt, die eigens für diesen Zweck entworfen worden waren.

Die in Tabelle 10 dargestellten Messverfahren sollen im Folgenden differenziert erläutert werden. Die Reihenfolge der Darstellung der Messverfahren orientiert sich an den zu evaluierenden Wirkungsebenen. Das Screeningverfahren wird losgelöst von den Wirkungsebenen vorgestellt, da es einzig dazu diente, die Schüler*innenstichprobe zu generieren.

Das *Kurzscreening schulbezogenen Verhaltens – Lehrkräfte* (Blumenthal & Blumenthal, in Vorb.) wurde an der Universität Rostock entwickelt. Der Fragebogen erfasst in vier Subskalen und insgesamt 19 positiv formulierten Items (Bsp. „Hält sich an soziale Regeln“, „Beendet angefangene Aufgaben“, „Wird von anderen Schülerinnen und Schülern akzeptiert“) das schulische Verhalten von Kindern anhand von Lehrkraftbeurteilungen. Die Subskalen Regelverhalten (RV, 6 Items), Emotional-Soziales Verhalten (ESV, 5 Items), Arbeitsverhalten (AV, 6 Items) und Allgemeine Schuleinstellung (ASE, 2 Items) werden auf einer fünfstufigen Likert-Skala (0 = immer, 1 = fast immer, 2 = mal so, mal so, 3 = selten, 4 = nie) eingeschätzt. Der Test kann als Gruppen- oder Einzeluntersuchung durchgeführt werden. Die Punkte werden addiert und bilden pro Subskala einen Gesamtrohwert. Ab einem Punkt über der Hälfte der maximal zu erreichenden Punkte ist der Wert in der betreffenden Subskala als kritisch einzuschätzen. Die Cut-Off-Werte lauten für die einzelnen Subskalen wie folgt: Regelverhalten > 12, emotional-soziales Verhalten > 10, Arbeitsverhalten > 12 und allgemeine Schuleinstellung > 4. Der Test kann für Kinder im Grundschulalter angewandt werden. Bei der Testdauer gibt es keine Zeitbegrenzung, das Ausfüllen eines Fragebogens dauert nur wenige Minuten.

Das Kurzscreening schulbezogenen Verhaltens wurde als Screening zum Anfang der Studie eingesetzt, um den Proband*innenpool der Schüler*innen zu bestimmen, anhand dessen die Stichprobe gezogen werden konnte. Bei der Stichprobenziehung wurden die Schüler*innen berücksichtigt, deren Ergebnisse sich in mindestens einer der drei Subskalen Regelverhalten, Emotional-Soziales Verhalten oder Arbeitsverhalten über den Cut-Off-Werten bewegten.

4.3.4.1 Erfassung der Schüler*innenoutcomes

Zur Erfassung der Schüler*innenoutcomes wurden zwei Messverfahren eingesetzt: das verlaufsdiagnostische Verfahren *Direct Behavior Rating* (Christ et al., 2009) und der statusdiagnostische Fragebogen *Strengths and Difficulties Questionnaire* (Goodman, 2005).

Das Messinstrument DBR (vgl. Kap. 2.1.3.4) wurde eingesetzt, um die Entwicklung des nichtregelkonformen Verhaltens und der spezifischen Verhaltensauffälligkeiten als Formen schulischer Verhaltensauffälligkeiten (s. Kap. 2.1.2) abzubilden. Dazu wurden sowohl eine Single-Item-Scale („Hält sich nicht an Absprachen und allgemeine Regeln“) als auch eine Multi-Item-Scale mit drei Items genutzt. Sowohl bei SIS als auch MIS wurde eine negative Itemformulierung gewählt, da dies die Genauigkeit der Einschätzung erhöht (Mahlau & Erdmann, 2020; Huber & Rietz, 2015; Chafouleas et al., 2013).

Die Identifizierung relevanter Items für die DBR-MIS erfolgte schrittweise. Zunächst wurden anhand der Ergebnisse des *Kurzscreening schulbezoge-*

nen Verhaltens die problematischsten, größtenteils externalisierenden Verhaltensweisen (s. Kap. 2.1) eines jeden Kindes der Proband*innengruppe identifiziert. Im zweiten Schritt wurden drei dieser positiv formulierten Items (Bsp. „Hält sich an individuelle Absprachen“, „Verhält sich gegenüber der Lehrkraft respektvoll/angemessen“, „Hält sich an soziale Regeln“) durch die Lehrkräfte stärker operationalisiert. Unter Anleitung notierten die Lehrkräfte konkrete(re), beobachtbare und messbare Verhaltensweisen ihrer Schüler*innen (s. Kap. 2.1.3.4). Erschienen den Lehrkräften bei diesem Schritt andere, durch das Screening nicht abgedeckte Verhaltensweisen eher als vordergründig im Schulalltag, so konnten sie auch diese operationalisieren. Während eines dritten Schrittes wurden die niedrig-inferenten Items durch die Studienleiterin so weit wie möglich systematisiert und sprachlich vereinheitlicht. Die dadurch entstandenen MIS-Items (spezifische Verhaltensauffälligkeiten) wurden dann mit den Lehrkräften besprochen, sodass diese genau wussten, welches konkrete Verhalten sie beurteilen sollten (vgl. Conley et al., 2014). Anhand dieser Verfahrensweise wurden insgesamt 36 Items entwickelt, die sich auf die spezifischen Verhaltensauffälligkeiten eines jeden Kindes im schulischen Alltag bezogen. Die MIS-Items sind Tabelle 12 zu entnehmen.

Tabelle 12: Operationalisierte Items der Multi-Item-Scales, geordnet nach deren Häufigkeit

Nr.	Items MIS	Häufigkeit
1	redet/ruft dazwischen	10
2	stört Mitschüler*innen beim Arbeiten	6
3	ist nicht bereit für den Unterricht	6
4	beschäftigt sich mit Nicht-Unterrichtsbezogenem	5
5	hält keinen Blickkontakt	4
6	ist im Unterricht abwesend/abgelenkt	4
7	beginnt zögernd zu arbeiten	4
8	beleidigt andere Schüler*innen	3
9	hält keine Ordnung am Platz	3
10	geht nicht sorgsam mit Arbeitsmaterialien um	3
11	flucht und schimpft	3
12	reagiert unangemessen auf Ansprache/Ermahnung	3

Nr.	Items MIS	Häufigkeit
13	unterbricht Arbeit	3
14	hört anderen nicht zu	2
15	gibt schnell auf	2
16	löst Konflikte unangemessen	2
17	steht unaufgefordert im Unterricht auf	2
18	fordert sofort Hilfe ein	2
19	schlägt andere	2
20	drückt Gefühle unangemessen aus	2
21	macht unangemessene Geräusche	2
22	widerspricht der Lehrkraft	2
23	sagt Unwahrheiten	2
24	stört/lenkt andere Schüler*innen ab	2
25	zeigt keine Gesprächsbereitschaft	1
26	spricht in unangemessener Lautstärke	1
27	hilft anderen nicht von sich aus	1
28	kippelt	1
29	ist sich über das Ausmaß seiner Handlungen nicht bewusst	1
30	isst unangemessen	1
31	mischt sich in Angelegenheiten anderer Schüler*innen ein	1
32	geht mit Gefühlen unangemessen um	1
33	respektiert Privatbesitz anderer nicht	1
34	hält sich nicht an Arbeitsanweisungen	1
35	möchte nach Hause	1
36	Hausaufgabe/Material fehlt	1

Anmerkungen. MIS = Multi-Item-Scale (Christ et al., 2009).

Die Einschätzung der qualitativen Ausprägungen der SIS und MIS erfolgte einmal täglich anhand einer zehnstufigen Likert-Skala (0 = nicht vorhanden, 1-2 = schwach vorhanden, 2-3 = gelegentlich vorhanden, 5 = mal so, mal so, 6-7 = meist vorhanden, 8-9 = stark vorhanden, 10 = immer vorhanden). Die DBR wurden zu unterschiedlichen Tageszeiten erhoben, um ein möglichst breites und aussagekräftiges Verhaltensbild zu erstellen: An zwei Wochentagen wurde das Verhalten zu Anfang, an zweien zum Ende und an einem zur Mitte eines Schultages eingeschätzt.

Der *Strengths and Difficulties Questionnaire* zu Stärken und Schwächen (SDQ; Goodman, 2005) ist ein kurzer Fragebogen zur Erfassung von Verhaltensauffälligkeiten und -stärken bei Kindern und Jugendlichen. Er liegt als Eltern- und Lehrkraftversion sowie als Selbstbericht vor und wurde in der vorliegenden Studie in der deutschen Version (SDQ-D) genutzt. Die 25 Items verteilen sich auf folgende fünf Skalen: „Emotionale Probleme“, „Verhaltensprobleme“, „Hyperaktivität“, „Verhaltensprobleme mit Gleichaltrigen“ und „Prosoziales Verhalten“. Jede Skala besteht aus fünf Items (Bsp. „Streitet sich oft“, „Leicht ablenkbar“, „Rücksichtsvoll“), die auf einer dreistufigen Likert-Skala (0 = nicht zutreffend, 1 = teilweise zutreffend, 2 = eindeutig zutreffend) eingeschätzt werden. Für jede Skala und den Gesamtproblemwert werden der Score sowie eine Interpretation (normal, grenzwertig oder auffällig) angegeben. Die Verteilung der Werte zur Interpretation und Definition einer Störung wurde von den Autor*innen dabei so gewählt, dass ca. 80 % der Kinder als normal, 10 % als grenzwertig und 10 % als auffällig eingestuft werden.

Die Reliabilität der englischen Originalform ist ausführlich überprüft worden. Die mittlere interne Konsistenz variierte in einer großen britischen Bevölkerungsstudie für die einzelnen Problemskalen zwischen $\alpha = .70$ und $.88$ und lag für den Gesamtproblemwert bei $\alpha = .80$ (Goodman, 2001). Die Retest-Reliabilitäten bewegten sich nach vier bis sechs Monaten in der britischen Bevölkerungsstudie zwischen $r_{tt} = .65$ und $.82$ (ebd.). Die bisherigen Evaluationen der deutschsprachigen Lehrkraftversion weisen auf eine angemessene interne Konsistenz sowie eine akzeptable Passung des 5-Faktorenmodells, nachgewiesen durch eine Hauptkomponentenanalyse, hin (Bettge, Ravens-Sieberer, Wietzker & Hölling, 2002; Saile, 2007). Die Validität der deutschen Selbstbeurteilungsversion des Fragebogens konnte außerdem für klinisch-praktische sowie für Forschungszwecke nachgewiesen werden (Lohbeck, Schultheiß, Petermann & Petermann, 2015). Der SDQ ist ein zeitökonomisches Instrument, dessen Beantwortung nur wenige Minuten in Anspruch nimmt.

Im Rahmen der vorliegenden Studie wurde der Fragebogen von den Klassenlehrkräften ausgefüllt. Bei der Auswertung wurden die Werte aller Skalen genutzt, um herauszufinden, ob sich die Ausprägungen der Verhaltensauffälligkeiten und des prosozialen Verhaltens der Schüler*innen mit der Zeit positiv verändern.

4.3.4.2 Erfassung des unterrichtlichen Lehrkrafthandelns

Im Folgenden sollen die Verfahren erläutert werden, die zur Erfassung des unterrichtlichen Lehrkrafthandelns entwickelt und eingesetzt wurden.

Den für den unterrichtlichen Einsatz ausgewählten Handlungsmöglichkeiten liegen unterschiedliche Erklärungs- und Handlungsansätze zugrunde (s. Kapitel 2.1.3.4). Drei Ausrichtungen können unterschieden werden: Handlungsmöglichkeiten mit einem eher lerntheoretischen, einem kognitionspsychologischen und einem schüler*innenzentrierten Erklärungs- und Handlungsansatz. Theoriegeleitet entschieden sich die Lehrkräfte während der Fortbildung II (s. Kap. 3.2.1.3) für bis zu drei Handlungsmöglichkeiten zum unterrichtlichen Einsatz. Hierbei konnten Handlungsmöglichkeiten eines Erklärungs- und Handlungsansatzes oder einer Kombination aus lerntheoretischem, kognitionspsychologischem und/oder schüler*innenzentriertem Ansatz gewählt werden. Ihre Auswahl teilten sie der Studienleiterin mit, die diese anhand ihrer Ausrichtung sowie der Häufigkeit (Wie oft wurde jede Handlungsmöglichkeit ausgewählt?) systematisierte. Eine differenzierte Darstellung der eingesetzten Handlungsmöglichkeiten erfolgt in Kapitel 5.1.2.

Die Lehrkräfte dokumentierten während der Interventionsphase an jedem angesetzten Ratingtag, ob sie in der jeweiligen Unterrichtssituation die Handlungsmöglichkeit/en eingesetzt hatten oder nicht. Um die gemittelte Frequenz des Einsatzes der 84 Handlungsmöglichkeiten zu bestimmen, wurde die Häufigkeit des Einsatzes ins Verhältnis zu den getätigten Ratings gesetzt. Wenn eine Lehrkraft bspw. angegeben hat, insgesamt 30-mal die Handlungsmöglichkeiten während der *Intervention* eingesetzt zu haben und insgesamt 40 Ratings vorlagen, so ergab dies eine Einsatzhäufigkeit von 75 % (30/40 = 0.75).

4.3.4.3 Erfassung der Lehrkraftkognitionen

Die Wirkungsebenen der Lehrkraftkognitionen wurde anhand des Selbstwirksamkeitsempfindens der Lehrkräfte untersucht. Die *Skala der Lehrer-Selbstwirksamkeitserwartung* (Schwarzer & Schmitz, 1999), die die individuelle Selbstwirksamkeitserwartungen von Lehrkräften erfasst, besteht aus 10 Items (Bsp.: "Ich bin mir sicher, dass ich auch mit den problematischen Schülern in guten Kontakt kommen kann, wenn ich mich darum bemühe", "Selbst wenn mein Unterricht gestört wird, bin ich mir sicher, die notwendige Gelassenheit bewahren zu können"). Das Antwortformat ist vierstufig (1 = stimmt nicht, 2 = stimmt kaum, 3 = stimmt eher, 4 = stimmt genau). Die Items sind, bis auf eines, positiv formuliert. Das eine, negativ formulierte Item muss demnach bei der Auswertung umgepolt werden.

Die Reliabilität ist in bisher drei Stichproben (jeweils $N \sim 300$) überprüft worden. Die innere Konsistenz der Skala lag zu den verschiedenen MZP zwischen $\alpha = .76$ und $.82$. Das Instrument ist demnach homogen und misst

zuverlässig. Die Retest-Reliabilität über den Zeitraum eines Jahres betrug $rtt = .67$ ($N = 158$) und $.76$ ($N = 193$), über den Zeitraum von zwei Jahren lag diese bei $rtt = .65$ ($N = 161$). Die bundesweite Erprobung des Messinstruments ergab gute psychometrische Kennwerte. Alle untersuchten Selbstwirksamkeitsmaße ließen sich durch Idealismus, Arbeitsüberforderung, Kontrolliertheitserleben und Arbeitsunzufriedenheit über ein Jahr hinweg vorhersagen. Dabei ergaben sich auch Hinweise zur Validität. Selbstwirksame Lehrkräfte waren eher gewohnt, einen Teil ihrer Freizeit für zusätzliches Engagement zu opfern, als dies bei weniger selbstwirksamen Lehrkräften der Fall war. Die Entwicklung des Ausbrennens im Beruf (Burnout) ließ sich über Jahre hinweg aufgrund von Selbstwirksamkeitserwartungen relativ gut vorhersagen (Korrelationen zwischen Lehrer-Selbstwirksamkeit und drei Burnoutdimensionen zu den MZP 2 bis 4 zwischen $r = .41$ und $r = .79$). Außerdem gab es empirische Hinweise auf eine kausale Prädominanz von Selbstwirksamkeitserwartung gegenüber Burnout. Die Skala ist für Erwachsene konzipiert und die Beantwortung dauert nur wenige Minuten (Schwarzer & Schmitz, 1999, 2002).

Die Skala zur Lehrer-Selbstwirksamkeit wurde zu drei MZP eingesetzt, um das Selbstwirksamkeitsempfinden der Lehrkräfte über die Zeit zu untersuchen.

4.3.4.4 Erfassung der Lehrkraftreaktionen

Die Lehrkraftreaktionen wurden anhand von Zufriedenheits- und Nützlichkeits einschätzungen abgefragt und mittels teilstandardisierter, qualitativer Befragungen erhoben. Die Entwicklung des Fragenkatalogs wird im Folgenden erläutert.

Die Konzipierung des Fragenkatalogs erfolgte in Anlehnung an die Handreichung von Scheffelt und Teichmann (2018). Die Autor*innen schlagen für die konkrete Durchführung eines Evaluationsvorhabens sieben Schritte vor: 1) Evaluationsziele klären und festlegen, 2) Evaluationsgegenstand bestimmen, 3) Gegenstand konkretisieren und Indikatoren festlegen, 4) Instrumente auswählen, 5) Daten sammeln und aufbereiten, 6) Daten analysieren und interpretieren sowie 7) Rückmeldungen geben, Konsequenzen ziehen und Maßnahmen planen.

Zunächst sollten die Evaluationsziele geklärt und festgelegt werden (Schritt 1). Bei den hier zu unterscheidenden Evaluationsfunktionen (Erkenntnis-, Kontroll-, Entwicklungs-, Legitimations-, Marketing- und Dialogfunktion) spielen in der vorliegenden Studie vor allem die Erkenntnis- und Entwicklungsfunktion eine wichtige Rolle: Es soll überprüft werden, inwieweit mit der Professionalisierungsmaßnahme bestimmte Ziele erreicht wurden (Erkenntnisfunktion), welchen besonderen Mehrwert die Lehrkräfte in dem Professionalisierungskonzept sehen (Erkenntnis- und Entwicklungsfunktion) und welche Verbesserungsvorschläge sie haben (Entwicklungsfunktion).

tion). Bei der Bestimmung des Evaluationsgegenstandes (Schritt 2) unterscheiden Scheffelt und Teichmann (2018) vier Gruppen: Inputs (Bedingungen und Ressourcen), Prozesse (Verlauf der Lehr- und Lernprozesse), Outputs (konkrete Ergebnisse) und Outcomes (längerfristige Wirkungen). Da das gesamte Professionalisierungskonzept, von der Planung über die Durchführung bis hin zu den Ergebnissen und Aspekten der Nachhaltigkeit, evaluiert werden sollte, wurden die Lehrkräfte zu allen vier Evaluationsgegenständen befragt, wobei ein besonderer Fokus auf den Outputs lag. Als Indikatoren (Schritt 3) für die Erreichung des Professionalisierungsziels wurden vor allem Zufriedenheits- („Wie haben Sie sich betreut gefühlt?“), Nützlichkeits- und Praktikabilitätseinschätzungen („Wie nützlich finden Sie das Gesamtkonzept?“, „Welchen Mehrwert hatte das Konzept für Sie?“) sowie Einschätzungen über die positive Wirkung der Professionalisierungsmaßnahmen („Haben Sie durch die Anwendung der direkten Verhaltensbeurteilung positive Veränderungen bei dem*der Schüler*in/der ganzen Klasse/bei sich selbst wahrgenommen?“, „In welchem Ausmaß hatte die *Beratung* positiven Einfluss auf Ihre Emotionsregulierung/Stressregulation/Sicherheit in der Klasse“) zugrunde gelegt. Auf der Ebene der Inputs wurde die Länge der Veranstaltungen abgefragt („Die Länge der Veranstaltungen war angemessen.“). Auf der Ebene der Prozesse wurde Feedback sowohl zur Durchführung der Fortbildungen („Als wie hilfreich schätzen Sie die Fortbildung ein?“) als auch zum Einsatz des Messinstruments *Direct Behavior Rating* („Bewerten Sie die Nützlichkeit des Messinstruments“, „Wie durchführbar/praxistauglich finden Sie das Messinstrument?“) eingeholt. Auf der Ebene der Outcomes wurde die Wahrscheinlichkeit der zukünftigen Nutzung der Professionalisierungselemente abgefragt („Wie wahrscheinlich ist es, dass Sie in Zukunft freiwillige Beratungen in Anspruch nehmen?“, „Wie wahrscheinlich ist es, dass Sie in Zukunft die Direkte Verhaltensbeurteilung anwenden?“) und Raum für Verbesserungsvorschläge gegeben („Wozu hätten Sie noch Informationen benötigt?“, „Welche Verbesserungsvorschläge haben Sie?“). Die Zufriedenheits- und Nützlichkeitseinschätzungen erfolgten anhand geschlossener und offener Fragen mittels Fragenkatalog (Schritt 4). Die offenen Fragen wurden anonym und schriftlich auf einem Fragebogen beantwortet. Der Befragungsteil bezüglich der stattgefundenen *Beratung* wurde außerdem mithilfe der Fünf-Finger-Methode durchgeführt. Bei dieser Methode sind in den Umriss einer Hand Satzanfänge neben die einzelnen Finger geschrieben. Die Satzanfänge fungieren als Impulse, die zu einer kritischen Auseinandersetzung mit dem Befragungsgegenstand anregen sollen. Erfragt werden auf diese Weise folgende Bereiche: „Das fand ich gut.“ (Daumen), „Darauf möchte ich hinweisen.“/„Das ist mir aufgefallen.“ (Zeigefinger), „Das gefällt mir nicht.“ (Mittelfinger), „Das nehme ich mit.“ (Ringfinger) und „Das würde ich mir für die Zukunft wünschen.“ (Kleiner Finger). Die geschlossenen Fragen wurden auf einer 10-er Likert-Skala (0 = totale Ablehnung; 10 = tota-

le Zustimmung) teilweise schriftlich und anonym, teilweise für alle sichtbar auf einem Smiley-Barometer erhoben (Schritt 5). Dieses Verfahren erwies sich als zeitökonomisch und praktikabel, da durch die vorgegebenen Fragen der Vergleich und die Systematisierung der Antworten möglich war. Die Auswertung der Antworten (Schritt 6) erfolgte im Falle der geschlossenen Fragen deskriptiv und, im Falle der offenen Fragen, anhand eines vereinfachten Modells der Qualitativen Inhaltsanalyse (Dresing & Pehl 2018, S. 37ff.). Dieses Auswertungsverfahren wird in Kapitel 4.3.5.3 genauer erläutert.

4.3.5 Auswertungsverfahren, statistische Prüfgrößen und Interpretationsmaße

Zur Beantwortung der Fragestellungen und zur Überprüfung der aufgestellten Hypothesen wurden verschiedene Auswertungs- und Analyseverfahren sowie statistische Prüfgrößen und Interpretationsmaße berücksichtigt. Als unabhängige Variable werden in der vorliegenden Studie, so wie bei Ein-Gruppen-Plänen üblich, die MZP (Vorher- versus Nachher-Messungen) angesehen (vgl. Bortz & Döring 2016, S. 738). Die Wahl der jeweiligen Auswertungsverfahren und Prüfgrößen erfolgten abhängig von der Anzahl der MZP und der Struktur der abhängigen Variablen.

Zur Überprüfung der Hypothese hinsichtlich der Verbesserung der Schüler*innenoutcomes (Hypothese 1) wurden sogenannte Stückweise Regressionsanalysen, Mehrebenenanalysen mit Kreuzprädiktor sowie einfaktorielle Varianzanalysen (ANOVA) durchgeführt. Die Stückweise Regressionsanalyse erfolgte nach Huitema und McKean (2000) sowie deren mehrebenenanalytische Erweiterung nach Van den Noortgate und Onghena (2003). Zur Überprüfung der Hypothese über das veränderte Lehrkraftthandeln (Hypothese 2) erfolgte die Quantifizierung der Einsatzhäufigkeit der *84 Handlungsmöglichkeiten* und die Analyse von deren Interaktionseffekten mit den Schüler*innenoutcomes anhand einer Mehrebenenanalyse mit Kreuzprädiktor. Zur Überprüfung der Hypothese über veränderte Lehrkraftkognitionen (Hypothese 3) wurde eine ANOVA sowie eine Mehrebenenanalyse mit Kreuzprädiktor eingesetzt. Die Überprüfung der Hypothese über positive Lehrkraftreaktionen (Hypothese 4) erfolgte anhand eines vereinfachten Modells der Qualitativen Inhaltsanalyse nach Dresing und Pehl (2018). Die im Rahmen der Evaluationsstudie zum Professionalisierungskonzept angewandten und in Tabelle 13 dargestellten Auswertungsverfahren und statistischen Prüfgrößen lassen sich in verlaufdiagnostische, statusdiagnostische sowie qualitative Auswertungsverfahren gliedern.

Tabelle 13: Auswertungsverfahren und statistische Prüfgrößen, geordnet nach den untersuchten Evaluationsebenen, den abhängigen Variablen und den eingesetzten Messverfahren

Evaluationsebenen mit abhängigen Variablen	Messverfahren	Auswertungsverfahren	statistische Prüfgrößen
Schüler*innenoutcomes			
spezifische Verhaltensauffälligkeiten, nicht-regelkonformes Verhalten	DBR	Stückweise (mehrebenenanalytische) Regressionsanalyse	<i>p</i> -Wert <i>B</i> -Parameter
Verhaltensauffälligkeiten und -stärken	SDQ	ANOVA	<i>p</i> -Wert
Interaktionseffekte je nach Erklärungs- und Handlungsansatz der HM	Systematisierung in lerntheoretische/kognitionspsychologische/schüler*innen-zentrierte HM + DBR	Mehrebenenanalyse mit Kreuzprädiktor	<i>p</i> -Wert <i>B</i> -Parameter
unterrichtliches Lehrkrafthandeln	Monitoring über Einsatzhäufigkeit der HM +	Verhältnis zwischen Rating- und Einsatzhäufigkeit der HM	quantifizierte Einsatzhäufigkeit
Interaktionseffekte mit Schüler*innenoutcomes	DBR	Mehrebenenanalyse mit Kreuzprädiktor	<i>p</i> -Wert <i>B</i> -Parameter
Lehrkraftkognitionen	WirkLehr +	ANOVA	<i>p</i> -Wert
Interaktionseffekte mit Schüler*innenoutcomes	DBR	Mehrebenenanalyse mit Kreuzprädiktor	<i>p</i> -Wert <i>B</i> -Parameter
Lehrkraftreaktionen	Evaluationsfragebogen	Vereinfachtes Modell der Qualitativen Inhaltsanalyse	Quantifizierte Ober- und Unterkategorien

Anmerkungen. DBR = *Direct Behavior Rating* (Christ et al., 2009), SDQ = *Strengths and Difficulties Questionnaire* – Lehrkraftversion (Goodman, 2005), WirkLehr = Skala Lehrer-Selbstwirksamkeitserwartung (Schwarzer & Schmitz, 1999), HM = *84 Handlungsmöglichkeiten* (Hartke et al., 2018), *p* = Signifikanzwert, *B* = regressionsanalytische Parameter Level (Veränderung des Levels mit Einsetzen der Interventionsrate) und Slope (Steigung in der Interventionsrate), ANOVA = Einfaktorielle Varianzanalyse.

4.3.5.1 Verlaufsdiagnostische Auswertungsverfahren und Prüfgrößen

Verlaufsdiagnostische Erhebungen fanden statt, um die Veränderungen der Schüler*innenoutcomes bezüglich nichtregelkonformen Verhaltens und spezifischen Verhaltensauffälligkeiten im Verlauf zu untersuchen (Hypothese 1a, 1b, 1c, 2b, 3c). Die kontrollierten Einzelfalldaten wurden anhand des Messverfahrens *Direct Behavior Rating* erhoben und – den Empfehlungen experimenteller Einzelfallforschung folgend (Kazdin, 2021; Kratochwill et al., 2010; Manolov et al., 2016; Manolov & Moeyaert, 2017; Tanios & Onghena, 2020; Tate et al., 2016; Vannest & Ninci, 2014) – anhand einer Kombination und Staffelung deskriptiver Interpretationsmaße (Mittelwert, Standardabweichung, Trend), nicht-parametrischer Effektstärkemaße, sogenannter Überlappungsmaße (Nonoverlap-Techniken; Parker et al., 2011) und inferenzstatistischer Verfahren und Prüfgrößen (Huitema & McKean, 2000; Van den Noortgate & Onghena, 2003) ausgewertet. Dieses Vorgehen ist in Tabelle 14 dargestellt. Die Analysen der kontrollierten Einzelfalldaten wurden mit dem R-Paket Scan (Wilbert & Lüke, 2021) durchgeführt.

DBR-SIS und DBR-MIS wurden getrennt voneinander ausgewertet. Da die DBR-MIS aus drei Items mit jeweils 10-er Likert-Skalen bestehen (s. Kap. 2.1.3.4), die für die Auswertung zu einem Wert aggregiert wurden, erhöhen sich hier die Skalenwerte auf insgesamt 30 Punkte.

Verschiedene Prüfgrößen und Parameter kommen bei der Auswertung der Daten zum Einsatz. Zur Interpretation der deskriptiven Ergebnisse werden die Anzahl der MZP in A- und B-Phase (N), der Median (Med), die Standardabweichung (SD), der Medianunterschied zwischen den Phasen (ΔMed), der Trend, der Trendunterschied zwischen den Phasen ($\Delta Trend$) sowie die Überlappungsmaße (s. Kap. 4.3.5.1) angegeben. Der Median ist der Wert, der genau in der Mitte einer Datenreihe liegt, die nach der Größe geordnet ist. Aufgrund dieser zentralen Lage wird er auch Zentralwert genannt. Die Standardabweichung (Streuung) ist das Maß für die Streubreite der Werte eines Merkmals rund um dessen Mittelwert. Der Trend beschreibt die Entwicklung von Werten über die Zeit.

Zur Interpretation der inferenzstatistischen Ergebnisse werden die Regressionsparameter (B) Intercept (Konstante), Trend (Trend in der Grundrate), Level (Veränderung des Niveaus zur Interventionsrate), Slope (Steigung in der Interventionsrate) sowie der Standardfehler (SE) und der Signifikanzwert (p) angegeben. Die B -Regressionsparameter werden im Unterkapitel zur Stückweisen Regressionsanalyse genauer erklärt. Der Standardfehler ist ein Maß dafür, wie stark ein beobachteter Parameter – etwa der Mittelwert oder der Median – in einer Stichprobe durchschnittlich vom wahren Parameter der Grundgesamtheit abweicht. Das Signifikanzniveau ist eine per Konvention festgelegte Höchstgrenze der Alpha-Fehler-Wahrscheinlichkeit, die markiert, wie unwahrscheinlich ein Stichprobenergebnis sein muss, damit man die

Tabelle 14: Vorgehen bei der Auswertung der kontrollierten Einzelfallstudie durch die Kombination deskriptiver Interpretationsmaße, Effektstärkemaße und inferenzstatistischer Interpretations- und Prüfgrößen der Stückweisen Regressionsanalyse

Parameter	Prüfgrößen/Interpretationsmaße		Vergleich zwischen GR und IR
	GR	IR	
deskriptive Interpretationsmaße			
Niveau	M	M	Wie groß ist die Differenz der Mittelwerte?
Varianz	SD	SD	Sind die Werte während der IR "stabiler" und „enger“ als während der GR?
Trend	Trend	Trend	Ist ein zunehmender/abnehmender/kein Trend erkennbar? Ist der Trend wünschenswert/unerwünscht? Beeinflusst der Trend in der GR den Interventionseffekt?
nicht-parametrische Effektstärkemaße			
(Nicht-) Überlappung	PAND (+ d) PET NAP TAU-U (+ p)		In welchem Ausmaß fallen die Datenpunkte der GR in den Bereich der Daten der IR? Keine Überlappung, etwas Überlappung, moderate Überlappung, große (signifikante) Überlappung
inferenzstatistische Interpretations- und Prüfgrößen			
Stückweise (mehrebenenanalytische) Regressionsanalyse	Level-Effekt (+ p)		Zeigen sich in der IR unmittelbare und sprunghafte (signifikante) Veränderungen in den Daten?
	Slope-Effekt (+ p)		Zeigen sich in der IR kontinuierliche und kumulative (signifikante) Veränderungen in den Daten?

Anmerkungen. *M* = Mittelwert, *SD* = Standardabweichung, PAND = Percentage of all nonoverlapping Data (Parker, Hagan-Burke & Vannest, 2007), PET = Percentage of Data points Exceeding the Trend (Wilbert & Lüke, 2021), NAP = Nonoverlap of all Pairs (Parker & Vannest, 2009), TAU-U = (Parker et al., 2011), GR = Grundrate, IR = Interventionsrate, *p* = Signifikanzwert, *d* = Effektstärkemaß (Cohen, 1988).

Nullhypothese ablehnt und die Alternativhypothese annimmt. In der vorliegenden Studie wurde das Signifikanzniveau (Alpha-Fehler-Niveau) mit *p* = .05 festgelegt. Dies ist in der Forschung üblich (Bortz & Döring 2016, S. 664). Demensprechend gilt folgender Interpretationsmaßstab für das Signifikanzniveau: *p* ≤ .05 = *; *p* ≤ .01 = **; *p* ≤ .001 = ***. Die in Tabelle 14 dargestellten Auswertungsschritte sollen im Folgenden vorgestellt sowie anhand relevanter Interpretationsmaße und Prüfgrößen erläutert werden.

Deskriptive Interpretationsmaße

Im Zuge der Auswertung von verlaufsdiagnostisch erhobenen Daten im Rahmen von kontrollierten Einzelfallstudien können in einem ersten Schritt einige deskriptive Parameter überprüft werden. Dies bietet sich im Rahmen der sogenannten visuellen Inspektion an, einem Verfahren, bei dem die Daten zunächst grafisch, in der Regel in Form eines einfachen Liniendiagramms, dargestellt werden. Die Häufigkeit des Zielverhaltens wird auf der y-Achse und die Zeiteinheiten werden auf der x-Achse über die A- und B-Phase eingetragen (Gast & Spriggs, 2010). Drei Prüfgrößen sind bei der visuellen Inspektion von besonderer Relevanz: die Betrachtung des Niveaus (Median) und der Variabilität der Daten (Standardabweichung) sowie die Analyse des Trends. Veränderungen im Niveau beziehen sich auf die Unterschiede, die zwischen zwei Untersuchungsphasen in der durchschnittlichen Ausprägung des Zielverhaltens festzustellen sind. „Ermittelt wird das (durchschnittliche) Niveau, indem der [...] Median des Verhaltensmerkmals über die Zeitpunkte jeder Phase gebildet wird“ (Brunstein & Julius 2014, S. 131). Wenn sich etwa das Niveau sprunghaft nach Einsetzen der B-Phase verändert, kann dies auf die Wirksamkeit der Intervention hindeuten. Veränderungen der Variabilität beziehen sich auf die Streubreite der Daten. Eine geringere Streubreite und damit Variabilität in den Daten der Interventionsrate kann ein Indiz für einen Interventionseffekt darstellen. Eine über beide Phasen gleichbleibend hohe Variabilität der Daten kann hingegen – je nach der Eigenschaft der untersuchten abhängigen Variable – auf eine eher schwache Interventionswirkung hinweisen. Veränderungen im Trend beziehen sich auf die systematische Zu- oder Abnahme des Zielverhaltens über die Zeitpunkte einer Untersuchungsphase, hierbei müssen sich die Kurvenverläufe zweier Phasen „deutlich voneinander unterscheiden“ (Brunstein & Julius 2014, S. 131). Ist der Trend in der A-Phase positiv, in der B-Phase aber negativ, kann dies, je nach erwarteter Verlaufsrichtung, ein Hinweis für einen Interventionserfolg sein. Ebenso verhält es sich mit einem eher flachen Trend in der Grund- und einem steilen Trend in der Interventionsphase – dies kann für die Wirksamkeit der Intervention sprechen. Setzt sich der Trend aus der A-Phase in der B-Phase lediglich fort, kann man hingegen davon ausgehen, dass die Intervention nichts bewirkt hat (Brunstein & Julius 2014, S. 131).

Die experimentelle Einzelfallforschung hat jedoch gezeigt, dass die visuelle Analyse zum einen eine sehr geringe Interrater-Reliabilität aufweist (Heyvaert & Onghena, 2014; Manolov et al., 2016) und zum anderen, insbesondere bei Daten mit einem Trend in der Grundrate, anfällig für einen Beurteilungsfehler vom Typ 1 ist (Wilbert, Bosch & Lüke, 2021), demnach die Nullhypothese zurückgewiesen wird, obwohl sie eigentlich bestätigt werden müsste. Somit besitzt die visuelle Analyse allein wenig Aussagekraft für die Auswertung von kontrollierten Einzelfallstudien und sollte deshalb um weitere Auswertungsverfahren und Prüfgrößen, auch im Sinne weiterer Sensitivi-

tätsanalysen (Kratochwill et al., 2010), erweitert werden. Dies soll anhand von nicht-parametrischen Effektstärkemaßen und inferenzstatistischen Verfahren realisiert werden. Erstere werden im folgenden Abschnitt vorgestellt.

Nicht-parametrische Effektstärkemaße: Überlappingsmaße

Sogenannte Überlappingsmaße als nicht-parametrische Effektstärkemaße messen den Grad oder Prozentsatz an (Nicht-)Überlappung der Daten in der A- und B-Phase. Im Rahmen der deskriptiven Analyse der kontrollierten Einzelfallstudie wurden folgende Überlappingsmaße herangezogen: Percentage of all nonoverlapping Data (PAND; Parker et al., 2007), Percentage of data points exceeding the trend (PET; Wilbert & Lüke, 2021), Nonoverlap of all Pairs (NAP; Parker & Vannest, 2009) und Tau-U (Parker et al., 2011), die allesamt mit dem R-Paket Scan (Wilbert & Lüke, 2021) berechnet wurden. Da im Zuge der *Intervention* von einer Reduktion der spezifischen Verhaltensauffälligkeiten und des nichtregelkonformen Verhaltens ausgegangen wird, erfolgte die Berechnung der Überlappingsmaße unter der Annahme abnehmender Werte (Wilbert & Lüke, 2021). Für die Interpretation der errechneten Effektstärken weisen Parker et al. (2011) darauf hin, dass Effektgrößen nicht nur als klein, mittel oder groß beschrieben werden, sondern sich immer auch auf die praktische und klinische Bedeutung des untersuchten Verhaltens oder Konstrukts beziehen. Sie sollten im Vergleich zu ähnlichen Anwendungen der Intervention gesehen werden und/oder im Kontext der erwarteten oder gewünschten Veränderung interpretiert werden. Obwohl der (Nicht-)Überlappung der Datenpunkte in der A- und B-Phase zwar unterschiedliche Kriterien zugrunde liegen (Vergleich jedes Datenpunkts mit jedem Datenpunkt in A- und B-Phase vs. Überprüfung der Überlappungsfreiheit anhand eines Kriteriums in der Grundrate, wie etwa des Trends), ist das Rechenprinzip der verschiedenen Überlappingsmaße ähnlich. Aus diesem Grund soll die Berechnung anhand eines einfach per Hand zu errechnenden Überlappingsmaßes, des PAND, exemplarisch aufgezeigt werden.

Der PAND erfasst den Grad der „Überlappungsfreiheit“ zwischen Grundraterhebung und Interventionsphase numerisch (Parker et al., 2007). Wie in Abbildung 16 veranschaulicht, werden dabei zunächst die Datenpunkte, die sich in beiden Untersuchungsphasen überlappen, aus dem Liniendiagramm entfernt. Es werden nur so viele Punkte entnommen, wie für die Überlappungsfreiheit unbedingt erforderlich ist. Anschließend wird die Anzahl der Überlappungen (z.B. 1) über alle Einzelfälle aufsummiert und zur Gesamtzahl der Testungen (z.B. 20) ins Verhältnis gesetzt ($1:20 = 0.05$ oder 5 %). Anschließend wird der Anteil überlappender Daten von 1 abgezogen ($1 - 0.05 = 0.95$ oder 95 %). Bei völliger Überlappungsfreiheit würde das Ergebnis $PAND = 1$ lauten. Ein PAND-Wert von 0.50 (die Hälfte der Datenpunkte müsste entfernt werden) bedeutet, dass die Wirksamkeit der Intervention rein

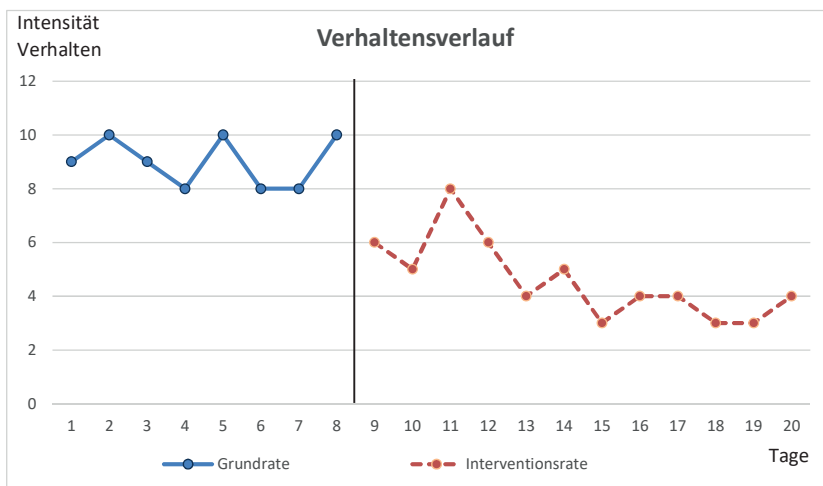


Abbildung 16: Bestimmung der Überlappungsfreiheit im Zuge der PAND-Berechnung (eigene Darstellung)

Anmerkung. PAND = Percentage of all nonoverlapping Data (Parker et al., 2007).

zufällig zustande kam. Bei einem PAND-Wert von < 0.90 spricht man von einem „starken“ Effekt und bei einem Wert von $0.70 - 0.90$ von einem „mittelstarken“ Effekt (Alresheed et al., 2013). Aus dem PAND-Wert kann auch das Effektstärkemaß d nach Jacob Cohen (1988) bestimmt werden (Parker & Hagan-Burke, 2007). Aus dem PAND-Wert wird zunächst ein Koeffizient abgeleitet (Phi), der die Stärke des Zusammenhangs zwischen den Untersuchungsphasen und der Höhe der Testergebnisse anzeigt. Danach kann d direkt berechnet werden. Die beiden Formeln dafür lauten:

$$(1) \text{Phi} = (2 \times \text{PAND}) - 1.$$

$$(2) d = (2 \times \text{Phi}) / \sqrt{(1 - \text{Phi}^2)}$$

Die Effektstärkeberechnung über PAND ist an zwei Voraussetzungen gebunden. Zum einen sollten die einzelnen Phasen mindestens 20 Datenpunkte umfassen. Zum anderen sollte, im Falle eines Multiple-Baseline-Designs, die Gesamtzahl der Messungen in den Interventionsphasen in etwa der Gesamtzahl der Messungen während der Grundraten entsprechen (Julius & Brunstein, 2014). Einschränkend muss erwähnt werden, dass bei der Berechnung von PAND keine Informationen über Trends berücksichtigt werden.

Der PET misst den Prozentsatz der Daten, die über oder unter dem in die Interventionsrate verlängerten Trend liegen. Es liegen keine Konventionen zur Interpretation des Koeffizienten vor, weshalb Kennwerte von vergleich-

baren Extended Celeration Line (ECL)-Techniken herangezogen werden. Hier kann man bei Werten ab 90 % von starken, bei Werten zwischen 70 % und 90 % von moderaten und bei Werten kleiner als 70 % von schwachen bzw. keinen Effekten ausgehen (vgl. Alresheed et al., 2013). Zusätzlich kann beim PET mit einem Binominaltest überprüft werden, ob die Anzahl der über dem vorausgesagten Trend liegenden Datenpunkte signifikant oberhalb/unterhalb des 95 %-Konfidenzintervalls des Erwartungswertes der Zufallsverteilung liegt (Wilbert & Lüke, 2021). Damit bezieht der PET den Trend der Grundrate in die Kalkulation mit ein.

Der NAP gibt den Prozentsatz aller paarweisen Vergleiche über die Baseline- und Interventionsphasen hinweg an. Für den NAP gelten nach Parker und Vannest (2009) Werte zwischen 93 und 100 % als großer Effekt, bei Werten zwischen 66 und 92 % ist von einem moderaten Effekt auszugehen und Werte kleiner als 66 % verweisen auf unzuverlässige bzw. keine Effekte.

Die Tau-U-Analyse erlaubt es, Interventionseffekte sowohl auf die Differenz zwischen den Phasen als auch auf den Trend innerhalb der Phasen zu untersuchen und bietet mindestens vier verschiedene Arten der Berechnung (Parker et al., 2011). Wenn untersucht werden soll, ob der Interventionseffekt sowohl zuverlässig als auch dauerhaft war, dann sollte auch der Trend der B-Phase berücksichtigt werden. Der Tau-U mit dem Zusatz „nonoverlap with Phase B trend with baseline trend controlled“ realisiert dies, indem er die Nicht-Überlappung aller Paare zwischen der Baseline- und der Interventionsphase plus den Trend der Interventionsphase minus den Trend der Baselinephase einbezieht (Parker et al. 2011, S. 291). Der TAU-U kann grundsätzlich Werte zwischen 1 und -1 annehmen, je nachdem, in welche Richtung der Trend der Datenpunkte verläuft. Nach Vannest und Vinci (2017) deuten Werte größer als .80 auf einen sehr großen, zwischen .60 und .80 auf einen großen, zwischen .20 und .60 auf einen moderaten und Werte kleiner als .20 auf einen kleinen Effekt hin. Da das Ziel der *Intervention* die Verringerung der spezifischen Verhaltensauffälligkeiten ist, würde sich deren Wirksamkeit durch negative TAU-U-Werte bemerkbar machen. Positive Werte würden hingegen für einen gegenteiligen Effekt sprechen. Die größte Einschränkung von Tau-U ist die schwere Vereinbarkeit mit der visuellen Analyse, sowohl in der Theorie als auch in der Praxis. So lässt sich das monotone TAU-U-Trendkontrollverfahren nur schwer visualisieren. Die Verwendung der Tau-U-Basistrendkorrektur ist daher eine Art „Black Box“-Prozess, der eine nuancierte Interpretation der Effektgrößenergebnisse erschwert (Brossart et al., 2018).

Zusammenfassend lässt sich festhalten, dass die nicht-parametrischen Effektstärkemaße in Form von Überlappungsmaßen das Berichten deskriptiver Parameter wie Niveau, Variabilität und Trend gut ergänzen. Sie können anhand ihrer jeweiligen Interpretationsmaße Hinweise über mögliche Interventionseffekte generieren. Die Effektstärkeberechnung anhand von Überlap-

pungsmaßen unterliegt aber auch Limitationen. Aus diesem Grund wurden in der vorliegenden Studie bei der Auswertung der kontrollierten Einzelfallstudie zusätzlich inferenzstatistische Verfahren eingesetzt, die im Folgenden erläutert werden sollen.

Inferenzstatistische Auswertungsverfahren und Prüfgrößen

Die Wirksamkeit einer Intervention kann neben der visuellen Inspektion und der Berechnung nicht-parametrischer Effektstärkemaße auch mittels inferenzstatistischer Verfahren wie Regressionsanalysen oder Randomisierungstests überprüft werden. Die inferenzstatistischen Verfahren ermöglichen Aussagen über wirksame Interventionen mit einer Irrtumswahrscheinlichkeit von weniger als 5 % (p -Wert). Randomisierungstests setzen ein bestimmtes Studiendesign voraus (u.a. Umkehrpläne und einen randomisierten Interventionsbeginn) und wurden daher in der vorliegenden Studie als mögliches Auswertungsverfahren ausgeschlossen. Regressionsbasierte Ansätze ermöglichen neben der Analyse eines Gesamteffekts der Intervention auch eine differenzierte Analyse von Effektmustern (Manolov & Moeyaert, 2017). Nur durch dieses Verfahren können sowohl Änderungen im Niveau als auch Änderungen im Trend adäquat berücksichtigt werden (Grosche 2014, S. 151). Außerdem können die Ergebnisse einer größeren Zahl an Fällen zusammengefasst und mehrbenenanalytisch ausgewertet werden, um verallgemeinernde Schlussfolgerungen über die Wirksamkeit der Intervention ziehen zu können (Van den Noortgate & Onghena, 2003). Aus diesen Gründen wurde die Stückweise Regressionsanalyse als inferenzstatistisches Auswertungsverfahren gewählt. Diese soll im folgenden Abschnitt explizit erläutert werden.

Die Stückweise Regressionsanalyse (*Piecewise Regression Model*, Huijtema & Mckean, 2000) ist ein inferenzstatistisches Verfahren, bei dem die unabhängige Variable in Intervalle unterteilt und die Regressionsfunktion in „Stücken“ modelliert wird. Bei einem AB-Versuchsplan, wie er in der vorliegenden Studie angewandt wurde, werden die Grund- und Interventionsrate als separate „Stücke“ behandelt. Dies geschieht, weil man davon ausgeht, dass an der Bruchstelle (Einsetzen der Intervention) eine qualitative Veränderung stattfindet, die sich auf das Intercept und die Steigung auswirkt. Für die Analyse von Einzelfalldaten, bei denen es sich aus statistischer Sicht um in Phasen unterteilte Zeitreihendaten handelt (Bortz & Döring, 2016), eignet sich dieser Ansatz sehr gut. Die Stückweise Regressionsanalyse ermöglicht, neben der Kontrolle von Entwicklungstrends in den Daten (Trendeffekte), die Untersuchung zweier Arten von Interventionseffekten: unmittelbare und abrupt einsetzende Level-Effekte und sich kumulativ und kontinuierlich realisierende Slope-Effekte. Während der Level-Effekt auf eine plötzliche Verhaltensänderung hinweist, die unmittelbar nach der Durchführung der Intervention auftritt, deutet der Slope-Effekt auf eine kontinuierliche und

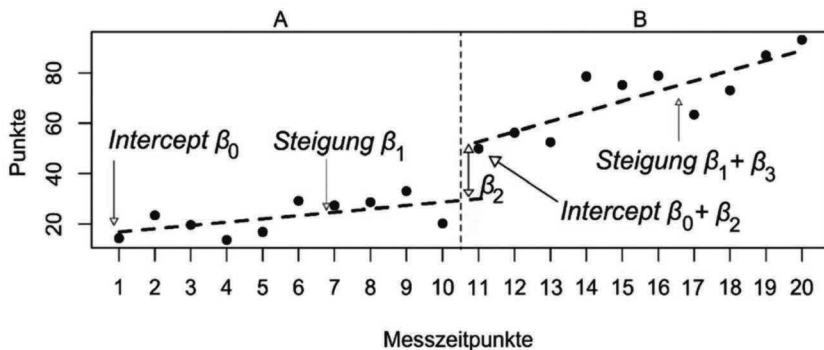


Abbildung 17: Regressionsmodell von Huitema und McKean (2000) zur Einzelfallanalyse, entnommen aus Wilbert (2014)

progressive Verhaltensänderung hin (Rindskopf & Ferron, 2014). Beide Interventionseffekte werden anhand der Interaktion zwischen dem Trend in der A-Phase und der Interventionsphase modelliert. Huitema und McKean schlagen zur Modellierung des Interaktionsterms vor, „dass mit Einsetzen der Intervention ein linearer Trend erwartet wird und zuvor der Term keinen Einfluss auf die Vorhersage der Messwerte ausübt“ (2000, S. 23). Das Prinzip der Stückweisen Regressionsanalyse, das in Abbildung 17 veranschaulicht ist, lässt sich folgendermaßen zusammenfassen:

Durch das Modell werden zwei Regressionsgeraden geschätzt. Das Ausgangsniveau zum Beginn der Intervention (MZP 1) wird durch das Intercept der ersten Regressionsgeraden β_0 überschlagen. Der Trend in der Phase vor der Intervention durch den Steigungsparameter β_1 , der Haupteffekt der Intervention durch β_2 . Die Zuwachsrate [...] als Folge der Intervention β_3 ergibt zusammen mit dem Trend β_1 den Steigungsparameter der zweiten Regressionsgerade. Das Intercept dieser Regressionsgeraden bildet sich aus der Summe des Ausgangsniveaus β_0 , dem Anstieg im Laufe der A-Phase und des Haupteffektes β_2 . Kommt es in der B-Phase zu bedeutsamen Abweichungen von dem auf Basis der A-Phase prognostizierten Regressionsmodell[s], so liegt ein Indiz für die Wirksamkeit der Maßnahme vor (Wilbert 2014, S. 23).

Mit der Stückweisen Regressionsanalyse können die Ergebnisse einer größeren Zahl an Fällen auch zusammengefasst und mehrerebenenanalytisch ausgewertet werden, um verallgemeinernde Schlussfolgerungen über die Wirksamkeit der Intervention ziehen zu können (Van den Noortgate & Onghena, 2003). Diese Kombination von Regressionsmodellen erlaubt Rückschlüsse auf die Interventionseffekte über eine Gruppe von Einzelfällen, indem alle Fälle als ein Einzelfall mit Messpunkten (Ebene 1) zu Schüler*innen (Ebene 2) geclustert werden (Hypothesen 1a und 1b).

Bei komplexeren Analysen (Hypothesen 1c, 2b und 3c) können zusätzliche Prädiktoren in das Stückweise Regressionsmodell aufgenommen werden. Auf diese Weise können beispielsweise Interaktionseffekte zwischen zwei Variablen, wie etwa dem regelkonformen Schüler*innenverhalten und dem Selbstwirksamkeitsempfinden der Lehrkräfte, untersucht werden. Eine der beiden Interaktionsvariablen fungiert hierbei als sogenannter Kreuzprädiktor. Ein starker Interaktionseffekt würde anhand eines signifikanten Level-Effekts sichtbar werden (Wilbert & Lüke, 2021).

4.3.5.2 Statusdiagnostische Auswertungsverfahren und Prüfgrößen

Die Berechnung der Unterschiede der Verhaltensauffälligkeiten und -stärken (Hypothesen 1d und 1e) der Schüler*innen (SDQ; Goodman, 2005) und des Selbstwirksamkeitsempfindens (Hypothesen 3a und 3b) der Lehrkräfte (WirkLehr; Schwarzer & Schmitz, 1999) über die Zeit erfolgte durch die einfaktorielle Varianzanalyse (ANOVA = *analysis of variance*) mit Messwiederholungen anhand der Statistiksoftware SPSS. Bei Eingruppen-Plänen und „mehr als zwei Messungen erfolgt die statistische Hypothesenprüfung über die einfaktorielle Varianzanalyse mit Messwiederholungen“ (Bortz & Döring 2016, S. 739). Varianzanalysen sind statistische Verfahren, die zur Prüfung von Mittelwertunterschieden einer abhängigen Variablen für verschiedene Stufen einer mehrklassigen unabhängigen Variablen oder mehrerer unabhängiger Variablen angewandt werden. Die einfaktorielle Varianzanalyse mit Messwiederholung testet, ob sich die Mittelwerte mehrerer abhängiger Gruppen (oder Stichproben) unterscheiden, wobei der MZP die unabhängige Variable darstellt. Die Signifikanzprüfung erfolgte nach der *Bonferroni-Korrektur*. Mit diesem Verfahren wird die globale Erhöhung der Alpha-Fehler-Wahrscheinlichkeit, die durch multiples Testen in ein und derselben Stichprobe entsteht, neutralisiert.

Zur Interpretation der statusdiagnostischen Ergebnisse werden die Anzahl (N), der Mittelwert (M), die Standardabweichung (SD), das Konfidenzintervall (KI), die Mittelwertsdifferenzen (M_{Diff}), die Standardfehlerdifferenz (SE_{Diff}), der Signifikanzwert (p) und, bei Vorliegen signifikanter Unterschiede, die Effektstärke (ES) nach Cohen (1988) dargestellt. Der Mittelwert kennzeichnet die Summe aller Messwerte dividiert durch die Anzahl der eingehenden Werte. Das Konfidenzintervall (Vertrauensintervall) gibt die Präzision der Schätzung des Mittelwerts an und kennzeichnet denjenigen Bereich der Merkmalsausprägungen, in dem sich 95/99 % aller möglichen Populationsparameter befinden, die den empirisch ermittelten Stichprobenkennwert erzeugt haben können. Dies bedeutet, dass mindestens 95/99 % aller auf Grundlage von Messdaten berechneten Konfidenzintervalle den wahren Wert der Grundgesamtheit beinhalten. Es wurde ein Konfidenzkoeffizient von 95 % gewählt (Bortz & Döring 2016, S. 673). Bei signifikanten Mittelwertsunterschieden erfolgt die Berechnung der Effektstärke nach Cohen (1988),

die sich in den Human- und Sozialwissenschaften als geeignete Schätzung etabliert hat (Bortz & Döring 2016, S. 816). Bei der vorliegenden ANOVA wird dabei für den Haupteffekt das partielle Eta-Quadrat (η^2) anhand des verwendeten SPSS-Programms berechnet. Das Maß gibt an, wieviel Varianz der abhängigen Variablen durch die unabhängige prozentual erklärt wird. Danach gilt $\eta^2 = .01$ als ein kleiner, $\eta^2 = .06$ als ein mittlerer und $\eta^2 = .14$ als ein großer Effekt.

4.3.5.3 Qualitative Auswertungsverfahren und Prüfgrößen

Zur Auswertung und Interpretation der qualitativen Daten bei geschlossenen Fragen wurden die Parameter Mittelwert (M) und Standardabweichung (SD) genutzt. Für die Interpretation der Daten wurde die 10-er Likert-Skala (0 = totale Ablehnung; 10 = totale Zustimmung) nochmals in kleinteiligere Interpretationsmaße gegliedert, deren Bezeichnung sich an den gängigen Begrifflichkeiten von Schulnoten orientiert: 10-9 = sehr gut, 8-7 = gut, 6-5 = befriedigend, 4-3 = ausreichend, 2-0 = ungenügend.

Zur Auswertung und Interpretation der Antworten des offenen Frageteils wurde ein vereinfachtes Modell der Qualitativen Inhaltsanalyse genutzt. Die Überprüfung der Hypothesen zu den Lehrkraftreaktionen (Hypothesen 4a und 4b) erfolgte im Falle der offenen Fragen regelgeleitet anhand eines vereinfachten Modells der Qualitativen Inhaltsanalyse (Dresing & Pehl 2018, S. 37ff.). Ganz allgemein lässt sich die qualitative Inhaltsanalyse als ein Verfahren zur systematischen und zusammenfassenden Beschreibung von Datenmaterial durch Zuordnung von relevanten Dateneinheiten in Kategorien mit relevanten Bedeutungsaspekten definieren (Schreier, 2014). Hierbei wird der vorhandene Inhalt zunächst verdichtet und strukturiert, um dann das Material zusammenzufassen und auf Besonderheiten, Unterschiede sowie Zusammenhänge hin zu untersuchen. Das für die Auswertung angewandte vereinfachte Modell sieht folgende drei Schritte der Auswertung vor: 1) Lesen und Entdecken, 2) Strukturieren, 3) Zusammenfassen und Beschreiben. Im Zuge dieses Vorgehens wurde der sinntragende Inhalt des offenen Antwortteils je nach Beurteilungsschwerpunkt (gesamte Professionalisierungsmaßnahme oder einzelne Professionalisierungselemente) in erste Themenkomplexe bzw. Kategorien unterteilt. Sodann erfolgte eine feinere thematische Unterteilung der Antworten in Unterkategorien. Nach der erfolgten Zuordnung in Unterkategorien wurden die sinntragenden Textstellen stichpunktartig zusammengefasst und anhand ihrer Auftretenshäufigkeit in den Ober- und Unterkategorien quantifiziert (n = Häufigkeit der Nennungen pro Kategorie). Zur Interpretation der Antworten wurden die quantifizierten Ober- und Unterkategorien hinsichtlich ihrer Relevanz (Häufigkeit der Nennungen) systematisiert und ausgewertet.

5 Ergebnisse

5.1 Hypothesenüberprüfung

5.1.1 Schüler*innenoutcomes

Spezifische Verhaltensauffälligkeiten vor und während der Intervention (DBR-MIS Summenwerte)

Begonnen werden soll mit den deskriptiven Ergebnissen, die in der Tabelle 15 für jeden Einzelfall zusammengefasst sind. Die Auswertung erfolgte anhand der gemittelten Werte über die Gesamtgruppe ($N = 27$), um eine Aussage über die Effektivität der *Intervention* treffen zu können. Die spezifischen Verhaltensauffälligkeiten wurden anhand einer DBR-MIS mit drei Ausprägungen auf jeweils 10-er Likert-Skalen erfasst und für die Auswertung auf 30 Skalenwerte aufsummiert. Dies muss bei der Auswertung der Ergebnisse beachtet werden.

Über alle Fälle sind in der Grundrate im Mittel $N_A = 17.15$ (min. = 11, max. = 22) und in der Interventionsrate $N_B = 41.26$ (min. = 27, max. = 50) und somit über beide Phasen $N_{A+B} = 58.41$ Datenpunkte vorhanden. Da an insgesamt 67 Tagen geratet werden konnte, entspricht dies einer durchschnittlichen Ratinghäufigkeit von 87 %. Der Median der Werte beträgt in der A-Phase im Mittel $Med_A = 15.43$ (min. = 6, max. = 25, $SD_A = 4.54$) und in der B-Phase $Med_B = 12.56$ (min. = 2, max. = 26, $SD_B = 4.59$). Über alle Fälle reduziert sich der Median von der Grund- zur Interventionsrate damit um $\Delta Med = -2.87$ Punkte, die gemittelten Standardabweichungen für Grund- und Interventionsraten sind dabei fast identisch hoch. Der Trend der Verhaltensverläufe in der Grundrate ist mit $n = 18$ Fällen ($M_{TrendA} = 0.32$, min. = 0.03, max. = 1.41) zu zwei Dritteln steigend, ein Drittel der Fälle ($n = 9$) weist in der Grundrate einen eher fallenden Trend auf ($M_{TrendB} = -0.15$, min. = -0.01, max. = -0.28). In der Interventionsrate zeigt sich ein umgekehrtes Bild: Während mit $n = 18$ ($M_{Trend} = -0.12$, min. = -0.01, max. = -0.28) zwei Drittel der Fälle einen fallenden Trend aufzeigen, ist bei $n = 6$ ein ganz leicht steigender Verhaltensverlauf ($M_{Trend} = 0.06$, min. = 0.01, max. = 0.10) und bei einigen wenigen Fällen ($n = 3$) keine Trendbewegung zu verzeichnen. Die Trenddifferenz zwischen Grund- und Interventionsrate über alle Fälle hinweg beträgt $\Delta Trend = -0.23$, das bedeutet, der Trend ist über alle Fälle und Phasen hinweg deutlich abfallend.

Tabelle 15: Deskriptive Ergebnisse der spezifischen Verhaltensauffälligkeiten (DBR-MIS Summenwerte) für jedes Kind und über alle Schüler*innen

Schüler*innen	N_A	N_B	$Med_A (SD)$	$Med_B (SD)$	$Trend_A$	$Trend_B$	$\Delta Trend$
Kind 1	21	36	11.0 (4.85)	3.5 (6.83)	0.04	-0.17	-0.21
Kind 2	17	43	15.0 (1.92)	14.0 (1.78)	-0.03	0.01	0.04
Kind 3	19	41	6.0 (5.37)	10.0 (4.55)	-0.13	0.00	0.12
Kind 4	21	36	25.0 (5.42)	18.5 (5.91)	0.25	-0.28	-0.53
Kind 5	15	46	22.0 (2.13)	14.5 (3.28)	0.07	-0.10	-0.17
Kind 6	17	50	9.0 (8.33)	2.0 (2.19)	1.41	0.04	-1.37
Kind 7	16	27	20.0 (1.96)	22.0 (2.42)	-0.01	0.10	0.11
Kind 8	20	37	21.5 (6.44)	22.0 (8.47)	0.09	0.10	0.00
Kind 9	17	48	8.0 (4.51)	5.5 (4.43)	-0.24	-0.07	0.17
Kind 10	18	40	15.5 (6.15)	14.0 (5.38)	0.60	0.07	-0.53
Kind 11	15	50	15.0 (4.61)	7.0 (2.36)	0.81	-0.10	-0.91
Kind 12	19	36	19.0 (3.51)	14.0 (5.07)	-0.08	-0.10	-0.02
Kind 13	16	50	22.0 (0.96)	14.0 (3.23)	0.15	-0.09	-0.24
Kind 14	21	44	10.0 (3.73)	12.0 (4.05)	0.35	0.00	-0.35
Kind 15	22	36	23.5 (3.53)	26.0 (6.98)	0.26	-0.17	-0.43
Kind 16	14	44	17.0 (7.48)	17.5 (7.54)	-0.28	0.00	0.28
Kind 17	22	36	22.0 (6.30)	21.0 (4.77)	0.18	-0.11	-0.29
Kind 18	22	37	6.0 (4.14)	5.0 (3.91)	0.15	-0.08	-0.22
Kind 19	22	40	9.5 (3.94)	9.0 (3.91)	0.23	-0.12	-0.35
Kind 20	15	38	16.5 (5.38)	16.0 (6.12)	-0.28	-0.11	0.16
Kind 21	18	41	15.5 (6.78)	15.0 (4.56)	0.60	-0.11	-0.72
Kind 22	11	39	23.0 (2.05)	12.0 (4.31)	-0.30	-0.02	0.28
Kind 23	20	43	17.0 (5.90)	12.0 (6.62)	0.18	-0.05	-0.23
Kind 24	14	47	11.5 (5.24)	6.0 (7.47)	0.18	-0.01	-0.19
Kind 25	17	50	11.0 (4.85)	8.5 (4.94)	0.14	-0.28	-0.42
Kind 26	14	37	16.5 (3.56)	13.0 (3.67)	0.12	-0.12	-0.24
Kind 27	22	42	8.5 (3.67)	5.0 (3.20)	-0.03	0.03	0.06
M über alle Schüler*innen	17.15	41.26	15.43 (4.54)	12.56 (4.59)	0.16	-0.06	-0.23

Anmerkungen. N_A = Anzahl MZP in der Baselinephase, N_B = Anzahl MZP in der Interventionsphase, Med = Median, SD = Standardabweichung, $\Delta Trend$ = Trendunterschied zwischen den Phasen, M = Mittelwert.

Anhand der deskriptiven Analyse der gemittelten Mediandifferenzen, Standardabweichungen und Trends in Grund- und Interventionsraten wird sichtbar, dass sich die spezifischen Verhaltensauffälligkeiten über die Gesamtgruppe der Schüler*innen während der Interventionsphase sowohl im Niveau als auch im Trend reduziert haben und dass in den Grund- und Interventionsphasen eine gleichbleibend hohe Streubreite der Werte um den Median besteht.

Betrachtet man zunächst die Überlappingsmaße der Einzelfälle, die Tabelle 16 zu entnehmen sind, so zeigen sich beim PAND vorwiegend schwache ($n = 15$), einige mittlere ($n = 6$) und wenige große ($n = 3$) sowie unzuverlässige ($n = 3$) Effektstärken. Laut PET sind starke ($n = 8$), moderate ($n = 9$) und fragwürdige Effekte ($n = 10$) über die Schüler*innengruppe hinweg recht gleichmäßig verteilt. Der Binominaltest zeigt jedoch darüber hinaus, dass bei zwei Dritteln der Fälle ($n = 18$) die Datenpunkte der Interventionsrate signifikant unterhalb des 95 %-Konfidenzintervalls des Erwartungswertes der Zufallsverteilung des vorausgesagten Trends liegen. Der NAP verweist bei den Einzelfällen auf eine hälftige Effektivität: Bei $n = 4$ zeigt sich ein großer und bei $n = 9$ ein mittlerer Effekt, bei der anderen Hälfte der Schüler*innen ($n = 14$) ist kein Effekt zu verzeichnen. Der TAU-U ergibt für die Mehrheit der Fälle eine moderate Effektstärke ($n = 11$), die zweitgrößte Gruppe bewegt sich bei einem kleinen Effekt ($n = 8$), bei $n = 7$ haben sich keine Effekte ergeben und bei einem Fall ($n = 1$) ist ein großer Effekt zu konstatieren.

Mittelt man die Überlappingsmaße aller Schüler*innen ($N = 27$) so verweisen PAND (67 %) und NAP (63.02 %) auf schwache, tendenziell moderate Veränderungen in der Interventionsphase, PET (70.53 %) und TAU-U (-.25) indizieren hingegen einen moderaten Effekt.

Betrachtet man die Einzelfälle, zusammengefasst in Tabelle 17, so kann zunächst festgehalten werden, dass sich bei nahezu allen Fällen ($n = 25$) die spezifischen Verhaltensauffälligkeiten zu Anfang der Grundrate signifikant von Null unterscheiden (Intercept). In der Grundrate weisen insgesamt zwei Drittel der Verhaltensverläufe ($n = 18$) positive Trends auf, wovon sich bei $n = 5$ die spezifischen Verhaltensauffälligkeiten signifikant erhöht haben. Mit Einsetzen der *Intervention* zeigen insgesamt zwei Drittel der Fälle ($n = 18$) sprunghafte Abfälle der Verhaltensverläufe, bei $n = 5$ sind diese Level-Effekte signifikant. Die Reduktion des auffälligen Verhaltens zeigt sich auch über die kontinuierliche Senkung der Verhaltensverlaufsgraphen. Dies ist bei insgesamt $n = 19$ Verläufen der Fall. Signifikante Slope-Effekte werden bei $n = 10$ (min. = -0.35, max. = -7.78) sichtbar.

Insgesamt werden bei nahezu der Hälfte der Einzelfälle ($n = 12$) signifikante Veränderungen in den Interventionsraten durch Slope- und/oder Level-Effekte sichtbar. Von diesen zeigen $n = 7$ Slope-Effekte, $n = 2$ Level-Effekte und $n = 3$ sowohl Level- als auch Slope-Effekte. Unter Letzteren weisen alle Fälle außerdem signifikante positive Trend-Effekte auf.

Tabelle 16: Überlappungsmaße hinsichtlich der spezifischen Verhaltensauffälligkeiten (DBR-MIS Summenwerte) für jedes einzelne Kind und gemittelt über alle Schüler*innen

Schüler*innen	PAND (%)	PET (%)	<i>p</i> (PET _{binominal-95%CI})	NAP (%)	TAU-U
Kind 1	68.42	86.11	<.001	75.86	-.37***
Kind 2	70.00	46.51	.73	66.21	-.12
Kind 3	50.00	9.76	1.00	33.31	.13
Kind 4	78.95	94.44	<.001	71.56	-.41***
Kind 5	96.72	100.00	<.001	97.68	-.54***
Kind 6	91.04	100.00	<.001	92.18	-.33***
Kind 7	39.53	25.93	.10	29.05	.32**
Kind 8	47.37	51.35	.50	44.86	.06
Kind 9	69.23	12.50	1.00	65.38	-.20*
Kind 10	62.07	97.50	<.001	46.32	.05
Kind 11	96.92	100.00	<.001	94.13	-.66***
Kind 12	63.64	58.33	.20	70.03	-.24**
Kind 13	84.85	98.00	<.001	93.50	-.44***
Kind 14	50.77	86.36	<.001	28.68	.14
Kind 15	41.38	83.33	<.001	35.67	.03
Kind 16	58.62	11.36	1.00	42.45	.06
Kind 17	62.07	83.33	<.001	53.91	-.11
Kind 18	59.32	89.19	<.001	66.40	-.24**
Kind 19	58.06	92.50	<.001	53.69	-.16
Kind 20	69.23	13.16	1.00	57.33	-.12
Kind 21	62.71	100.00	<.001	51.56	-.16
Kind 22	88.00	61.54	.10	95.10	-.18
Kind 23	61.90	81.40	<.001	59.53	-.14
Kind 24	60.66	80.85	<.001	65.20	-.12
Kind 25	67.16	88.00	<.001	73.59	-.58***
Kind 26	72.55	83.78	<.001	68.73	-.29**
Kind 27	78.12	69.05	.01	69.59	-.13
M über alle Schüler*innen	67.50***	70.53	<.01	63.02	-.25***

Anmerkungen. PAND = Percentage of all nonoverlapping data, PET = Percentage of data points exceeding the trend, p (PET_{binominal-95%CI}) = p -Wert des Binominaltests des PET gegen den Zufall, NAP = Nonoverlap of all pairs, Tau-U = Parker's Tau-U (A vs. B + Trend B – Trend A), M = Mittelwert.

* ($p \leq .05$); ** ($p \leq .01$); *** ($p \leq .001$).

Tabelle 17: Stückweise Regressionsanalyse hinsichtlich der spezifischen Verhaltensauffälligkeiten (DBR-MIS Summenwerte) für jedes Kind

Schüler*innen	Intercept (SE)	Trend (SE)	Level (SE)	Slope (SE)
Kind 1	10.80*** (2.72)	0.04 (0.20)	-1.32 (3.17)	-0.21 (0.22)
Kind 2	15.54*** (0.94)	-0.03 (0.09)	-0.95 (1.04)	0.04 (0.09)
Kind 3	9.31*** (2.25)	-0.13 (0.17)	4.45 (2.57)	0.12 (0.18)
Kind 4	20.39*** (2.15)	0.25 (0.17)	-0.20 (2.64)	-0.53** (0.18)
Kind 5	21.09*** (1.41)	0.07 (0.14)	-5.35** (1.60)	-0.17 (0.14)
Kind 6	-1.14 (1.45)	1.41*** (0.14)	-21.67*** (1.56)	-1.37*** (0.14)
Kind 7	20.20*** (1.16)	-0.01 (0.11)	0.44 (1.36)	0.11 (0.12)
Kind 8	18.25 (3.68)	0.09 (0.28)	-2.84 (4.28)	0.00 (0.30)
Kind 9	10.55*** (2.23)	-0.24 (0.22)	1.12 (2.42)	0.17 (0.22)
Kind 10	7.72* (2.92)	0.60** (0.22)	-7.15* (2.99)	-0.53* (0.23)
Kind 11	8.05*** (0.92)	0.81*** (0.09)	-12.38*** (1.03)	-0.91*** (0.09)
Kind 12	19.69*** (2.15)	-0.08 (0.16)	0.00 (2.52)	-0.02 (0.17)
Kind 13	20.06*** (1.37)	0.15 (0.14)	-6.14*** (1.54)	-0.24 (0.14)
Kind 14	6.02*** (1.67)	0.35** (0.13)	-0.21 (1.96)	-0.35* (2.60)
Kind 15	19.96*** (2.48)	0.26 (0.19)	1.67 (2.92)	-0.43* (0.20)
Kind 16	18.64** (6.33)	-0.28 (0.47)	3.74 (4.70)	0.28 (0.48)
Kind 17	17.78*** (2.34)	0.18 (0.18)	0.80 (2.78)	-0.29 (0.19)
Kind 18	6.01** (1.74)	0.15 (0.13)	-1.91 (2.08)	-0.22 (0.14)
Kind 19	6.90*** (1.62)	0.23 (0.12)	-0.16 (1.93)	-0.35** (0.13)
Kind 20	20.86*** (4.92)	-0.28 (0.34)	3.07 (3.45)	0.16 (0.34)
Kind 21	7.18* (2.69)	0.60** (0.20)	-3.40 (2.75)	-0.72** (0.21)
Kind 22	26.70*** (2.20)	-0.30 (0.21)	-0.30 (0.21)	-7.78** (2.46)
Kind 23	13.77*** (2.95)	0.18 (0.23)	-3.22 (3.46)	-0.23 (0.24)
Kind 24	10.05** (3.76)	0.18 (0.38)	-4.21 (4.23)	-0.19 (0.38)
Kind 25	11.51*** (1.74)	0.14 (0.17)	1.93 (1.87)	-0.42* (0.17)
Kind 26	14.59*** (2.90)	0.12 (0.20)	-0.38 (2.04)	-0.24 (0.20)
Kind 27	8.30*** (1.50)	-0.03 (0.11)	-2.87 (1.81)	0.06 (0.12)

Anmerkungen. Intercept = Konstante, Trend = Trend in der Grundrate, Level = Veränderung des Levels zur Interventionsrate, Slope = Steigung in der Interventionsrate, SE = Standardfehler.

* ($p \leq .05$); ** ($p \leq .01$); *** ($p \leq .001$).

Tabelle 18: Mehrebenenanalyse hinsichtlich der spezifischen Verhaltensauffälligkeiten (DBR-MIS Summenwerte) über alle Schüler*innen

Parameter	B	SE	df	t	p
Intercept	12.95	1.04	1568	12.42	<.001
Trend	0.21	0.04	1568	5.25	<.001
Level	-2.45	0.53	1568	-4.59	<.001
Slope	-0.28	0.04	1568	-6.92	<.001

Anmerkungen. Intercept = Konstante, Trend = Trend in der Grundrate, Level = Veränderung des Levels zur Interventionsrate, Slope = Steigung in der Interventionsrate, SE = Standardfehler, df = Freiheitsgrade, t = t-Wert, p = Signifikanzwert.
 Schätzungsmethode für Slope: B&L-B/6.

Die Mehrebenenanalyse der 27 Einzelfälle als ein aggregierter Einzelfall mit Messpunkten (Ebene 1) geclustert zu Schüler*innen (Ebene 2) ist in Tabelle 18 dargestellt. Die Analyse ergibt für alle Parameter signifikante Effekte. Das Intercept verweist mit einem Wert von 12.95 darauf, dass sich über alle Fälle die spezifischen Verhaltensauffälligkeiten zu Anfang der Grundrate signifikant von Null unterscheiden. Der signifikante Trend-Effekt macht deutlich, dass der Trend in der Grundrate über alle Fälle um 0.21 Skalenspunkte ansteigt. Das bedeutet, dass sich die spezifischen Verhaltensauffälligkeiten aller Schüler*innen vor der *Intervention* tendenziell verstärkt haben. Der Level-Effekt über alle Fälle verweist auf eine Verhaltensverbesserung unmittelbar nach Beginn der *Intervention* von -2.45 Skalenspunkten. Relativiert man diesen Effekt anhand der maximal zu erreichenden 30 Skalenspunkte, so bedeutet dies eine beachtliche sprunghafte Reduktion der spezifischen Verhaltensauffälligkeiten um etwa 10 % mit Einsetzen der Intervention. Der Slope-Effekt zeigt auf, dass während der Interventionsphase eine kontinuierliche Reduzierung der auffälligen Verhaltensweisen über alle 27 Schüler*innen um -0.28 Skalenspunkten pro Messung einsetzte. Dies bedeutet konkret, dass sich durchschnittlich mit jedem MZP das Verhalten um etwa 1 % verbesserte. Dieser Slope-Effekt prognostiziert außerdem, dass nach 25 Interventionstagen eine Reduktion der Verhaltensauffälligkeiten um 7,5 Skalenspunkte erreicht wird, nach 50 Interventionstagen bereits um 15 Skalenspunkte.

16 B&L-B = Berry and Lewis-Beck Modell (Huitema & McKean 2000, S. 45); Erläuterung: Zur Schätzung des Slope-Effekts jeder Phase muss eine andere Art von Dummy-Variablen erstellt werden. Wie bei den Dummy-Variablen für Level-Effekte werden die Werte für alle Messungen mit Ausnahme derjenigen der betreffenden Phase auf Null gesetzt. Hier beginnen die Werte mit jeder Messung bis zum Ende der Phase zu steigen. Für die Art und Weise, wie diese Werte ansteigen, sind verschiedene Vorschläge gemacht worden. Das Berry and Lewis-Beck Modell (B&L-B-Modell) beginnt mit einer Eins bei der ersten Messung der Phase und erhöht sich mit jeder Messung, während etwa das H-M-Modell mit einer Null beginnt (Wilbert & Lüke, 2021; Huitema & McKean, 2000).

Regelverhalten vor und während der Intervention (DBR-SIS)

Das Regelverhalten vor und während der *Intervention* wurde ebenfalls mithilfe des Messinstruments DBR anhand der Single-Item-Scale (SIS) „Hält sich nicht an Absprachen und allgemeine Regeln“ auf einer 10-er Likert-Skala erfasst. Die Auswertung erfolgte, wie bei den DBR-MIS auch, anhand deskriptiver Statistiken, nicht-parametrischer Effektstärkemaße und Regressionsanalysen.

Begonnen werden soll mit den deskriptiven Ergebnissen, abgebildet in Tabelle 19. Die Auswertung erfolgte anhand der gemittelten Werte über die Gesamtgruppe ($N = 27$), um eine Aussage über die Effektivität der *Intervention* treffen zu können. In der Grundrate sind im Durchschnitt $N_A = 17.81$ (min. = 11, max. = 22) und in der Interventionsrate $N_B = 41.37$ (min. = 27, max. = 50) Datenpunkte vorhanden. Der Median der Werte beträgt in der A-Phase im Mittel $Med_A = 5.15$ (min. = 0.5, max. = 8, $SD = 2.02$) und in der B-Phase $Med_B = 4.37$ (min. = 1, max. = 8, $SD = 1.85$). Der Median über alle Fälle reduziert sich damit von Grund- zu Interventionsrate um $\Delta Med = -0.78$ Punkte. Die gemittelten Standardabweichungen in den Grund- und Interventionsraten sind ähnlich hoch, in der Interventionsrate ist jedoch eine leichte Reduzierung der Streuung zu verzeichnen. Der Trend der Verhaltensverläufe in der Grundrate verläuft bei knapp zwei Dritteln der Fälle ($n = 17$) steigend ($M_{Trend} = 0.14$, min. = 0.01, max. = 0.22). Bei $n = 10$ zeigt sich in der Grundrate ein fallender Trend ($M_{Trend} = -0.06$, min. = -0.01, max. = -0.13). In der Interventionsrate wird ein umgekehrtes Bild sichtbar, hier werden vorrangig fallende Trends in den Verhaltensverläufen erkennbar. So ist bei $n = 19$ ein sinkender Trend zu verzeichnen ($M_{Trend} = -0.14$, min. = -0.02, max. = -0.56), bei $n = 8$ zeigt sich ein steigender Verhaltensverlauf ($M_{Trend} = 0.06$, min. = 0.02, max. = 0.18). Die gemittelte Trenddifferenz über alle Schüler*innen beträgt $\Delta Trend = -0.08$, das bedeutet, der Trend ist über alle Fälle und Phasen hinweg leicht abfallend.

Die deskriptive Analyse der gemittelten Werte macht deutlich, dass sich das Regelverhalten der Schüler*innen in der Interventionsrate über alle Fälle hinweg verbessert hat. Sichtbar wird dies anhand der Mediandifferenzen, der durchschnittlichen Entwicklung der Trends in den A- und B-Phasen und der Verringerung der Streubreite der Werte während der Interventionsrate.

Im Folgenden sollen die nicht-parametrischen Überlappungsmaße ausgewertet werden. Betrachtet man zunächst die Überlappungsmaße der Einzelfälle, die Tabelle 20 zu entnehmen sind, so zeigen sich beim PAND vorwiegend schwache ($n = 12$) und mittlere ($n = 10$) sowie wenige große ($n = 2$) sowie unzuverlässige ($n = 3$) Effektstärken. Laut PET ergeben sich vorrangig starke ($n = 12$) und einige moderate ($n = 6$) Effektstärken. Ein Drittel der Fälle zeigt beim PET eher fragwürdige Effekte ($n = 9$). Durch den Binominaltest des PET wird jedoch deutlich, ähnlich wie bei den MIS-Summen-

Tabelle 19: Deskriptive Ergebnisse des Regelverhaltens (DBR-SIS) für jedes Kind

Schüler*innen	N_A	N_B	Med_A (SD)	Med_B (SD)	$Trend_A$	$Trend_B$	$\Delta Trend$
Kind 1	21	36	5.0 (2.25)	2.0 (2.60)	-0.03	-0.06	-0.03
Kind 2	17	42	5.0 (1.10)	4.5 (0.77)	0.03	-0.01	-0.03
Kind 3	19	41	7.0 (2.61)	5.0 (2.11)	0.04	-0.01	-0.05
Kind 4	21	36	7.0 (2.27)	5.0 (2.40)	0.14	-0.11	-0.24
Kind 5	15	46	8.0 (0.41)	5.0 (0.67)	-0.01	-0.03	-0.02
Kind 6	17	50	2.0 (3.34)	1.0 (0.92)	0.58	0.02	-0.56
Kind 7	17	27	8.0 (1.20)	6.0 (0.87)	0.01	0.04	0.03
Kind 8	20	37	7.0 (3.23)	8.0 (2.93)	0.20	0.06	-0.14
Kind 9	17	48	2.0 (2.12)	1.0 (1.60)	-0.09	-0.02	0.07
Kind 10	18	40	5.5 (3.08)	5.0 (1.66)	0.18	0.00	-0.19
Kind 11	15	50	7.0 (1.29)	4.0 (1.36)	0.17	-0.05	-0.22
Kind 12	19	36	6.0 (1.93)	5.0 (1.65)	-0.07	0.02	0.09
Kind 13	16	50	7.0 (0.96)	5.0 (1.48)	0.15	-0.03	-0.18
Kind 14	21	44	4.0 (2.06)	4.0 (1.85)	0.06	-0.01	-0.06
Kind 15	22	36	0.5 (2.59)	7.0 (3.00)	-0.13	0.05	0.18
Kind 16	14	44	3.5 (2.42)	3.5 (3.10)	0.07	0.03	-0.04
Kind 17	22	36	6.5 (2.05)	6.0 (2.07)	0.09	-0.03	-0.12
Kind 18	22	37	3.0 (2.40)	3.0 (2.23)	0.13	-0.02	-0.15
Kind 19	22	40	5.0 (2.78)	5.0 (2.24)	0.14	-0.03	-0.17
Kind 20	14	38	5.0 (1.73)	6.0 (1.97)	-0.07	-0.02	0.05
Kind 21	14	45	3.0 (2.62)	5.0 (1.89)	-0.09	-0.07	0.02
Kind 22	11	39	8.0 (0.83)	4.0 (1.33)	-0.05	-0.02	0.03
Kind 23	20	43	5.0 (2.49)	4.0 (1.98)	0.15	-0.05	-0.20
Kind 24	14	47	5.0 (1.98)	2.0 (2.57)	-0.03	0.01	0.04
Kind 25	17	50	4.0 (2.01)	3.0 (1.76)	0.22	-0.09	-0.30
Kind 26	14	38	6.0 (1.07)	5.0 (1.48)	-0.01	-0.06	-0.05
Kind 27	22	41	4.0 (1.59)	4.0 (1.45)	0.07	0.02	-0.05
M über alle Schüler*innen	17.81	41.37	5.15 (2.02)	4.37 (1.85)	0.06	-0.02	-0.08

Anmerkungen. N_A = Anzahl MZP in der Baselinephase, N_B = Anzahl MZP in der Interventionsphase, Med = Median, SD = Standardabweichung, $\Delta Trend$ = Trendunterschied zwischen den Phasen, M = Mittelwert.

werten, dass bei gut zwei Dritteln der Fälle ($n = 19$) die Datenpunkte in der Interventionsrate signifikant unterhalb des 95%-Konfidenzintervalls des Erwartungswertes der Zufallsverteilung des vorausgesagten Trends lagen. Der NAP ergibt für gut zwei Drittel der Fälle ($n = 19$) eine fragwürdige Effektivität, ein mittlerer Effekt wird bei $n = 6$ Fällen und ein großer Effekt bei $n = 2$ Fällen sichtbar. Der TAU-U ergibt für die größte Gruppe einen kleinen Effekt ($n = 12$), die zweitgrößte Gruppe bewegt sich bei einer moderaten Effektstärke ($n = 9$) und bei einem Fall ($n = 1$) ist ein großer Effekt zu konstatieren. Bei $n = 5$ können keine Effekte nachgewiesen werden.

Die Mittelwerte der Überlappingsmaße über alle Schüler*innen ($N = 27$) zeichnen ein nicht ganz einheitliches Bild: PAND (70 %), PET (72.21 %) und TAU-U (-.219) verweisen auf moderate Veränderungen in der Interventionsphase, der NAP (60.05 %) indiziert hingegen einen eher fragwürdigen Effekt.

Es folgt die inferenzstatistische Analyse anhand der Stückweisen Regressionsanalyse (Huitema & McKean, 2000), zunächst auf Ebene der Einzelfälle, anschließend als Mehrebenen-Analyse.

Auf der Ebene der Einzelfälle, zusammengefasst in Tabelle 21, wird deutlich, dass sich bei nahezu allen Fällen ($n = 26$) das nichtregelkonforme Verhalten zu Anfang der Grundrate signifikant von Null unterscheidet. In der Grundrate weisen insgesamt zwei Drittel der Verhaltensverläufe ($n = 18$) positive Trends auf, die bei $n = 7$ (min. = 0.14, max. = 0.69) zusätzlich signifikant sind. Mit Einsetzen der *Intervention* zeigen insgesamt gut zwei Drittel der Fälle ($n = 20$) sprunghafte Abfälle der Verhaltensverläufe, bei $n = 7$ (min. = -1.74, max. = -7.52) ist dieser Level-Effekt signifikant. Die Reduktion des auffälligen Verhaltens zeigt sich zudem anhand der kontinuierlichen Senkung der Verhaltensverlaufsgraphen. Dies ist bei insgesamt $n = 20$ Verläufen der Fall. Signifikante Slope-Effekte werden bei $n = 7$ (min. = -0.18, max. = -0.56) sichtbar.

Insgesamt werden bei mehr als einem Drittel der Einzelfälle ($n = 11$) signifikante Veränderungen in der Interventionsrate durch Slope- und/oder Level-Effekte sichtbar. Von diesen zeigen $n = 4$ Level-Effekte, $n = 4$ Slope-Effekte und $n = 3$ sowohl Level- als auch Slope-Effekte. Unter Letzteren weisen alle Fälle außerdem signifikante positive Trend-Effekte auf.

Tabelle 20: Überlappingsmaße hinsichtlich des Regelverhaltens (DBR-SIS) für jedes Kind und über alle Schüler*innen

Schüler*innen	PAND (%)	PET (%)	p (PETbinominal-95%CI)	NAP (%)	Tau-U
Kind 1	64.91	66.67	.03	68.92	-.28**
Kind 2	79.66	85.71	<.001	54.69	-.07
Kind 3	73.33	78.05	<.001	60.65	-.14
Kind 4	71.93	94.44	<.001	66.73	-.37***
Kind 5	100.00	100.00	<.001	100.00	-.71***
Kind 6	79.10	100.00	<.001	71.35	-.13
Kind 7	86.36	92.59	<.001	79.96	-.25*
Kind 8	43.86	97.30	<.001	35.41	.19
Kind 9	66.15	8.33	1.00	59.68	-.14
Kind 10	68.97	95.00	<.001	50.97	-.02
Kind 11	96.92	100.00	<.001	90.00	-.57***
Kind 12	67.27	22.22	1.00	62.50	-.06
Kind 13	84.85	92.00	<.001	80.94	-.29**
Kind 14	56.92	72.73	<.001	36.15	.08
Kind 15	31.03	2.78	1.00	13.57	.45***
Kind 16	62.07	68.18	.01	51.87	.04
Kind 17	62.07	83.33	<.001	52.15	-.07
Kind 18	59.32	89.19	<.001	53.07	-.09
Kind 19	58.06	87.50	<.001	46.99	-.04
Kind 20	69.23	18.42	1.00	53.01	-.05
Kind 21	66.10	0.00	<.001	35.08	-.13
Kind 22	92.00	92.31	<.001	96.27	-.30**
Kind 23	68.25	97.67	<.001	61.80	-.25**
Kind 24	70.49	63.83	.04	65.20	-.09
Kind 25	70.15	96.00	<.001	59.53	-.43***
Kind 26	73.08	55.26	.31	63.16	-.36***
Kind 27	71.43	90.24	<.001	51.61	0
M über alle Schüler*innen	70.00	72.21	<.001	60.05	-.219***

Anmerkungen. PAND = Percentage of all nonoverlapping data, PET = Percentage of data points exceeding the trend, p = p -Wert des Binominaltests des PET gegen den Zufall, NAP = Nonoverlap of all pairs, Tau-U = Parker's Tau-U (A vs. B + Trend B – Trend A), M = Mittelwert. ^

* ($p \leq .05$); ** ($p \leq .01$); *** ($p \leq .001$).

Tabelle 21: Stückweise Regressionsanalyse hinsichtlich des Regelverhaltens (DBR-SIS) für jedes Kind

Schüler*innen	Intercept (SE)	Trend (SE)	Level (SE)	Slope (SE)
Kind 1	4.90*** (1.10)	-0.03 (0.08)	0.13 (1.29)	-0.03 (0.09)
Kind 2	4.46*** (0.45)	0.03 (0.04)	-0.22 (0.50)	-0.03 (0.05)
Kind 3	5.61*** (1.06)	0.04 (0.08)	-0.83 (1.22)	-0.05 (0.08)
Kind 4	5.08*** (0.89)	0.14* (0.07)	-0.24 (1.10)	-0.24** (0.07)
Kind 5	8.24*** (0.25)	-0.01 (0.03)	-2.46*** (0.29)	-0.02 (0.03)
Kind 6	-1.82** (0.57)	0.58*** (0.06)	-7.52*** (0.61)	-0.56*** (0.06)
Kind 7	7.15*** (0.51)	0.01 (0.05)	-1.74** (0.62)	0.03 (0.06)
Kind 8	3.81** (1.36)	0.20 (0.10)	-2.01 (1.59)	-0.14 (0.11)
Kind 9	3.46*** (0.89)	-0.09 (0.09)	0.43 (0.96)	0.07 (0.09)
Kind 10	2.96* (1.18)	0.18* (0.09)	-1.85 (1.21)	-0.19 (0.09)
Kind 11	5.16*** (0.56)	0.17** (0.06)	-2.53*** (0.63)	-0.22*** (0.06)
Kind 12	6.27*** (0.83)	-0.07 (0.06)	-0.43 (0.97)	0.09 (0.07)
Kind 13	5.06*** (0.67)	0.15* (0.07)	-1.91* (0.75)	-0.18* (0.07)
Kind 14	3.42*** (0.85)	0.06 (0.07)	0.59 (1.00)	-0.06 (0.07)
Kind 15	3.12* (1.23)	-0.13 (0.09)	4.88** (1.45)	0.18 (0.10)
Kind 16	3.30 (2.47)	0.07 (0.18)	-1.16 (1.84)	-0.04 (0.19)
Kind 17	5.19*** (0.90)	0.09 (0.07)	-0.48 (1.07)	-0.12 (0.07)
Kind 18	2.42* (1.00)	0.13 (0.08)	-1.32 (1.20)	-0.15 (0.08)
Kind 19	3.34** (1.06)	0.14 (0.08)	-0.47 (1.26)	-0.17 (0.09)
Kind 20	6.63*** (1.62)	-0.07 (0.11)	0.55 (1.14)	0.05 (0.11)
Kind 21	3.34** (1.11)	0.11 (0.08)	0.81 (1.13)	-0.18* (0.09)
Kind 22	8.34*** (0.70)	-0.05 (0.07)	-2.71** (0.78)	0.03 (0.07)
Kind 23	3.16** (0.93)	0.15* (0.07)	-1.55 (1.09)	-0.20* (0.08)
Kind 24	4.41** (1.36)	-0.03 (0.14)	-1.15 (1.53)	0.04 (0.14)
Kind 25	2.12** (0.69)	0.69** (0.07)	-0.18 (0.74)	-0.30*** (0.07)
Kind 26	6.07*** (1.01)	-0.01 (0.07)	0.96 (0.71)	-0.05 (0.07)
Kind 27	2.90*** (0.66)	0.07 (0.05)	-1.07 (0.79)	-0.05 (0.05)

Anmerkungen. Intercept = Konstante, Trend = Trend in der Grundrate, Level = Veränderung des Levels zur Interventionsrate, Slope = Steigung in der Interventionsrate, SE = Standardfehler.

* ($p \leq .05$); ** ($p \leq .01$); *** ($p \leq .001$).

Tabelle 22: Mehrebenenanalyse hinsichtlich des Regelverhaltens (DBR-SIS) über alle Schüler*innen

Parameter	<i>B</i>	<i>SE</i>	<i>df</i>	<i>t</i>	<i>p</i>
Intercept	4.07	0.29	1568	13.99	<.001
Trend	0.09	0.02	1568	5.74	<.001
Level	-0.90	0.22	1568	-4.09	<.001
Slope	-0.11	0.02	1568	-6.73	<.001

Anmerkungen. Intercept = Konstante, Trend = Trend in der Grundrate, Level = Veränderung des Levels zur Interventionsrate, Slope = Steigung in der Interventionsrate, *SE* = Standardfehler, *df* = Freiheitsgrade, *t* = t-Wert, *p* = Signifikanzwert. Schätzungsmethode für Slope: B&L-B.

Die Mehrebenenanalyse der aggregierten 27 Einzelfälle als ein Einzelfall mit Messpunkten (Ebene 1) geclustert zu Schüler*innen (Ebene 2) zeigt signifikante Trend-, Level- und Slope-Effekte (Tabelle 22). Das Intercept verweist mit einem Wert von 4.07 darauf, dass sich über alle Fälle das nichtregelkonforme Verhalten zu Anfang der Grundrate signifikant von Null unterscheidet. Der Trend-Effekt zeigt, dass über alle Fälle hinweg der Trend in der Grundrate um 0.09 Skaleneinheiten ansteigt, das bedeutet, dass sich das Regelverhalten aller Schüler*innen vor der *Intervention* tendenziell verschlechterte. Der Level-Effekt über alle Fälle weist auf eine Verhaltensverbesserung von -0.9 Skaleneinheiten hin, die unmittelbar nach Beginn der *Intervention* auftrat. Genau wie den DBR-MIS entspricht dies einer bemerkenswerten sprunghaften Verhaltensverbesserung um etwa 10 %. Der Slope-Effekt indiziert, dass es während der Interventionsphase zu einer kontinuierlichen Reduzierung des nichtregelkonformen Verhaltens über alle 27 Schüler*innen um -0.11 Skaleneinheiten pro Messung kam. Damit spiegelt auch der Slope-Effekt das Ergebnis der MIS wider: an jedem Ratingtag verbesserte sich das nichtregelkonforme Verhalten der Schüler*innen um gut 1 %.

*Interaktionseffekte zwischen 84 Handlungsmöglichkeiten in Abhängigkeit zu deren Erklärungs- und Handlungsansatz und den Schüler*innenoutcomes*

Da im Interventionszeitraum lerntheoretische, kognitionspsychologische und schüler*innenzentrierte Methoden Anwendung fanden, werden im Folgenden die Verläufe der spezifischen Verhaltensauffälligkeiten (DBR-MIS) und des Regelverhaltens (DBR-SIS) der Schüler*innen gruppiert nach der Ausrichtung der *Intervention* analysiert. Unterschieden wird hierbei zwischen dem Einsatz ausschließlich lerntheoretischer Handlungsmöglichkeiten (*n* = 15) und der Kombination aus lerntheoretischem, kognitionspsychologischem und/oder schüler*innenzentriertem Erklärungs- und Handlungsansatz (*n* = 12).

Tabelle 23: Mehrebenenanalyse hinsichtlich der spezifischen Verhaltensauffälligkeiten (DBR-MIS Summenwerte) mit lerntheoretischer Intervention über $n = 15$ Schüler*innen

Parameter	<i>B</i>	<i>SE</i>	<i>df</i>	<i>t</i>	<i>p</i>
Intercept	12.28	1.40	890	8.75	<.001
Trend	0.23	0.06	890	4.05	<.001
Level	-3.86	0.73	890	-5.32	<.001
Slope	-0.30	0.06	890	-5.19	<.001

Anmerkungen. Intercept = Konstante, Trend = Trend in der Grundrate, Level = Veränderung des Levels zur Interventionsrate, Slope = Steigung in der Interventionsrate, *SE* = Standardfehler, *df* = Freiheitsgrade, *t* = t-Wert, *p* = Signifikanzwert. Schätzungsmethode für Slope: B&L-B.

Begonnen wird mit den Ergebnissen in Bezug auf die spezifischen Verhaltensauffälligkeiten (DBR-MIS) der Schüler*innen. Die Mehrebenenanalyse der aggregierten 15 Einzelfälle mit lerntheoretisch ausgerichteter Intervention, dargestellt in Tabelle 23, zeigt für alle *B*-Parameter signifikante Effekte. Das Intercept verweist mit einem Wert von 12.28 darauf, dass sich über alle Fälle die spezifischen Verhaltensauffälligkeiten zum Beginn der Grundrate signifikant von Null unterscheiden. Der Trend-Effekt zeigt, dass der Trend in der Grundrate um 0.23 Skalenpunkte ansteigt und sich die spezifischen Verhaltensauffälligkeiten vor Einsetzen der Intervention somit kontinuierlich intensivierten. Der Level-Effekt verweist auf eine Verhaltensverbesserung von -3.86 Skalenpunkten unmittelbar nach Einsetzen der *Intervention* und der Slope-Effekt indiziert eine kontinuierliche Reduktion der spezifischen Verhaltensauffälligkeiten um -0.30 Skalenpunkte pro Messung. Die Untergruppe mit lerntheoretisch ausgerichteter *Intervention* zeigt somit sowohl unmittelbare als auch kontinuierliche Interventionseffekte. Die *B*-Werte liegen dabei leicht über denen der Gesamtgruppe ($N = 27$).

Die Mehrebenenanalyse der aggregierten 12 Einzelfälle mit kombinierter *Intervention* (lerntheoretisch/kognitionspsychologisch/schüler*innenzentriert), dargestellt in Tabelle 24, ergibt signifikante Trend- und Slope-Effekte. Das Intercept verweist mit dem bisher höchsten Wert von 14.16 darauf, dass sich über alle Fälle die spezifischen Verhaltensauffälligkeiten zu Beginn der Grundrate signifikant von Null unterscheiden. Der Trend-Effekt zeigt, dass über alle 12 Fälle hinweg der Trend in der Grundrate um 0.16 Skalenpunkte ansteigt und sich die spezifischen Verhaltensauffälligkeiten somit auch hier kontinuierlich steigern, wenn auch weniger intensiv als in der Gruppe mit lerntheoretischer *Intervention*. Der Slope-Effekt indiziert eine kontinuierliche Reduktion der spezifischen Verhaltensauffälligkeiten um -0.23 Skalenpunkte pro Messung in der Interventionsrate. Der Level-Effekt war mit einem Wert von -0.64 nicht signifikant.

Tabelle 24: Mehrebenenanalyse hinsichtlich der spezifischen Verhaltensauffälligkeiten (DBR-MIS Summenwerte) bei kombinierter Intervention (= kognitionspsychologisch/schüler*innenzentriert/lerntheoretisch) über $n = 12$ Schüler*innen

Parameter	<i>B</i>	<i>SE</i>	<i>df</i>	<i>t</i>	<i>p</i>
Intercept	14.16	1.43	675	9.87	<.001
Trend	0.16	0.05	675	2.92	<.01
Level	-0.64	0.78	675	-0.81	.41
Slope	-0.23	0.06	675	-4.09	<.001

Anmerkungen. Intercept = Konstante, Trend = Trend in der Grundrate, Level = Veränderung des Levels zur Interventionsrate, Slope = Steigung in der Interventionsrate, *SE* = Standardfehler, *df* = Freiheitsgrade, *t* = t-Wert, *p* = Signifikanzwert. Schätzungsmethode für Slope: B&L-B.

Die Untergruppe mit kombinierter *Intervention* zeigt somit in erster Linie kontinuierliche Interventionseffekte. Mit Ausnahme des Intercepts, liegen die *B*-Werte insgesamt unter denen der Gruppe mit lerntheoretischer *Intervention* und denen der Gesamtgruppe ($N = 27$).

Es folgen die Ergebnisse bezüglich des Regelverhaltens der Schüler*innen (DBR-SIS). Die Mehrebenenanalyse der aggregierten 15 Einzelfälle mit lerntheoretisch ausgerichteter *Intervention*, dargestellt in Tabelle 25, zeigt für alle *B*-Parameter signifikante Effekte. Das Intercept verweist darauf, dass sich über alle Fälle das Regelverhalten zum Beginn der Grundrate mit einem Wert von 3.96 signifikant von Null unterscheidet. Der Trend-Effekt impliziert, dass über alle Fälle hinweg der Trend in der Grundrate um 0.08 Skalenpunkte ansteigt und sich das nicht-regelkonforme Verhalten vor Einsetzen der *Intervention* somit kontinuierlich intensivierte. Der Level-Effekt verweist auf eine Verhaltensverbesserung von -1.11 Skalenpunkten unmittelbar nach Einsetzen der *Intervention* und der Slope-Effekt indiziert eine kontinuierliche Reduktion des nicht-regelkonformen Verhaltens während der *Intervention* um -0.30 Skalenpunkte pro Messung. Die Untergruppe mit lerntheoretisch ausgerichteter *Intervention* zeigt somit auch im Bereich Regelverhalten sowohl unmittelbare als auch kontinuierliche Interventionseffekte. Der Level-Effekt (-1.11) übersteigt darüber hinaus deutlich den Wert der Gesamtgruppe ($N = 27$; Level-Effekt = 0.9).

Die Mehrebenenanalyse der aggregierten 12 Einzelfälle mit kombinierter *Intervention* (lerntheoretisch/kognitionspsychologisch/schüler*innenzentriert), dargestellt in Tabelle 26, ergibt signifikante Trend- und Slope-Effekte. Das Intercept verweist mit dem bisher höchsten Wert von 4.36 darauf, dass sich über alle Fälle das nicht-regelkonforme Verhalten zum Beginn der Grundrate signifikant von Null unterscheidet. Der Trend-Effekt zeigt, dass über alle 12 Fälle hinweg der Trend in der Grundrate um 0.09 Skalenpunkte ansteigt und sich das nicht-regelkonforme Verhalten vor Einsetzen der Intervention auch

Tabelle 25: Mehrebenenanalyse hinsichtlich des Regelverhaltens (DBR-SIS Summenwerte) mit lerntheoretischer Intervention über $n = 15$ Schüler*innen

Parameter	<i>B</i>	<i>SE</i>	<i>df</i>	<i>t</i>	<i>p</i>
Intercept	3.96	0.41	890	9.72	<.001
Trend	0.08	0.02	890	3.63	<.001
Level	-1.11	0.30	890	-3.69	<.001
Slope	-0.10	0.02	890	-4.28	<.001

Anmerkungen. Intercept = Konstante, Trend = Trend in der Grundrate, Level = Veränderung des Levels zur Interventionsrate, Slope = Steigung in der Interventionsrate, *SE* = Standardfehler, *df* = Freiheitsgrade, *t* = t-Wert, *p* = Signifikanzwert.

Schätzungsmethode für Slope: B&L-B.

hier kontinuierlich intensivierte. Der Slope-Effekt indiziert eine kontinuierliche Reduktion des nicht-regelkonformen Verhaltens um -0.11 Skalenpunkte pro Messung. Der Level-Effekt wird mit einem Wert von -0.56 nicht signifikant. Die Untergruppe mit kombinierter *Intervention* zeigt bezüglich des Regelverhaltens somit in erster Linie kontinuierliche Interventionseffekte.

Tabelle 26: Mehrebenenanalyse hinsichtlich des Regelverhaltens (DBR-SIS Summenwerte) bei kombinierter Intervention (= kognitionspsychologisch/schüler*innen-zentriert/lerntheoretisch) über $n = 12$ Schüler*innen

Parameter	<i>B</i>	<i>SE</i>	<i>df</i>	<i>t</i>	<i>p</i>
Intercept	4.36	0.38	675	11.47	<.001
Trend	0.09	0.02	675	4.09	<.001
Level	-0.56	0.32	675	-1.72	.08
Slope	-0.11	0.02	675	-4.82	<.001

Anmerkungen. Intercept = Konstante, Trend = Trend in der Grundrate, Level = Veränderung des Levels zur Interventionsrate, Slope = Steigung in der Interventionsrate, *SE* = Standardfehler, *df* = Freiheitsgrade, *t* = t-Wert, *p* = Signifikanzwert.

Schätzungsmethode für Slope: B&L-B.

Verhaltensauffälligkeiten nach dem Treatment

Neben dem verlaufsdiagnostischen Monitoring mittels DBR wurden die Verhaltensauffälligkeiten auch statusdiagnostisch mithilfe des SDQ (Goodman, 2005) zu drei MZP erfasst. Auf diese Weise sollte überprüft werden, ob sich die Verhaltensauffälligkeiten nach der *Intervention* reduzieren und ob der Effekt vier Monate später noch anhält. Von 27 Schüler*innen konnten bei der

Analyse 21 Kinder berücksichtigt werden. Bei einem Kind lag ein unvollständiger Datensatz zu MZP 3 vor, der deshalb nicht auswertbar war, drei Lehrkräfte waren zu MZP 3 nicht (mehr) an der Schule (Gründe: Referendariat/Schwangerschaft/privat) und zwei Schüler*innen hatten zu MZP 3 die Schule bzw. Klasse gewechselt und ein Vergleich der Lehrkräfteeinschätzungen zu drei MZP war damit nicht möglich.

Die Mittelwerte des Gesamtproblemwerts des SDQ betragen zum MZP 1 $M = 20.71$ ($SD = 5.31$), zum MZP 2 $M = 16.38$ ($SD = 5.87$) und zum MZP 3 $M = 17.33$ ($SD = 4.14$). Zu allen drei Erhebungszeitpunkten zeigen sich damit auffällige Gesamtproblemwerte, wobei eine Reduzierung der Verhaltensauffälligkeiten über die Zeit sichtbar wird und bei MZP 2 und 3 eine Tendenz hin zu einem grenzwertigen Gesamtproblemwert zu verzeichnen ist. Betrachtet man die Subskalen des Gesamtproblemwerts, wird ersichtlich, dass in jedem Bereich eine Reduzierung der Verhaltensauffälligkeiten über die Zeit stattgefunden hat. So betragen die Mittelwerte der Subskala Emotionale Probleme (EP) zu MZP 1 $M = 3.41$ ($SD = 2.94$), zum MZP 2 $M = 2.64$ ($SD = 2.59$) und zum MZP 3 $M = 2.68$ ($SD = 1.89$) und zeigen damit eine leichte Reduktion der Werte. Bei der Subskala Verhaltensprobleme (VP) ist die Veränderung über die Zeit ein wenig größer: Hier betragen die Mittelwerte zum MZP 1 $M = 5.54$ ($SD = 2.06$), zum MZP 2 $M = 4.08$ ($SD = 2.23$) und zum MZP 3 $M = 4.04$ ($SD = 1.97$). Im Bereich Hyperaktivität (H) betragen die Mittelwerte zum MZP 1 $M = 8.14$ ($SD = 1.56$), zum MZP 2 $M = 7.14$ ($SD = 2.44$) und zum MZP 3 $M = 7.14$ ($SD = 2.29$). In der Subskala Verhaltensprobleme mit Gleichaltrigen (VPG) betragen die Mittelwerte zum MZP 1 $M = 3.91$ ($SD = 2.25$), zum MZP 2 $M = 2.95$ ($SD = 1.99$) und zum MZP 3 $M = 2.95$ ($SD = 2.13$).

Die Signifikanzprüfung der Mittelwertunterschiede erfolgte mit der einfaktoriellen Varianzanalyse (ANOVA), nachdem geprüft worden war, ob Sphärizität gegeben ist ($p = .92$). Der Haupteffekt zeigt, dass die Mittelwerte sich mit der Zeit signifikant verändert haben, $F(2, 40) = 9.563$, $p < .001$, partielles $\eta^2 = .323$. Es handelt es sich hierbei um einen großen Effekt. Ein Bonferroni-korrigierter Post-hoc-Test zwischen den Gruppen zeigte einen signifikanten Unterschied zwischen den Mittelwerten von MZP 1 und 2 ($p = .001$) sowie MZP 1 und 3 ($p = .006$). Die Mittelwerte von MZP 2 und 3 unterschieden sich nicht signifikant voneinander ($p = .50$). Anhand des SDQ-Gesamtproblemwertes kann somit eine nachhaltige signifikante Reduktion der Verhaltensauffälligkeiten nach der *Intervention* konstatiert werden. Mit genauerem Blick auf die Subskalen zeigt sich, dass dies insbesondere für die Bereiche Verhaltensprobleme ($p = .001$ und $p < .001$) und Verhaltensprobleme mit Gleichaltrigen ($p = .006$ und $p = .02$) gilt. Die Subskala Hyperaktivität zeigt ebenfalls eine signifikante Veränderung zwischen MZP 1 zu MZP 2 ($p = .017$) und MZP 1 und 3 ($p = .019$).

Tabelle 27: Deskriptive Statistik und Ergebnisse des Mittelwertvergleichs über drei MZP hinsichtlich der Verhaltensprobleme der Schüler*innen, erhoben mit dem Strengths and Difficulties Questionnaire – Lehrkraftversion (SDQ-D; Goodman, 2005)

MZP	N	M (SD)	95% KI	Gruppen	M _{Diff}	SE _{Diff}	p (1-seitig)
SDQ Gesamtproblemwert							
MZP 1	21	20.71 (5.31)	[18.30, 23.13]	MZP 2	4.33	1.09	.001**
				MZP 3	3.38	1.02	.006*
MZP 2	21	16.38 (5.87)	[13.71, 19.05]	MZP 1	-4.33	1.09	.001**
				MZP 3	-0.95	1.01	.50
MZP 3	21	17.33 (4.14)	[15.45, 19.22]	MZP 1	-3.38	1.02	.006*
				MZP 2	0.95	1.01	.50
SDQ Subskala Emotionale Probleme							
MZP 1	21	3.41 (2.94)	[2.11, 4.71]	MZP 2	0.77	0.52	.08
				MZP 3	0.73	0.46	.07
MZP 2	21	2.64 (2.59)	[1.49, 3.79]	MZP 1	-0.77	0.52	.08
				MZP 3	-0.05	0.43	.46
MZP 3	21	2.68 (1.89)	[1.85, 3.52]	MZP 1	-0.73	0.46	.07
				MZP 2	0.05	0.43	.46
SDQ Subskala Verhaltensprobleme							
MZP 1	21	5.54 (2.06)	[4.67, 6.41]	MZP 2	1.46	0.58	.01**
				MZP 3	1.50	0.36	<.001***
MZP 2	21	4.08 (2.23)	[3.14, 5.02]	MZP 1	-1.46	0.58	.01**
				MZP 3	0.04	0.52	.47
MZP 3	21	4.04 (1.97)	[3.21, 4.87]	MZP 1	-1.50	0.36	<.001***
				MZP 2	-0.04	0.52	.47
SDQ Subskala Hyperaktivität							
MZP 1	21	8.14 (1.56)	[7.43, 8.85]	MZP 2	1.00	0.44	.017*
				MZP 3	1.00	0.45	.019*
MZP 2	21	7.14 (2.44)	[6.04, 8.25]	MZP 1	-1.00	0.44	.017*
				MZP 3	0	0.35	.50
MZP 3	21	7.14 (2.29)	[6.10, 8.18]	MZP 1	-1.00	0.45	.019*
				MZP 2	0	0.35	.50
SDQ Subskala Verhaltensprobleme mit Gleichaltrigen							
MZP 1	21	3.91 (2.25)	[2.91, 4.90]	MZP 2	0.96	0.35	.006**
				MZP 3	0.96	0.43	.02*
MZP 2	21	2.95 (1.99)	[2.07, 3.84]	MZP 1	-0.96	0.35	.006**
				MZP 3	0	0.33	.50
MZP 3	21	2.95 (2.13)	[2.01, 3.90]	MZP 1	-0.96	0.43	.02*
				MZP 2	0	0.33	.50

Anmerkungen. N = Anzahl, M = Mittelwert, KI = Konfidenzintervall, SD = Standardabweichung; min. = Minimum, max. = Maximum, Diff = Differenz, SE = Standardfehler, p = Signifikanzwert.

* ($p \leq .05$); ** ($p \leq .01$); *** ($p \leq .001$).

Prosoziales Verhalten nach dem Treatment

Das prosoziale Verhalten wurde mit der Subskala des SDQ gemessen. Äquivalent zu den Verhaltensauffälligkeiten sollte anhand dreier sich über ein Schuljahr verteilender MZP die Entwicklung des prosozialen Verhaltens untersucht werden. Anders als beim Gesamtproblemwert des SDQ konnten $N = 22$ Schüler*innen berücksichtigt werden, da der eine unvollständige Datensatz (s. Auswertung Gesamtproblemwert) nicht die Skala für prosoziales Verhalten betraf.

Der Mittelwert des prosozialen Verhaltens steigt vom Prätest ($M = 3.27$, $SD = 1.86$) zum Posttest ($M = 4.77$, $SD = 2.62$) an und wechselt damit vom auffälligen in den grenzwertigen Bereich. Zum Follow-up-Test fällt der Wert wieder leicht ab, bleibt dabei jedoch oberhalb der Schwelle zum auffälligen Bereich ($M = 4.68$, $SD = 2.21$). Die Signifikanzprüfung der Mittelwertunterschiede erfolgte mit einer ANOVA, nachdem geprüft wurde, ob Sphärizität gegeben ist ($p = .95$). Der Haupteffekt zeigt, dass sich die Mittelwerte prosozialen Verhaltens mit der Zeit signifikant verändern, $F(2, 42) = 9.97$, $p < .001$, partielles $\eta^2 = .322$, d.h. die Effektstärke verweist auf einen großen Effekt. Ein Bonferroni-korrigierter Post-hoc-Test zwischen den Gruppen zeigt einen signifikanten Unterschied zwischen den Mittelwerten von MZP 1 und 2 ($p = .002$) und MZP 1 und 3 ($p = .002$). Die Mittelwerte von MZP 2 und 3 unterscheiden sich nicht signifikant voneinander ($p = .50$). Das prosoziale Verhalten der Schüler*innen war somit nach Beendigung des *Treatments* stärker ausgeprägt als vor dessen Beginn. Der grenzwertige Bereich wurde auch vier Monate nach Beendigung des *Treatments* nicht mehr unterschritten, somit hielt der Effekt an.

Tabelle 28: Deskriptive Statistik und Ergebnisse des Mittelwertvergleichs über drei MZP hinsichtlich des prosozialen Verhaltens der Schüler*innen, erhoben mit dem Strengths and Difficulties Questionnaire – Lehrkraftversion (SDQ-D; Goodman, 2005)

MZP	N	M (SD)	Gruppen	M_{Diff}	SE_{Diff}	95 % KI _{Diff}	p (1-seitig)
MZP 1	22	3.27 (1.86)	MZP 2	-1.50	0.39	[-2.51, -0.49]	.002**
			MZP 3	-1.41	0.36	[-2.36, -.46]	.002**
MZP 2	22	4.77 (2.62)	MZP 1	1.50	0.39	[0.49, 2.51]	.002**
			MZP 3	0.09	0.38	[-0.89, 1.07]	.50
MZP 3	22	4.68 (2.12)	MZP 1	0.36	0.36	[0.46, 2.36]	.002**
			MZP 2	0.38	0.38	[-1.07, -1.07]	.50

Anmerkungen. N = Anzahl, M = Mittelwert, SD = Standardabweichung, min. = Minimum, max. = Maximum, Diff = Differenz, SE = Standardfehler, KI = Konfidenzintervall, p = Signifikanzwert.

* ($p \leq .05$); ** ($p \leq .01$); *** ($p \leq .001$).

5.1.2 Lehrkrafthandeln

Veränderung des Lehrkrafthandelns durch eine angemessene Einsatzhäufigkeit der 84 Handlungsmöglichkeiten

Ziel der Professionalisierungsmaßnahme war es unter anderem, das unterrichtliche Lehrkrafthandeln zu verändern. Während der zweiten Fortbildung entschieden sich die Lehrkräfte theoriegeleitet für bis zu vier Handlungsmöglichkeiten, die sie während der Interventionsphase einsetzen würden. In Abbildung 18 sind alle auf diese Weise ausgewählten Handlungsmöglichkeiten,

Eingesetzte Handlungsmöglichkeiten (Anzahl) geordnet nach deren Erklärungs- und Handlungsansatz
<i>Lerntheoretisch</i> – Verstärkung positiven Verhaltens (9) – Token-Programm (8) – Signale einsetzen (u.a. Ampelsystem) (4) – Verhaltensverträge schließen (4) – klare Instruktionen (3) – absehbare Risikosituationen entschärfen (2) – Interesse an positiven Aktivitäten (2) – geplantes Ignorieren (2) insgesamt: 34
<i>Kognitionspsychologisch</i> – Entspannungsübungen (2) – reflektierendes Krisengespräch (1) – die „Stopp-Technik“ (1) – individuelle Fortschritte anerkennen (1) – klare Strukturen schaffen (1) – Arbeitsbesprechung im Morgenkreis (1) – Erkennen von Ängsten im Schulalltag (1) – Wochenplan (1) insgesamt: 9
<i>Schüler*innenzentriert</i> – Brücken bauen zwischen Konfliktparteien (3) – eine konkrete, realistische Bitte formulieren (1) – Wie geht's Dir? (1) – Talenten Raum geben (1) – Senden von Ich-Botschaften (1) insgesamt: 7

Abbildung 18: Auflistung der während der Intervention eingesetzten Handlungsmöglichkeiten nach Erklärungs- und Handlungsansatz sowie die Abstufung der Einsatzhäufigkeit (eigene Darstellung)

Anmerkungen. HM = 84 Handlungsmöglichkeiten (Hartke et al., 2018).

geordnet nach deren Erklärungs- und Handlungsansatz, aufgelistet. Die Übersicht zeigt, dass die Lehrkräfte insbesondere Handlungsmöglichkeiten mit lerntheoretischem Erklärungs- und Handlungsansatz für die *Intervention* auswählten (34x), besonders beliebt waren hier die Methoden Verstärkung positiven Verhaltens (9x) und Tokenprogramme (8x). Kognitionspsychologisch orientierte (9x) und schüler*innenzentrierte (7x) Handlungsmöglichkeiten wurden hingegen wesentlich seltener ausgesucht.

Die Analyse der Einsatzhäufigkeit der 84 Handlungsmöglichkeiten erfolgt deskriptiv (Abbildung 19). Bei dem Großteil der Schüler*innen ($n = 11$) war eine sehr hohe (80-100 %) Einsatzfrequenz zu verzeichnen, bei $n = 5$ Schüler*innen erfolgte der Einsatz oft (60-80 %), bei $n = 7$ Schüler*innen wurden die Handlungsmöglichkeiten gelegentlich (40-60 %) eingesetzt und bei wenigen Schüler*innen ($n = 4$) wurde eher selten (20-40 %) von der neuen Methodik Gebrauch gemacht. Werte < 20 % kamen nicht vor. Über die gesamte Schüler*innengruppe ($N = 27$) gesehen wurden die 84 Handlungsmöglichkeiten im Durchschnitt an 67 % (min. = 26 %, max. = 100 %) der Ratingtage eingesetzt und kamen somit oft zum Einsatz.

Schüler*in	Handlungsmöglichkeit (LT/KP/SZ)	Einsatz der HM/ Ratings	Einsatzhäufigkeit HM
Kind 1	klare Instruktionen (LT) Verstärkung positiven Verhaltens (LT)	36/36	100 % (sehr oft)
Kind 2	Senden von Ich-Botschaften (SZ)	36/42	86 % (sehr oft)
Kind 3	Token-Programm (LT)	41/41	100 % (sehr oft)
Kind 4	Token-Programm (LT) Entspannung (KP)	24/36	67 % (oft)
Kind 5	Token-Programm (LT)	25/46	54 % (gelegentlich)
Kind 6	Verhaltensverträge schließen (LT)	25/50	50 % (gelegentlich)
Kind 7	Token-Programm (LT) Wochenplan (KP)	27/27	100 % (sehr oft)
Kind 8	Arbeitsbesprechung im Morgenkreis (KP) Talente Raum geben (SZ)	24/37	65 % (oft)
Kind 9	Signale einsetzen (LT) Verstärkung positiven Verhaltens (LT) Verhaltensverträge schließen (LT)	24/48	50 % (gelegentlich)
Kind 10	Verhaltensverträge abschließen (LT) Verstärkung positiven Verhaltens (LT) Interesse an positiven Aktivitäten (LT)	34/40	85 % (sehr oft)
Kind 11	Token-Programm (LT)	26/50	52 % (gelegentlich)
Kind 12	klare Strukturen schaffen (KP) positive Verstärkung (LT) Verhaltensverträge abschließen (LT)	34/36	94 % (sehr oft)

Schüler*in	Handlungsmöglichkeit (LT/KP/SZ)	Einsatz der HM/ Ratings	Einsatzhäufigkeit HM
Kind 13	Ampelsystem (LT)	24/50	48 % (gelegentlich)
Kind 14	absehbare Risikosituationen entschärfen (LT) reflektierendes Krisengespräch (KP)	15/44	34 % (selten)
Kind 15	absehbare Risikosituationen entschärfen (LT)	20/36	56 % (gelegentlich)
Kind 16	geplantes Ignorieren (LT) Verstärkung positiven Verhaltens (LT)	14/44	32 % (selten)
Kind 17	Token-Programm (LT) Entspannung (KP)	26/36	72 % (oft)
Kind 18	Klare Instruktionen (LT) Verstärkung positiven Verhaltens (LT) Brücken bauen zw. Konfliktparteien (SZ)	32/37	86 % (sehr oft)
Kind 19	Stopp-Technik (KP) Wie geht's dir? (SZ)	38/40	95 % (sehr oft)
Kind 20	Individ. Fortschritte anerkennen (KP)	31/38	82 % (sehr oft)
Kind 21	Brücken bauen zw. Konfliktparteien (SZ) Erkennen v. Ängsten i. Schulalltag (KP)	11/41	27 % (selten)
Kind 22	Ampelsystem (LT)	25/39	64 % (oft)
Kind 23	Brücken bauen zw. Konfliktparteien (SZ) Konkr., realist. Bitte formulieren (SZ)	26/43	60 % (oft)
Kind 24	Interesse an positiven Aktivitäten (LT) Verstärkung positiven Verhaltens (LT) geplantes Ignorieren (LT) Token-Programm (LT)	19/47	40 % (gelegentlich)
Kind 25	klare Instruktionen (LT) Verstärkung positiven Verhaltens (LT)	13/50	26 % (selten)
Kind 26	Token-Programm (LT)	35/37	95 % (sehr oft)
Kind 27	Signale einsetzen (LT) Verstärkung positiven Verhaltens (LT)	36/42	86 % (sehr oft)
Häufigkeit (Ø)	LT: 34 KP: 9 SZ: 7	28/42	67 % (oft)

Abbildung 19: Einsatzhäufigkeit der 84 Handlungsmöglichkeiten während der Interventionsphase für jeden Einzelfall und im Durchschnitt über alle Schüler*innen (eigene Darstellung)

Anmerkungen. LT = lerntheoretisch, KP = kognitionspsychologisch, SZ = schüler*innenzentriert, HM = 84 Handlungsmöglichkeiten (Hartke et al., 2018), Einsatzhäufigkeit = Verhältnis zwischen der Anzahl der Einsatztage der Handlungsmöglichkeiten und den getätigten Ratings.

*Interaktionseffekte zwischen der Einsatzhäufigkeit der 84 Handlungsmöglichkeiten und den Schüler*innenoutcomes*

Nach der deskriptiven Analyse des Einsatzes der 84 *Handlungsmöglichkeiten* soll überprüft werden, ob sich Interaktionseffekte zwischen der Wirkung der *Intervention* und der Einsatzhäufigkeit der Handlungsmöglichkeiten feststellen lassen. Um etwaige Interaktionseffekte des Lehrkrafthandelns zu überprüfen, wurden anhand einer Mehrebenenanalyse die Level-Effekte zwischen den Schüler*innenoutcomes (DBR-SIS und DBR-MIS) und dem Kreuzprädiktor Einsatzhäufigkeit der Handlungsmöglichkeit („Handlungsmöglichkeit“) analysiert (Wilbert & Lüke, 2021). Die Einsatzhäufigkeit der Handlungsmöglichkeit ist ein Faktor mit den Werten „sehr oft“, „oft“, „gelegentlich“ und „selten“. Die Abstufung „kaum“ wurde nicht berücksichtigt, da sie im Datensatz nicht vorkam. Der Parameter „Handlungsmöglichkeit sehr oft“ spiegelt demnach beispielsweise den Level-Effekt auf das Verhalten der Schüler*innen wider, wenn die Handlungsmöglichkeiten sehr oft eingesetzt wurden. Für die Parameter „Handlungsmöglichkeit oft“, „Handlungsmöglichkeit gelegentlich“ und „Handlungsmöglichkeit selten“ gilt dies in vergleichbarer Weise.

Die mehrebenenanalytische Auswertung der spezifischen Verhaltensauffälligkeiten (DBR-MIS) aller Schüler*innen mit dem Kreuzprädiktor „Handlungsmöglichkeit“ ergibt für die Abstufung „gelegentlich“ (40-60 %), wie in Tabelle 29 dargestellt, einen Wert von -5.03 ($p < .001$). Wurden die Handlungsmöglichkeiten „oft“ (60-80 %) eingesetzt, so ergab der Interventionseffekt einen Wert von -3.11 ($p < .05$). Bei einem sehr hohen Einsatz (80-100 %) der Handlungsmöglichkeiten liegt der Interventionseffekt bei -1.38 ($p < .001$). Bei seltenem Einsatz der Handlungsmöglichkeiten (20-40 %) ergab der Interventionseffekt einen Wert von 0.2 ($p < .001$). Für alle Häufigkeitsabstufungen zeigen sich signifikante Interaktionseffekte.

Bezieht man die Werte der Mehrebenenanalyse der DBR-MIS mit ein, bleibt bezüglich der Interaktionseffekte zwischen der Einsatzhäufigkeit der Handlungsmöglichkeiten und den spezifischen Verhaltensauffälligkeiten der Schüler*innen folgendes festzuhalten: Der größte Level-Effekt zeigt sich beim gelegentlichen, gefolgt vom häufigen Einsatz der Handlungsmöglichkeiten. Beide Werte überschreiten den Level-Effekt bei den DBR-MIS (-2.45). Bei einem sehr häufigen Einsatz zeigen die Verhaltensverläufe ebenfalls einen sprunghaften Abfall von -1.38 und somit eine sofort eintretende Verhaltensverbesserung, wenn auch nicht ganz so stark, wie bei gelegentlichem und häufigem Einsatz der Handlungsmöglichkeiten. Bei seltener Anwendung ist keine positive Auswirkung auf das Schüler*innenverhalten mehr bzw. sogar eine Verschlechterung anhand von positiven Level-Effekten erkennbar. Es ist damit eindeutig von einem negativen Interaktionseffekt auszugehen. Darüber hinaus kann bei den MIS auch ein positiver Interaktions-

Tabelle 29: Interaktionseffekte zwischen der Einsatzhäufigkeit der 84 Handlungsmöglichkeiten und den spezifischen Verhaltensauffälligkeiten (DBR-MIS Summenwerte) aller Schüler*innen ($N = 27$) anhand einer Mehrebenenanalyse

Parameter	B	SE	df	t	p
Intercept	14.15	1.74	1565	8.13	<.001
Trend	0.19	0.04	1565	4.99	<.001
Level: Handlungsmöglichkeit gelegentlich	-5.03	0.70	1565	-7.20	<.001
Level: Handlungsmöglichkeit oft	-3.11	0.84	1565	2.28	<.05
Level: Handlungsmöglichkeit sehr oft	-1.38	0.70	1565	5.23	<.001
Level: Handlungsmöglichkeit selten	0.20	0.90	1565	5.83	<.001

Anmerkungen. Intercept = Konstante, Trend = Trend in der Grundrate, Level = Veränderung des Levels zur Interventionsrate, Handlungsmöglichkeit gelegentlich = Handlungsmöglichkeit wurde gelegentlich (40-60 %) eingesetzt; Handlungsmöglichkeit oft = Handlungsmöglichkeit wurde oft (60-80 %) eingesetzt, Handlungsmöglichkeit sehr oft = Handlungsmöglichkeit wurde sehr oft (80-100 %) eingesetzt, Handlungsmöglichkeit selten = Handlungsmöglichkeit wurde selten (20-40 %) eingesetzt, SE = Standardfehler, df = Freiheitsgrade, t = t-Wert, p = Signifikanzwert. Schätzungsmethode für Slope: B&L-B.

effekt angenommen werden, denn bei gelegentlichem, häufigem und sehr häufigem Einsatz der Handlungsmöglichkeiten zeigen sich signifikante negative Interaktionseffekte und damit die unmittelbare Verbesserung des Schüler*innenverhaltens nach Einsetzen der *Intervention*.

Bei der mehrebenenanalytischen Auswertung des nichtregelkonformen Verhaltens (DBR-SIS) aller Schüler*innen mit dem Kreuzprädiktor „Handlungsmöglichkeit“ ergeben sich folgende Ergebnisse: Bei gelegentlichem Einsatz der Handlungsmöglichkeiten ergibt die Analyse einen Level-Effekt von -1.07 ($p < .001$). Erfolgte der Einsatz sehr oft, so ergab sich ein Level-Effekt von -0.88, wurden die Handlungsmöglichkeiten oft eingesetzt, so ergab dies einen Wert von -1.18 und erfolgte der Einsatz eher selten, so zeigte sich ein Level-Effekt von -0.16 (Wilbert & Lüke, 2021). Die Interaktionen mit den Kreuzprädiktoren „gelegentlich“ und „selten“ ergaben signifikante Werte, die anderen Effekte waren nicht signifikant. Während der Level-Effekt bei sehr häufigem, häufigem und gelegentlichem Einsatz der Handlungsmöglichkeiten mindestens -0.88 Punkte beträgt, liegt der Effekt bei seltenem Einsatz bei lediglich -0.16.

Bezieht man den Level-Effekt der Mehrebenenanalyse des Regelverhaltens (DBR-SIS) aller Schüler*innen mit ein, so fällt folgendes für die Auswertung der Interaktionseffekte auf: Die Interaktionseffekte bei häufigem und gelegentlichem Einsatz der Handlungsmöglichkeiten liegen deutlich über und die bei sehr häufigem Einsatz ganz leicht unter dem Wert aus der Mehrebenenanalyse des Regelverhaltens (DBR-SIS) aller Schüler*innen (-0.9).

Tabelle 30: Interaktionseffekte zwischen der Einsatzhäufigkeit der 84 Handlungsmöglichkeiten und dem Regelverhalten (DBR-SIS) aller Schüler*innen ($N = 27$) anhand einer Mehrebenenanalyse

Parameter	B	SE	df	t	p
Intercept	3.70	0.45	1565	8.19	<.001
Trend	0.09	0.02	1565	5.57	<.001
Level: Handlungsmöglichkeit gelegentlich	-1.07	0.29	1565	-3.67	<.001
Level: Handlungsmöglichkeit oft	-1.18	0.35	1565	-0.32	.74
Level: Handlungsmöglichkeit sehr oft	-0.88	0.29	1565	0.65	.51
Level: Handlungsmöglichkeit selten	-0.16	0.37	1565	2.43	<.05

Anmerkungen. Intercept = Konstante, Trend = Trend in der Grundrate, Level = Veränderung des Levels zur Interventionsrate, Handlungsmöglichkeit gelegentlich = Handlungsmöglichkeit wurde gelegentlich (40-60 %) eingesetzt; Handlungsmöglichkeit oft = Handlungsmöglichkeit wurde oft (60-80 %) eingesetzt, Handlungsmöglichkeit sehr oft = Handlungsmöglichkeit wurde sehr oft (80-100 %) eingesetzt, Handlungsmöglichkeit selten = Handlungsmöglichkeit wurde selten (20-40 %) eingesetzt, SE = Standardfehler, df = Freiheitsgrade, t = t-Wert, p = Signifikanzwert. Schätzungsmethode für Slope: B&L-B.

Dies bedeutet, dass sowohl der sehr häufige, der häufige, als auch der gelegentliche Einsatz der 84 Handlungsmöglichkeiten sich positiv auf das Regelverhalten der Schüler*innen ausgewirkt haben. Erfolgte dieser nur selten, so war die Interventionswirkung eher gering (-0.16). Ein klar positiver Interaktionseffekt – erkennbar wäre dieser an einer signifikanten Interaktion zwischen Verhaltensverbesserung auf Seiten der Schüler*innen und dem häufigen Einsatz der Handlungsmöglichkeiten – kann jedoch hier ebenso nicht konstatiert werden. Stattdessen scheint es aber auch hier einen negativen Interaktionseffekt zu geben: Werden die Handlungsmöglichkeiten eher selten eingesetzt (20-40 %), zeigt sich keinerlei positive Wirkung auf Seiten der Schüler*innen.

5.1.3 Lehrkraftkognitionen

Selbstwirksamkeitsempfinden der Lehrkräfte nach der Intervention mit und ohne Erhalt einer kooperativ-systemischen Fallberatung

Anhand der deskriptiven Statistik wird bereits ersichtlich, dass der Mittelwert des Selbstwirksamkeitsempfindens zwischen dem ersten ($M = 31.11$, $SD = 2.80$) und dem zweiten MZP ($M = 32.94$, $SD = 3.21$) ansteigt. Zum dritten MZP ($M = 31.25$, $SD = 3.24$) fällt der Wert wieder ab und bewegt sich dabei leicht über dem Ausgangswert. Der Normwert der Skala zur Lehrer-Selbst-

wirksamkeit liegt bei 28.78, somit sind die Mittelwerte zu allen drei MZP als überdurchschnittlich hoch einzuordnen.

Die Signifikanzprüfung der Mittelwertunterschiede erfolgte anhand der ANOVA, nachdem geprüft wurde, ob Sphärizität gegeben ist ($p = .69$). Der Haupteffekt zeigt, dass die Mittelwerte des Selbstwirksamkeitsempfindens sich mit der Zeit signifikant verändert haben, $F(2, 30) = 3.47$, $p = .02$, partielles $\eta^2 = .19$. Es handelt sich hierbei um einen großen Effekt (Cohen, 1988). Ein Bonferroni-korrigierter Post-hoc-Test zwischen den Gruppen zeigt einen signifikanten Unterschied zwischen den Mittelwerten von MZP 1 und 2 ($p = .03$) sowie MZP 2 und 3 ($p = .02$). Die Mittelwerte von MZP 1 und 3 unterscheiden sich nicht signifikant voneinander ($p = .25$). Mittels der Ergebnisse kann somit eine signifikante Steigerung des Selbstwirksamkeitsempfindens über die gesamte Gruppe der Lehrkräfte unmittelbar nach der *Intervention* festgehalten werden, vier Monate später war der Effekt allerdings abgeklungen.

Für die Analyse des Effekts der *kooperativ-systemischen Beratung* wurde die Gruppe der Lehrkräfte aufgeteilt in diejenigen, die eine *Beratung* erhalten ($n = 9$) und diejenigen, die keine erhalten hatten ($n = 7$). Der Mittelwert der Gruppe mit *Beratung* steigt vom Prä- ($M = 31.11$, $SD = 2.80$) zum Posttest ($M = 32$, $SD = 3.50$) leicht an, zum Follow-up-Test ($M = 31.11$, $SD = 3.14$) sinkt er auf den Ausgangswert zurück. Es gibt keinen statistisch signifikanten Unterschied zwischen den drei Bedingungen, $F(2, 16) = .49$, $p = .31$.

Für die Lehrkräfte, die keine *Beratung* erhalten hatten, ergibt der Vergleich von MZP 1 ($M = 32.43$, $SD = 3.05$), MZP 2 ($M = 34.14$, $SD = 2.55$) und MZP 3 ($M = 31$, $SD = 3.11$) eine merkliche Steigung des Mittelwerts, dieser sinkt zum Follow-up-Test jedoch wieder unter das Ausgangsniveau. Der Haupteffekt zeigt, dass sich die Mittelwerte mit der Zeit signifikant verändert haben, $F(2, 30) = 13$, $p < .001$, partielles $\eta^2 = .68$. Es handelt sich hierbei um einen großen Effekt. Ein Bonferroni-korrigierter Post-hoc-Test zwischen den Gruppen zeigt einen signifikanten Unterschied zwischen den Mittelwerten von MZP 1 und 2 ($p = .02$) sowie MZP 2 und 3 ($p < .001$). Die Mittelwerte von MZP 1 und 3 unterscheiden sich nicht signifikant voneinander ($p = .06$).

Anhand der Ergebnisse wird ersichtlich, dass sich das Selbstwirksamkeitsempfinden nach der *Intervention* über die gesamte Gruppe der Lehrkräfte ($N = 16$) verbessert hat. Betrachtet man die Subgruppen, so wird deutlich, dass vor allem Lehrkräfte, die keine *Beratung* erhalten hatten, ihr Selbstwirksamkeitsempfinden deutlich steigern konnten, wohingegen bei denjenigen, die eine *Beratung* erhalten hatten, kein signifikanter Anstieg zu verzeichnen ist.

Tabelle 31: Ergebnisse der Angaben zum Selbstwirksamkeitsempfinden der Lehrkräfte vor und nach der Intervention sowie mit und ohne Beratung

MZP	N/n	M (SD)	95% KI	Gruppen	M_{Diff}	SE_{Diff}	p (1-seitig)
Selbstwirksamkeitsempfinden (gesamt)							
MZP 1	16	31.11 (2.80)	[29.61, 35.25]	MZP 2	-1.25	0.61	.03*
				MZP 3	0.44	0.65	.25
MZP 2	16	32.94 (3.21)	[31.79, 36.50]	MZP 1	1.25	0.61	.03*
				MZP 3	1.69	0.73	.02*
MZP 3	16	31.25 (3.24)	[28.13, 33.88]	MZP 1	-0.44	0.65	.25
				MZP 2	-1.69	0.73	.02*
Selbstwirksamkeitsempfinden (mit Beratung)							
MZP 1	9	31.11 (2.80)	[28.96, 33.27]	MZP 2	-.89	0.98	.20
				MZP 3	0	0.96	.50
MZP 2	9	32.00 (3.50)	[29.31, 34.69]	MZP 1	.89	0.98	.20
				MZP 3	.89	1.16	.23
MZP 3	9	31.11 (3.14)	[28.70, 33.53]	MZP 1	0	0.96	.50
				MZP 2	-.89	1.16	.23
Selbstwirksamkeitsempfinden (ohne Beratung)							
MZP 1	7	32.43 (3.05)	[29.61, 35.25]	MZP 2	-1.71	0.64	.02*
				MZP 3	1.43	0.78	.06
MZP 2	7	34.14 (2.55)	[31.79, 36.50]	MZP 1	1.71	0.64	.02*
				MZP 3	3.14	0.34	<.001***
MZP 3	7	31.00 (3.11)	[28.13, 33.88]	MZP 1	-1.43	0.78	.02*
				MZP 2	-3.14	0.34	<.001***

Anmerkungen. N/n = Anzahl, M = Mittelwert, SD = Standardabweichung, KI = Konfidenzintervall, M_{Diff} = Mittelwertsdifferenz, SE_{Diff} = Standardfehlerdifferenz, p = Signifikanzwert. * ($p \leq .05$); ** ($p \leq .01$); *** ($p \leq .001$).

*Interaktionseffekte zwischen dem Selbstwirksamkeitsempfinden der Lehrkräfte und den Schüler*innenoutcomes*

Im Folgenden werden die Interaktionseffekte zwischen dem Selbstwirksamkeitsempfinden der Lehrkräfte und den Schüler*innenoutcomes untersucht. Dazu wurden die unmittelbaren Interventionseffekte (Level-Effekte) auf das Schüler*innenverhalten (DBR-SIS und DBR-MIS) mit den Selbstwirksamkeitswerten der Lehrkräfte unmittelbar nach der *Intervention* (Posttest) als Kreuzprädiktor in einer mehrbenenanalytischen Regression modelliert. Statt $N = 27$ konnten bei der Modellierung nur $n = 26$ Einzelfälle berücksichtigt werden, da eine Lehrkraft zum Posttest die Skala zur Lehrer-Selbstwirksamkeit nicht ausfüllen konnte/wollte. Um eine differenziertere Aussage über

Tabelle 32: Interaktionseffekte zwischen dem Selbstwirksamkeitsempfinden der Lehrkräfte und den spezifischen Verhaltensauffälligkeiten (DBR-MIS) aller Schüler*innen ($n = 26$) anhand einer Mehrebenenanalyse

Parameter	<i>B</i>	<i>SE</i>	<i>df</i>	<i>t</i>	<i>p</i>
Intercept	14.75	1.44	1507	10.27	<.001
Trend	0.22	0.04	1507	5.47	<.001
Level: Selbstwirksamkeit hoch	-3.06	0.64	1507	-4.77	<.001
Level: Selbstwirksamkeit mittel	-2.79	0.69	1507	0.39	.69
Level: Selbstwirksamkeit niedrig	-1.42	0.72	1507	2.3	<.05

Anmerkungen. Intercept = Konstante, Trend = Trend in der Grundrate, Level = Veränderung des Levels zur Interventionsrate, Selbstwirksamkeit hoch = „hohes“ Selbstwirksamkeitsempfinden (Punkte: 35-37), Selbstwirksamkeit mittel = „mittleres“ Selbstwirksamkeitsempfinden (Punkte: 30-34), Selbstwirksamkeit niedrig = „niedriges“ Selbstwirksamkeitsempfinden (Punkte: 27-29), *SE* = Standardfehler, *df* = Freiheitsgrade, *t* = t-Wert, *p* = Signifikanzwert. Schätzungsmethode für Slope: B&L-B.

etwaige Interaktionseffekte zwischen dem Selbstwirksamkeitsempfinden und den Schüler*innenoutcomes treffen zu können, wurden die im Mittel im überdurchschnittlichen Bereich liegenden Posttestwerte ($M = 32.94$, $SD = 3.21$, min. = 27, max. = 37) dreifach abgestuft. Zwei Kriterien wurden bei der Abstufung angelegt: Zum einen sollte die Anzahl der Fälle ähnlich und zum anderen die Spannweite der Punkte annähernd gleich verteilt sein. Auf diese Weise wurden Werte zwischen 27-29 als „niedrig“ ($n = 7$), Werte zwischen 30-34 als „mittel“ ($n = 7$) und Werte zwischen 35-37 als „hoch“ ($n = 12$) eingestuft. Die Selbstwirksamkeit stellt damit einen dreifach gestuften Faktor dar. Der Parameter „Selbstwirksamkeit hoch“ zeigt beispielsweise die Level-Effekte auf das Verhalten der Schüler*innen auf, wenn sich die Lehrkräfte nach der *Intervention* als besonders selbstwirksam einschätzten. Für die Parameter „Selbstwirksamkeit mittel“ und „Selbstwirksamkeit niedrig“ gilt dies in vergleichbarer Weise.

Für die einzelnen Abstufungen ergibt die Mehrebenenanalyse mit dem Kreuzprädiktor „Selbstwirksamkeit“ im Bereich der spezifischen Verhaltensauffälligkeiten der Schüler*innen (DBR-MIS) folgende Ergebnisse, die in Tabelle 32 abgebildet sind: Bei „hoher“ Selbstwirksamkeit (Punkte: 35-37) ergibt die Analyse einen Level-Effekt von -3.06 ($p < .001$). Bei „mittlerem“ Selbstwirksamkeitsempfinden (Punkte: 30-34) ergibt sich ein Level-Effekt von -2.79 und bei einem „niedrigen“ Selbstwirksamkeitsempfinden (Punkte: 27-29) ergab die Analyse einen Wert von -1.42 ($p < .05$).

Die Ergebnisse zeigen ein erwartungskonformes Bild: Ein „hohes“ und „mittleres“ Selbstwirksamkeitsempfinden nach der *Intervention* geht mit einem positiven Interventionseffekt einher. Beide Interaktionseffekte übersteigen außerdem den Level-Effekt der Mehrebenenanalyse der spezifischen

Tabelle 33: Interaktionseffekte zwischen dem Selbstwirksamkeitsempfinden der Lehrkräfte und dem Regelverhalten (DBR-SIS) aller Schüler*innen ($n = 26$) anhand einer Mehrebenenanalyse

Parameter	<i>B</i>	<i>SE</i>	<i>df</i>	<i>t</i>	<i>p</i>
Intercept	4.43	0.37	1508	11.89	<.001
Trend	0.09	0.02	1508	5.72	<.001
Level: Selbstwirksamkeit hoch	-0.74	0.26	1508	-2.82	<.01
Level: Selbstwirksamkeit mittel	-1.24	0.28	1508	-1.76	.07
Level: Selbstwirksamkeit niedrig	-0.90	0.30	1508	-0.55	.58

Anmerkungen. Intercept = Konstante, Trend = Trend in der Grundrate, Level = Veränderung des Levels zur Interventionsrate, Selbstwirksamkeit hoch = „hohes“ Selbstwirksamkeitsempfinden (Punkte: 35-37), Selbstwirksamkeit mittel = „mittleres“ Selbstwirksamkeitsempfinden (Punkte: 30-34), Selbstwirksamkeit niedrig = „niedriges“ Selbstwirksamkeitsempfinden (Punkte: 27-29), *SE* = Standardfehler, *df* = Freiheitsgrade, *t* = t-Wert, *p* = Signifikanzwert. Schätzungsmethode für Slope: B&L-B.

Verhaltensauffälligkeiten aller Schüler*innen (-2.45). Ein „niedriges“ Selbstwirksamkeitsempfinden nach der *Intervention* ist hingegen eher ein Indiz für eine schwächere Interventionswirkung (-1.42). Zusammenfassend kann festgehalten werden, dass bei den DBR-MIS ein positiver Interaktionseffekt zwischen Lehrkraftkognitionen und Schüler*innenoutcomes feststellbar ist.

Als nächstes folgt die Analyse der Interaktionseffekte zwischen dem Selbstwirksamkeitsempfinden der Lehrkräfte und dem nichtregelkonformen Verhalten (DBR-SIS). Für die einzelnen Abstufungen ergibt die Mehrebenenanalyse mit dem Kreuzprädiktor „Selbstwirksamkeit“ im Bereich des Regelverhaltens der Schüler*innen (DBR-SIS) folgende Ergebnisse, die in Tabelle 33 abgebildet sind: Bei hoher Selbstwirksamkeit (Punkte: 35-37) im Posttest ergibt sich ein Level-Effekt von -0.74 ($p < .01$). Bei mittlerem Selbstwirksamkeitsempfinden (Punkte: 30-34) ergibt sich ein Level-Effekt von -1.24 und bei einem niedrigen Selbstwirksamkeitsempfinden (Punkte: 27-29) liegt der Wert bei -0.9.

Bezieht man den Level-Effekt der Mehrebenenanalyse des Regelverhaltens (DBR-SIS) aller Schüler*innen mit ein, so fällt Folgendes für die Auswertung der Interaktionseffekte auf: Die Level-Effekte bei „mittlerem“ und „niedrigem“ Selbstwirksamkeitsempfinden liegen deutlich über bzw. auf dem Wert der Mehrebenenanalyse des Regelverhaltens (DBR-SIS) aller Schüler*innen (-0.9). Interessanterweise zeigt der Interaktionseffekt bei einem „hohen“ Selbstwirksamkeitsempfinden nicht nur den niedrigsten Wert, dieser unterschreitet außerdem den Wert der Mehrebenenanalyse des Regelverhaltens (DBR-SIS) aller Schüler*innen. Somit kann bei den DBR-SIS kein (positiver) Interaktionseffekt zwischen dem Selbstwirksamkeitsempfinden der Lehrkräfte und den Schüler*innenoutcomes festgestellt werden.

5.1.4 Lehrkraftreaktionen

Zufriedenheit und Nützlichkeitseinschätzung bezüglich der gesamten Professionalisierungsmaßnahme

Die Lehrkraftreaktionen wurden bei der Abschlussveranstaltung abgefragt, die einen Monat nach dem Posttest und somit nach der Beendigung des Professionalisierungskonzepts stattfand und an der insgesamt 30 Lehrkräfte teilnahmen. Neben den 21 Lehrkräften, deren Verhaltensbeurteilungen Eingang in die Analysen finden konnten, waren neun weitere Lehrerinnen anwesend, die zwar bis zum Ende des Professionalisierungskonzepts an der Studie teilgenommen hatten, deren erhobene DBR-Datenmenge mit $< 80\%$ aber dem Ausschlusskriterium (s. Kap. 4.3.3.1) entsprach.

Die Lehrkraftreaktionen bezüglich der gesamten Professionalisierungsmaßnahme werden in Tabelle 34 zusammengefasst dargestellt. Neben der Darstellung der im Rahmen der Analyse der Antworten entwickelten Ober- und Unterkategorien, werden auch deren quantitative Häufigkeit ($= n$) berichtet.

Die Lehrkräfte haben sich während der gesamten Professionalisierungsmaßnahme gut betreut gefühlt ($M = 8.27$, $SD = 2.13$), die Einschätzung des Gesamtkonzepts fiel befriedigend aus ($M = 6.80$, $SD = 2.29$).¹⁷ Zur offenen Frage nach dem Mehrwert des Professionalisierungskonzepts („Welchen Mehrwert hatte das Konzept für Sie?“) gingen insgesamt $n = 59$ Antworten ein. Nach deren thematischer Clusterung konnten folgende sechs Kategorien aufgestellt werden: Zufriedenheitsbekundungen und Nützlichkeitseinschätzungen, Erweiterung des pädagogischen Wissens, Veränderung der Einstellungen/Überzeugungen, Veränderung des unterrichtspraktischen Handelns sowie Initiierung von Reflexionsprozessen. Die gebildeten Kategorien sollen im Folgenden genauer betrachtet werden.

Generelle Zufriedenheits- und Nützlichkeitsbekundungen ($n = 23$) wurden neben Danksagungen und positiven Attributen („hilfreich“, „lehrreich“, „toll“, und „positiv“) durch folgende Einschätzungen geäußert: „Gut einsetzbar für neue junge Kollegen, denen die Erfahrung im Umgang mit schwierigen Schülern fehlt“, „kaum mehr Aufwand“ und „Kennenlernen von Leuten, an die man sich vertrauensvoll wenden kann“. Neben diesen grundsätzlichen Einschätzungen bekundeten die Lehrkräfte ihre Zufriedenheit auch bezüglich spezifischer Elemente des Professionalisierungskonzepts und stellten deren Nützlichkeit heraus. Besonders positiv gewürdigt wurde der Erhalt und Einsatz der *84 Handlungsmöglichkeiten* ($n = 10$), vor allem der Aspekt des Haltgebens („Hilfe“, „Unterstützung“, „Stütze“, „Sicherheit“, „Begleiter“, „strukturiert“) wurde in diesem Zusammenhang oft erwähnt ($n = 5$). Auch das

17 Interpretationsmaße bei geschlossenen Fragen: 10-9 = sehr gut, 8-7 = gut, 6-5 = befriedigend, 4-3 = ausreichend, 2-0 = ungenügend.

Kennenlernen und Anwenden von Verhaltensmonitoring mittels *Direct Behavior Rating* erschien den Lehrkräften in der Gesamtrückschau als fruchtbar ($n = 4$), insbesondere die Möglichkeit von dessen Einsatz bei zukünftigen Elterngesprächen wurde positiv hervorgehoben. In jeweils $n = 2$ Antworten wurden explizit die Fortbildungen und die *Beratung* positiv gewürdigt. Die Lehrkräfte äußerten sich außerdem vermehrt positiv zur Erweiterung ihres pädagogischen Wissens ($n = 13$). So beschrieben sie, dass die Erweiterung des Wissens um neue Kenntnisse und Methoden zur Förderung von Schüler*innen mit Verhaltensschwierigkeiten ($n = 5$), das Erlernen neuer Bewertungs- und Analysevarianten durch die Entwicklung der eigenen Beobachtungs- und Beurteilungsfähigkeiten ($n = 5$) und die Sensibilisierung für die Ursachen von Verhaltensschwierigkeiten ($n = 3$) stattgefunden hätten. Des Weiteren nahmen viele Lehrkräfte bei sich eine positive Veränderung ihrer Einstellungen/Überzeugungen ($n = 10$) wahr. Die Pädagog*innen äußerten, dass sie sich durch die Professionalisierungsmaßnahme stärker mit einzelnen Schüler*innen beschäftigten ($n = 5$), etwa durch das bewusstere Beobachten und Aufbauen einer stärkeren Bindung. Sie gaben außerdem an, einen neuen Blick auf sogenanntes „Problemverhalten“ gewonnen zu haben ($n = 5$), mit mehr Verständnis für und weniger Vorurteilen gegenüber Kindern mit Verhaltensauffälligkeiten sowie die Stärkung des eigenen Selbstvertrauens und der Selbstwirksamkeit zu spüren ($n = 2$). Im Rahmen der Professionalisierungsmaßnahme hat sich laut Angabe der Lehrkräfte auch das unterrichtspraktische Handeln verändert ($n = 8$). So fand durch das Verhaltensmonitoring mittels *Direct Behavior Rating* eine regelmäßige und kontinuierliche Auseinandersetzung mit dem*r Schüler*in statt ($n = 4$), der Einsatz der 84 *Handlungsmöglichkeiten* ermöglichte eine strukturierte und kleinschrittigere Förderung ($n = 2$) und im Rahmen von Kolleg*innen- und Elterngesprächen wurde die Förderung der Schüler*innen weiter forciert ($n = 2$). Des Weiteren konnten während und nach der Professionalisierungsmaßnahme Reflexionsprozesse bei den Lehrkräften initiiert werden ($n = 5$). Die Lehrkräfte sahen sich vermehrt dazu angehalten, das eigene Verhalten fortwährend kritisch zu prüfen und gegebenenfalls anzupassen sowie sich mit den eigenen Befindlichkeiten auseinanderzusetzen.

Im Rahmen der Evaluation des Professionalisierungskonzepts hatten die Lehrkräfte auch die Möglichkeit, Anregungen für eine Weiterentwicklung des Professionalisierungskonzepts zu geben („Welche Verbesserungsvorschläge haben Sie?“). Im Rahmen der Auswertung der eingegangenen Modifikationsvorschläge ($n = 32$) konnten vier Kategorien gebildet werden. Die meisten Verbesserungsvorschläge gab es zum Einsatz des Verhaltensmonitorings mittels *Direct Behavior Rating* ($n = 13$). Viele Lehrkräfte sprachen sich dafür aus, den Beurteilungszeitraum auf mehrere Stunden bzw. den gesamten Schultag auszudehnen ($n = 7$). Einige gaben technische Hinweise zur Ver-

Tabelle 34: Gemittelte und kategorisierte Lehrkräfteeinschätzungen ($N = 30$) bezüglich der Professionalisierungsmaßnahme nach deren Beendigung

Frage/Item	Antworten (gemittelt)
1. Wie haben Sie sich betreut gefühlt?    $M = 8.27, SD = 2.13$	
2. Wie nützlich finden Sie das Gesamtkonzept?    $M = 6.80, SD = 2.29$	
Antworten (kategorisiert, zusammengefasst, quantifiziert)	
3. Welchen Mehrwert hatte das Konzept für Sie? ($n = 59$)	– Zufriedenheit und Nützlichkeit ($n = 23$): Erhalt und Einsatz der 84 HM ($n = 10$), Grundsätzliches ($n = 5$), Verhaltensmonitoring ($n = 4$), Fortbildungen ($n = 2$), Beratung ($n = 2$). – pädagogisches Wissen ($n = 13$): neue Kenntnisse, Ideen und Methoden ($n = 5$), Verhaltensmonitoring ($n = 5$), Sensibilisierung ($n = 3$). – Einstellungen/Überzeugungen ($n = 10$): stärkere Auseinandersetzung mit einzelnen Schüler*innen ($n = 5$), anderer Blick auf „Problemverhalten“ ($n = 3$), Stärkung des Vertrauens in eigene Selbstwirksamkeit ($n = 2$). – unterrichtspraktisches Handeln ($n = 8$): Verhaltensmonitoring ($n = 4$), 84 HM ($n = 2$), Kolleg*innen- und Elterngespräche ($n = 2$). – Reflexionsprozesse ($n = 5$): eigenes Verhalten kritisch betrachten und ggf. neue Ansätze finden.

4. Welche Verbesserungs-vorschläge haben Sie? (n = 32)

- Verhaltensmonitoring (n = 13): Ausdehnung des Beurteilungszeitraumes (n = 7), Beurteilung per App (n = 3), Einsatz bei vielen Schüler*innen (n = 1), zwischenzeitliche Auswertung der Beurteilungen (n = 1), Skalenbreite reduzieren (n = 1).
- Fortbildungen (n = 10): mehr Trainings- und Praxisbeispiele/Übungen (n = 7), Wissensvertiefung/mehr Informationen (n = 3).
- Konzepterweiterung (n = 6): Fremdhospitationen organisieren (n = 3), Austausch zwischen Kolleg*innen initiieren (n = 2), Hort einbeziehen (n = 1).
- zeitliche Rahmung (n = 3): Zeitrahmen verlängern (n = 1)/verkürzen (n = 2).

Anmerkungen. *M* = Mittelwert, *SD* = Standardabweichung, *n* = Anzahl der Nennungen innerhalb der Ober-/Unter-Kategorie, SuS = Schüler*innen, 84 HM = 84 *Handlungsmöglichkeiten* (Hartke et al., 2018).

Die Antworten zu den geschlossenen Fragen 1 und 2 wurden über die Gesamtgruppe gemittelt.

Die Antworten zu den offenen Fragen 3 und 4 wurden anhand der Inhalte kategorisiert (induktiv), in Stichpunkten zusammengefasst und quantifiziert (Ober-/Unter-Kategorie).

haltensbeurteilung via App (n = 3) und vereinzelt wurde der Einsatz von DBR bei mehr als zwei Schüler*innen einer Klasse (n = 1), die zwischenzeitliche Auswertung der Beurteilungen (n = 1) und die Reduzierung der Skalenbreite (n = 1) angeregt. Ein weiteres Thema, zu dem es viele Anregungen gab, waren die Fortbildungen (n = 10). Die Lehrkräfte haben sich vor allem mehr Trainings- und Praxisbeispiele bzw. Übungen (n = 7) und zu bestimmten Bereichen mehr Informationen (n = 3) gewünscht. So wurde angeregt, Praxisbeispiele und Erfahrungen bei der Anwendung der 84 *Handlungsmöglichkeiten* zu thematisieren und mehr Trainingsbeispiele und Übungen einzubauen, die Fortbildungsinhalte noch konkreter auf einzelne Schüler*innen auszurichten und die Wissensvermittlung im Bereich Entwicklungspsychologie auszubauen. Vorschläge und Ideen zur inhaltlichen Ergänzung der Professionalisierungsmaßnahme wurden unter der Kategorie Konzepterweiterung (n = 6) geclustert. Hier wurde angeregt, Fremdhospitationen zu organisieren (n = 3), den Austausch zwischen Kolleg*innen stärker zu forcieren (n = 2) und den Hort systematisch einzubeziehen (n = 1). In einer letzten Kategorie wurden Aussagen zur zeitlichen Rahmung (n = 3) der Professionalisierungsmaßnahme zusammengefasst, die recht widersprüchlich ausfielen. Während einerseits vorgeschlagen wurde, den Interventionszeitraum zu verlängern (n = 1), plädierten andere dafür, ihn zu verkürzen (n = 2).

Lehrkrachteinschätzungen bezüglich einzelner Elemente der Professionalisierungsmaßnahme

Im Folgenden sollen einzelne Professionalisierungselemente genauer betrachtet werden. Begonnen wird mit den Lehrkrafterückmeldungen zu den Fortbildungen und dem eingesetzten Verhaltensmonitoring mittels *Direct Behavior Rating*, die Tabelle 35 zu entnehmen sind. Beide Bereiche wurden durch ein geschlossenes Antwortformat mit einer 10-er Likert-Skala (0 = totale Ablehnung; 10 = totale Zustimmung) erfasst.

Die Fortbildungen schätzten die Lehrkräfte im Mittel als hilfreich ein ($M = 7.5$, $SD = 1.37$) und deren Länge als angemessen ($M = 7.3$, $SD = 2.22$). Das Verhaltensmonitoring mittels DBR wurde weniger positiv bewertet. Während sowohl die Nützlichkeitseinschätzung des Instruments ($M = 6.63$, $SD = 2.77$) als auch die Einschätzung zu dessen Durchführbarkeit/Praxistauglichkeit ($M = 6.60$, $SD = 2.97$) als befriedigend zu interpretieren sind, ist die Wahrscheinlichkeit, dass die Lehrkräfte das Verhaltensmonitoring per *Direct Behavior Rating* in Zukunft einsetzen, nur noch ausreichend hoch ($M = 4.73$, $SD = 3.86$).

Alle neun Lehrkräfte, die eine *kooperativ-systemische Beratung* erhalten hatten, konnten bei der Abschlussveranstaltung diesbezüglich befragt werden. Die Befragung erfolgte über geschlossene und offene Antwortformate, die der Tabelle 36 zu entnehmen sind. Begonnen wird mit den geschlossenen Antwortformaten, die auf einer 10-er Likert-Skala die Spannbreite von totaler Ablehnung (= 0) und totaler Zustimmung (= 10) abdecken.

Die Wahrscheinlichkeit, dass die Lehrkräfte auch in Zukunft freiwillige *Beratungen* in Anspruch nehmen werden, ist tendenziell hoch ($M = 7.89$, $SD = 2.02$). Dies lässt auf eine grundsätzlich hohe Zufriedenheit der Lehrkräfte mit den Beratungssitzungen schließen. Positive Effekte der *Beratung* auf die Emotionsregulierung ($M = 6.6$, $SD = 2.45$), Stressregulation ($M = 6.11$, $SD = 2.54$) und Sicherheit in der Klasse ($M = 6.11$, $SD = 3.65$) schätzen die Lehrkräfte für sich als eher befriedigend ein.

Es folgen die Ergebnisse zur Fünf-Finger-Methode. Die Antworten zu „Das fand ich gut.“ ($n = 11$) wurden in die Kategorien Atmosphäre, Stil und Art der Beratung sowie Grundsätzliches unterteilt. Die meisten Kommentare gab es zum Stil und zur Art der Beratung ($n = 5$). Besonders hervorgehoben wurden hierbei die Führung und Aufbereitung des Gesprächs, das Eingehen auf die eigene Person und Rolle als Lehrkraft, die umfangreiche und inhaltliche Beratung sowie, dass sich die Beraterin mit Ratschlägen zurückgehalten habe. Die Atmosphäre während der *Beratung* wurde ebenfalls häufig erwähnt ($n = 4$), die von den Lehrkräften als offen und positiv beschrieben wurde. Dem Bereich Grundsätzliches ($n = 2$) wurden die Kommentare zugeordnet, die grundsätzlich positiv betonten, dass die *Beratung* (in der Schule) stattgefunden hat („Zeit für persönliche Einzelfallberatung“, „Beratung in

Tabelle 35: Gemittelte Lehrkrifteinschätzung ($N = 30$) bezüglich der durchlaufenen Fortbildungen und des eingesetzten Verhaltensmonitorings nach Beendigung der Professionalisierungsmaßnahme

Frage/Item	Antworten (gemittelt)																																	
Fortbildungen																																		
1. Als wie hilfreich schätzen Sie die Fortbildungen ein?	 <table><tr><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td>x</td><td></td><td></td><td></td></tr><tr><td>0</td><td>1</td><td>2</td><td>3</td><td>4</td><td>5</td><td>6</td><td>7</td><td>8</td><td>9</td><td>10</td></tr></table> 																			x				0	1	2	3	4	5	6	7	8	9	10
									x																									
	0	1	2	3	4	5	6	7	8	9	10																							
$M = 7.5, SD = 1.37$																																		
2. Die Länge der Veranstaltungen war angemessen.	 <table><tr><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td>x</td><td></td><td></td><td></td></tr><tr><td>0</td><td>1</td><td>2</td><td>3</td><td>4</td><td>5</td><td>6</td><td>7</td><td>8</td><td>9</td><td>10</td></tr></table> 																			x				0	1	2	3	4	5	6	7	8	9	10
									x																									
	0	1	2	3	4	5	6	7	8	9	10																							
$M = 7.3, SD = 2.22$																																		
Verhaltensmonitoring mittels <i>Direct Behavior Rating</i>																																		
3. Bewerten Sie die Nützlichkeit der DBR.	 <table><tr><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td>x</td><td></td><td></td><td></td></tr><tr><td>0</td><td>1</td><td>2</td><td>3</td><td>4</td><td>5</td><td>6</td><td>7</td><td>8</td><td>9</td><td>10</td></tr></table> 																			x				0	1	2	3	4	5	6	7	8	9	10
									x																									
	0	1	2	3	4	5	6	7	8	9	10																							
$M = 6.63, SD = 2.77$																																		
4. Wie durchführbar/praxistauglich finden Sie die DBR?	 <table><tr><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td>x</td><td></td><td></td><td></td></tr><tr><td>0</td><td>1</td><td>2</td><td>3</td><td>4</td><td>5</td><td>6</td><td>7</td><td>8</td><td>9</td><td>10</td></tr></table> 																			x				0	1	2	3	4	5	6	7	8	9	10
									x																									
	0	1	2	3	4	5	6	7	8	9	10																							
$M = 6.60, SD = 2.97$																																		
5. Wie wahrscheinlich ist es, dass Sie die DBR in Zukunft anwenden?	 <table><tr><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr><tr><td>0</td><td>1</td><td>2</td><td>3</td><td>4</td><td>5</td><td>6</td><td>7</td><td>8</td><td>9</td><td>10</td></tr></table> 																							0	1	2	3	4	5	6	7	8	9	10
	0	1	2	3	4	5	6	7	8	9	10																							
$M = 4.73, SD = 3.86$																																		

Anmerkungen. M = Mittelwert, SD = Standardabweichung, DBR = *Direct Behavior Rating* (Christ et al., 2009).

Tabelle 36: Gemittelte Lehrkräfteeinschätzung ($n = 9$) bezüglich der erhaltenen systemisch-kooperativen Beratungssitzung nach Beendigung der Professionalisierungsmaßnahme

Frage/Item	Antworten (gemittelt und kategorisiert)																																																																								
1. Wie wahrscheinlich ist es, dass Sie in Zukunft freiwillige Beratungen in Anspruch nehmen?	  <table><tr><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td>x</td><td></td><td></td><td></td><td></td></tr><tr><td>0</td><td>1</td><td>2</td><td>3</td><td>4</td><td>5</td><td>6</td><td>7</td><td>8</td><td>9</td><td>10</td><td></td></tr></table> $M = 7.89, SD = 2.02$								x					0	1	2	3	4	5	6	7	8	9	10																																																	
							x																																																																		
0	1	2	3	4	5	6	7	8	9	10																																																															
2. In welchem Ausmaß hatte die Beratung positiven Einfluss auf Ihre...	Emotionsregulierung   <table><tr><td></td><td></td><td></td><td></td><td></td><td></td><td>x</td><td></td><td></td><td></td><td></td><td></td></tr><tr><td>0</td><td>1</td><td>2</td><td>3</td><td>4</td><td>5</td><td>6</td><td>7</td><td>8</td><td>9</td><td>10</td><td></td></tr></table> $M = 6.67, SD = 2.45$ Stressregulation   <table><tr><td></td><td></td><td></td><td></td><td></td><td></td><td>x</td><td></td><td></td><td></td><td></td><td></td></tr><tr><td>0</td><td>1</td><td>2</td><td>3</td><td>4</td><td>5</td><td>6</td><td>7</td><td>8</td><td>9</td><td>10</td><td></td></tr></table> $M = 6.11, SD = 2.54$ Sicherheit in der Klasse   <table><tr><td></td><td></td><td></td><td></td><td></td><td></td><td>x</td><td></td><td></td><td></td><td></td><td></td></tr><tr><td>0</td><td>1</td><td>2</td><td>3</td><td>4</td><td>5</td><td>6</td><td>7</td><td>8</td><td>9</td><td>10</td><td></td></tr></table> $M = 6.11, SD = 3.65$							x						0	1	2	3	4	5	6	7	8	9	10								x						0	1	2	3	4	5	6	7	8	9	10								x						0	1	2	3	4	5	6	7	8	9	10	
						x																																																																			
0	1	2	3	4	5	6	7	8	9	10																																																															
						x																																																																			
0	1	2	3	4	5	6	7	8	9	10																																																															
						x																																																																			
0	1	2	3	4	5	6	7	8	9	10																																																															

Anmerkungen. M = Mittelwert, SD = Standardabweichung, n = Anzahl der Nennungen innerhalb der Ober-/Unter-Kategorie.

Die Antworten zu den geschlossenen Fragen 1 und 2 wurden über die Gesamtgruppe gemittelt.

Frage/Item	Antworten (kategorisiert und zusammengefasst)
3. Das fand ich gut. (n = 11)	– Atmosphäre (n = 4): angenehm (n = 2), offen (n = 2). – Stil und Art der Beratung (n = 5): Gesprächsführung, keine Ratschläge, Aufbereitung, Eingehen auf Person und Rolle als Lehrkraft, sehr umfangreiche, inhaltliche Beratung. – Grundsätzliches (n = 2): Zeit für persönliche Einzelfallberatung, Beratung in Schule nach Unterricht.
4. Darauf möchte ich hinweisen. / Das ist mir aufgefallen. (n = 2)	– erhaltene Wertschätzung in Beratung. – eine Beratungssitzung ist zu wenig („Tropfen auf den heißen Stein“).
5. Das gefällt mir nicht. (n = 0)	
6. Das nehme ich mit. (n = 8)	– Methoden (n = 3): in die Rolle des Kindes versetzen, Fragen im Indikativ beantworten, visuelle Darstellung der Problemsituation, bspw. durch Puppen. – Erkenntnis (n = 3): eigene Einstellung zu Schüler*innen kann Auswirkungen auf deren Sozialverhalten haben, positive Verstärkung hilft da, wo man es nicht erwartet, öfter mal den Blickwinkel ändern/Neues ausprobieren. – Selbstvertrauen (n = 2): Gelassenheit, an die eigenen Stärken glauben.
7. Das würde ich mir für die Zukunft wünschen. (n = 7)	– Unterstützung (n = 4): kurzfristige Hilfen und lehrerbegleitende Beratung in Stresssituationen durch kompetentes Fachpersonal (n = 2), psychologische Unterstützung bei Elterngesprächen, mehr Unterstützung seitens des Schulumtes (bspw. durch Supervision). – höhere Professionalität (n = 2): in schwierigen Situationen gelassen(er) bleiben, Erfolg bei der Anwendung. – Einbindung der Eltern (n = 1): Eltern mit „ins Boot“ holen.

Anmerkung. Die Antworten zu den offenen Fragen 3 bis 7 wurden anhand der Inhalte kategorisiert (induktiv), in Stichpunkten zusammengefasst und quantifiziert (Ober-/Unter-Kategorie).

Schulenach Unterricht“). Bei „Darauf möchte ich hinweisen.“ / „Das ist mir aufgefallen.“ ($n = 2$) wurde einerseits für die erhaltene Wertschätzung in der *Beratung* gedankt und andererseits angemerkt, dass eine Beratung zwar helfe, jedoch über einen längeren Zeitraum betrachtet, wie ein „Tropfen auf den heißen Stein“ wirke. Unter der Rubrik „Das gefällt mir nicht.“ gab es keinerlei Anmerkungen ($n = 0$). Die Antworten zum Item „Das nehme ich mit.“ waren vielfältig ($n = 8$). Zum einen erachteten die Lehrkräfte das Kennenlernen bestimmter Beratungsmethoden als hilfreich und nützlich für die eigene Arbeit ($n = 3$). Dazu gehörte, sich in die Rolle des Kindes zu versetzen, Fragen im Indikativ zu beantworten und die Problemsituation, bspw. durch Puppen, visuell darzustellen. Zum anderen nahmen die Lehrkräfte folgende neue Erkenntnisse ($n = 3$) aus dem Beratungsgespräch mit: Die eigene Einstellung zu Schüler*innen kann Auswirkungen auf deren Sozialverhalten haben; positive Verstärkung hilft da, wo man es nicht erwartet; es lohnt sich, öfter mal den Blickwinkel zu ändern und Neues auszuprobieren. Des Weiteren äußerten einige Lehrkräfte hinzugewonnenes Selbstvertrauen ($n = 2$), das durch „Gelassenheit“ und „den Glauben an die eigenen Stärken“ umschrieben wurde. Für die Zukunft wünschten sich die Lehrkräfte ($n = 7$) mehr Unterstützung anhand von kurzfristigen Hilfen, lehrerbegleitenden Beratungen in Stresssituationen und bei Elterngesprächen durch kompetentes Fachpersonal sowie mehr Unterstützung seitens des Schulamtes, etwa durch Supervision. Vereinzelt wünschten sich die Lehrkräfte eine noch höhere Professionalität ($n = 2$), um auch in schwierigen Situationen gelassen(er) bleiben zu können. Eine Lehrkraft wünschte sich eine stärkere Einbindung der Eltern.

5.2 Ergebniszusammenfassung

Nach der ausführlichen Darstellung der Ergebnisse der Evaluationsstudie zum Professionalisierungskonzept werden diese im Folgenden tabellarisch zusammengefasst (Tabelle 37).

Tabelle 37: Zusammenfassende Darstellung der Ergebnisse der Evaluationsstudie zum Professionalisierungskonzept, geordnet nach den Wirkungsebenen von Lehrer*innenfortbildungen, den abhängigen Variablen/Untersuchungsschwerpunkten und den Analyseverfahren

Wirksamkeitsebene	Abhängige Variable/Untersuchungsschwerpunkt	Verfahren	Ergebnisse
Schüler*innenoutcomes	spezifische Verhaltensauffälligkeiten	DBR-MIS, Mehrebenenanalyse	signifikante sprunghafte und kontinuierliche Reduktion der spezifischen Verhaltensauffälligkeiten während der Intervention
	Nichtregelkonformes Verhalten	DBR-SIS, Mehrebenenanalyse	signifikante sprunghafte und kontinuierliche Reduktion des nichtregelkonformen Verhaltens während der Intervention
	Effekte der 84 HM in Abhängigkeit vom Erklärungs- und Handlungsansatz: lerntheoretische vs. „kombinierte“ Intervention	DBR-MIS und DBR-SIS, Mehrebenenanalyse	bei lerntheoretischer Intervention: signifikante sprunghafte und kontinuierliche Reduktion der spezifischen Verhaltensauffälligkeiten und des nichtregelkonformen Verhaltens während der Intervention bei kombinierter Intervention: signifikante kontinuierliche Reduktion der spezifischen Verhaltensauffälligkeiten und des nichtregelkonformen Verhaltens während der Intervention
	Verhaltensauffälligkeiten	Gesamtproblemwert SDQ, ANOVA	signifikante und nachhaltige Reduktion der Verhaltensauffälligkeiten nach der Intervention
	prosoziales Verhalten	Prosoziales Verhalten SDQ, ANOVA	signifikante und nachhaltige Steigerung des prosozialen Verhaltens nach der Intervention

Wirksam- keitsebene	Abhängige Variable/Unter- suchungs- schwerpunkt	Verfahren	Ergebnisse
Lehrkrafthandeln	Einsatzhäufig- keit der 84 HM	Quantifizierung der Einsatzhäufigkeit der 84 HM während Intervention: sehr oft = 80-100 % oft = 60-80 % gelegentlich = 40-60 % selten = 20-40 % kaum < 20 %	Einsatzhäufigkeit: 67 % (= oft)
	Interaktions- effekt zwischen Einsatzhäufig- keit der 84 HM und Schüler*in- nenoutcomes	DBR-MIS und DBR- SIS, Mehr- ebenenanalyse mit dem vierfach ge- stuften Kreuzprädik- tor „Einsatzhäufig- keit“ (sehr oft/oft/ gelegentlich/selten)	signifikant negativer Inter- aktionseffekt: bei seltenem Einsatz der 84 HM (= 20-40 %) keine positiven Auswirkungen auf DBR-MIS und DBR-SIS fest- stellbar signifikant positiver Inter- aktionseffekt: bei häufigem Einsatz der 84 HM ($\geq 60\%$) positive Auswirkung auf DBR-MIS feststellbar
Lehrkraftkognitionen	Selbstwirksam- keitsempfinden	Lehrer-Selbst- wirksam- keitserwartung, ANOVA	signifikante Steigerung des Selbstwirksamkeitsempfindens nach der Intervention; keine nachhaltige Wirkung feststellbar
	Selbstwirksam- keitsempfinden ohne Beratung		signifikante Steigerung des Selbstwirksamkeitsempfindens nach der Intervention; keine nachhaltige Wirkung feststellbar
	Selbstwirksam- keitsempfinden mit Beratung		keine signifikante Steigerung des Selbstwirksamkeitsempfin- dens nach der Intervention; keine Follow-up-Effekte

Wirksam- keitsebene	Abhängige Variable/Unter- suchungs- schwerpunkt	Verfahren	Ergebnisse
	Interaktions- effekt zwischen Selbstwirksam- keitsempfinden und Schüler*in- nen-outcomes	DBR-MIS und DBR- SIS, Mehrebenen- analyse mit dem dreifach gestuften Kreuzprädiktor „Selbstwirksam- keitsempfinden“	positive Interaktionseffekte zwischen Selbstwirksamkeits- empfinden und spezifischen Verhaltensauffälligkeiten der Schüler*innen (DBR-MIS) kein positiver Interaktionseffekt zwischen Selbstwirksamkeits- empfinden und Regelverhalten der Schüler*innen (DBR-SIS)
Lehrkraftreaktionen	Zufriedenheit und Nützlich- keitseinschät- zungen zur Professionalisie- rungsmaßnahme und ihren Ele- menten	Evaluationsfragebo- gen zum Professio- nalisierungskonzept, deskriptive Auswer- tung	Gesamte Professionalisie- rungsmaßnahme gute Betreuung Nützlichkeitseinschätzung des Gesamtkonzepts befriedigend Professionalisierungselemente Fortbildungen: hilfreich und Länge angemessen DBR: Nützlichkeit sowie Pra- xistauglichkeit befriedigend kooperativ-systemische Bera- tung: hohe Zufriedenheit mit Beratung; Einschätzung von positiven Effekten auf eigene Person befriedigend

Anmerkungen. DBR-MIS = *Direct Behavior Rating-Multi-Item-Scale*, DBR-SIS = *Direct Behavior Rating-Single-Item-Scale* (Christ et al., 2009), SDQ = *Strengths and Difficulties Questionnaire – Lehrkraftversion* (Goodman, 2005), ANOVA = *Einfaktorielle Varianzanalyse*, 84 HM = *84 Handlungsmöglichkeiten* (Hartke et al., 2018), Evaluationsfragebogen zum Professionalisierungskonzept, „kombinierte“ Intervention = Intervention mit lerntheoretischem/kognitions-psychologischem/schüler*innen-zentriertem Erklärungs- und Handlungsansatz.

6 Diskussion

Bevor die in Kapitel 5 vorgestellten Ergebnisse der Evaluationsstudie zum Professionalisierungskonzept diskutiert werden, erfolgt zunächst eine kritische Reflexion der Untersuchungsmethoden. Auf diese Weise sollen etwaige Einschränkungen der Aussagekraft der erhobenen Daten vorab diskutiert und eine Grundlage für die Interpretation der dargestellten Ergebnisse geschaffen werden.

6.1 Methodenkritische Reflexion der Ergebnisse

Bei der vorliegenden Evaluationsstudie handelt es sich um eine explanative, quasi-experimentelle Feldstudie im Ein-Gruppen-Design mit Messwiederholungen, bei der eine kontrollierte Einzelfalluntersuchung, standardisierte statusdiagnostische Verfahren und tentative, vertiefende Analysen zur Erhebung von Schüler*innen- und Lehrkraftdaten eingesetzt wurden. Es ist demnach eine komplexere Studie im Mixed-Method-Design, bei der es gilt, die jeweils spezifisch eingesetzten Erhebungs- und Auswertungsmethoden kritisch zu beurteilen. Für eine strukturierte Analyse erfolgen separate Betrachtungen des quasi-experimentellen Ein-Gruppen-Designs, der kontrollierten Einzelfallstudie im Multi-Baseline-Design und der qualitativen Befragungen.

Ein essentielles Kriterium für die Gültigkeit der Untersuchungsergebnisse stellt die interne und externe Validität dar. Eine experimentelle Untersuchung ist intern valide, wenn die kausale Interpretation ihrer Ergebnisse inhaltlich eindeutig ist – wenn also ein aufgetretener Effekt als Unterschied im Wert der abhängigen Variable eindeutig auf das *Treatment* und/oder die *Intervention* zurückgeführt werden kann und der Einfluss von Störvariablen dabei kontrolliert oder ausgeschaltet wurde. Traditionell (Cook & Campbell, 1979; vgl. Shadish et al. 2002, S. 55) werden folgende neun Bedrohungen der internen Validität unterschieden:

1. Unklare zeitliche Abfolge
2. Selektionseffekte
3. Externe zeitliche Einflüsse
4. Reifungsprozesse
5. Statistische Regressionseffekte
6. Experimentelle Mortalität

7. Testübung
8. Mangelnde instrumentelle Reliabilität
9. Kombination der genannten Bedrohungen der internen Validität

Bedrohungen der internen Validität führen dazu, dass der statistisch nachgewiesene Zusammenhang zwischen Variablen nicht eindeutig als Ursache-Wirkungs-Zusammenhang interpretiert werden kann (Bortz & Döring, 2006).

Als externe Validität wird die Verallgemeinerbar- oder Generalisierbarkeit der Untersuchungsergebnisse bezeichnet. Besonders hoch ist diese, wenn die Untersuchung in der natürlichen Umgebung stattfindet (Bortz & Döring, 2006). Die externe Validität kann, den Komponenten eines Experimentes entsprechend, in vier Aspekte unterteilt werden (vgl. Shadish et al., 2002):

1. Personen: Sind die Ergebnisse auf eine Grundgesamtheit von Personen übertragbar, aus denen die Versuchspersonen stammen, z.B. auf alle Schüler*innen des Bundeslandes M-V oder Schüler*innen im FSP ESE der Bundesrepublik? Eine Generalisierbarkeit auf Personen einer Grundgesamtheit ist dann möglich, wenn eine repräsentative Stichprobe dieser Grundgesamtheit gezogen und in einem Experiment untersucht wurde. Alle Faktoren, die die Repräsentativität der Stichprobe beeinträchtigen, wie z.B. der Verlust von Probanden durch vorzeitigen Abbruch, schränken dann auch die Generalisierbarkeit ein.
2. Treatments: Sind die Ergebnisse auf ähnliche Treatments übertragbar? Kann sich etwa der Effekt der Professionalisierungsmaßnahme und seiner Elemente auch bei einer anderen Variante desselben Treatments oder Kombination mit anderen Interventionen oder eine begrenzte Implementierung zeigen?
3. Abhängige Variablen (outcomes): Sind die Ergebnisse auch für andere abhängige Variablen gültig? Zeigt das entwickelte Professionalisierungskonzept neben der Stärkung des Selbstwirksamkeitsempfindens der Lehrkräfte auch einen positiven Effekt auf andere Variablen, wie ihr unterrichtliches Handeln oder die Schüler*innenoutcomes?
4. Untersuchungsbedingungen (settings): Gelten die Ergebnisse auch unter anderen situativen Bedingungen und kulturellen Kontexten? Gibt es etwa Unterschiede bei der Durchführung in eher ländlichen im Gegensatz zu städtischen Gebieten? Können die Ergebnisse auf andere Grundschulen übertragen werden?

Nach Bortz und Döring (2016) gelingt es bei Studien nur selten, sowohl die interne als auch die externe Validität voll zu erfüllen. Dies ist auch in der vorliegenden Studie der Fall, wobei vor allem Einschränkungen der internen Validität zu diskutieren sind. Da die vorliegende Feldstudie nicht unter Laborbedingungen, sondern in der natürlichen Umgebung der Proband*innen durchgeführt wurde, waren Störvariablen nur schwer zu kontrollieren und können so Einfluss auf die Untersuchungsergebnisse genommen haben – die

interne Validität ist im Vergleich zu anderen experimentellen Anordnungen in der vorliegenden Studie niedriger. Die externe Validität und somit die Repräsentativität und Generalisierbarkeit der Ergebnisse kann hingegen als eher hoch eingeschätzt werden. Im Folgenden sollen zunächst die interne und externe Validität des quasi-experimentellen Ein-Gruppen-Designs systematisch betrachtet werden. Dieser Teil der Evaluationsstudie fokussiert die Wirkungsebenen Lehrkraftkognitionen und Schüler*innenoutcomes (vgl. Kap. 4.1) und betrachtet die jeweiligen Merkmalsausprägungen zu drei MZP, vor der Intervention (Prä-), unmittelbar nach der Intervention (Post-) und vier Monate nach der Intervention (Follow-up-Testung).

Die wohl schwerwiegendste Gefährdung der internen Validität stellt das Ein-Gruppen-Design dar. Aufgrund des Fehlens einer Kontrollgruppe kann nicht ausgeschlossen werden, dass Veränderungen ursächlich nicht ausschließlich auf das *Treatment* und seine Elemente zurückzuführen sind (Bortz & Döring 2016, S. 202). So können sich Veränderungen des sozialen außerschulischen Umfelds der Kinder auswirken und damit *externe zeitliche Einflüsse* die Ergebnisse beeinflussen. Eine weitere Gefährdung der internen Validität stellen in diesem Zusammenhang *Reifungsprozesse* dar. Es kann nicht ausgeschlossen werden, dass sich die Schüler*innen im Verlauf der Evaluationsstudie von allein in ihrem Verhalten verändert haben (Bortz & Döring 2016, S. 103). Des Weiteren können sich die Messungen aus formal-statistischen Gründen verändern. Diese sogenannten *Regressionseffekte* betreffen vorzugsweise Extremwerte, die bei wiederholten Messungen zur Mitte tendieren. Alle drei Aspekte können fälschlich als kausale Treatmenteffekte interpretiert werden und müssen bei der Interpretation der Ergebnisse berücksichtigt werden. Davon abgesehen, kann das Studiendesign uneingeschränkt verwendet werden, um zu zeigen, ob während der Intervention bedeutsame Merkmalsänderungen eingetreten sind.

Die Proband*innen der Untersuchungsgruppen wurden außerdem nicht zufällig ausgewählt, sondern befanden sich in ihren natürlichen Gruppen, ihren Schulen bzw. Schulklassen. Die Lehrkräfte meldeten sich auf freiwilliger Basis und aus Interesse für die Studienteilnahme an. Bei den Schüler*innen war die Bedingung für die Aufnahme in die Untersuchungsgruppe das Aufweisen auffälliger Werte im *Kurzscreening schulischen Verhaltens* (Blumenthal & Blumenthal, in Vorb.). Die Lehrkräfte, die Interesse an der Studienteilnahme signalisiert hatten, schätzten dafür diejenigen Schüler*innen ihrer Klassen ein, für die ein Einverständnis der Erziehungsberechtigten vorlag (vgl. Kap. 4.3.3). Bei dieser Art der Untersuchungsgruppenbildung können *Selektionseffekte* nicht ausgeschlossen werden. Die Versuchspersonen können sich etwa anhand gewisser Dispositionen wie Alter und Geschlecht systematisch von der Zielpopulation unterscheiden. Bei den vorliegenden Untersuchungsgruppen zeigen sich vornehmlich unausgewogene Geschlechterverhältnisse. So ist die Untersuchungsgruppe Lehrkräfte ausnahmslos

weiblich. Laut repräsentativer Zahlen beträgt der Frauenanteil bei Grundschullehrkräften in Deutschland jedoch rund 90 % (Statistisches Bundesamt 2019, S. 95) – somit ist der Selektionseffekt bezüglich des Geschlechts der Lehrkräfte als marginal einzuschätzen. Die Untersuchungsgruppe der Schüler*innen zeigt ebenfalls ein unausgewogenes Geschlechterverhältnis: Sowohl in der Gesamtstichprobe aller gescreenten Kinder als auch in der Untersuchungsgruppe wurden die Jungen in den Subskalen Regelverhalten, emotional-soziales Verhalten oder Arbeitsverhalten signifikant, im Schnitt drei bis viermal häufiger als auffällig eingeschätzt als die Mädchen. Auch dies spiegelt die in anderen Studien festgestellte Geschlechterverteilung wider. Jungen haben ein drei bis vier Mal größeres Risiko, Verhaltensauffälligkeiten oder emotional-soziale Entwicklungsstörungen zu entwickeln als Mädchen (Brezinka, 2003). Diese sind somit auch in der Grundgesamtheit aller Kinder mit Verhaltensauffälligkeiten und emotional-sozialen Entwicklungsstörungen überproportional häufig vertreten. Selektionseffekte sind in der vorliegenden Studie somit eher gering.

Da es zwischen Prä- und Posttest bereits zu einem Dropout von 11 Lehrkräften (~ 5 %) kam und für die Bildung der Untersuchungsgruppe weitere 24 Lehrkraftdaten (~ 50 %) aufgrund theoretischer Überlegungen nicht in die finalen Berechnungen eingingen (vgl. Kap. 4.3.3.1), muss außerdem die *experimentelle Mortalität* diskutiert werden. Die Wirksamkeit eines Trainings kann stark überschätzt werden, „wenn bei der Nachher-Messung nur noch die besonders motivierten und leistungsstarken Teilnehmenden präsent sind, weil die anderen die Maßnahme inzwischen abgebrochen haben“ (Bortz & Döring 2016, S. 103). Aus diesem Grund sollte festgehalten werden, wann und warum Untersuchungspersonen aus der Studie herausgefallen sind, um dies bei der Interpretation berücksichtigen zu können. Dies ist vorliegend der Fall. Der Zeitpunkt und die Gründe für den Abbruch bzw. den Ausschluss aus der Untersuchungsgruppe wurden in der Studie genau dokumentiert (vgl. Kap. 4.3.3.1) und können damit bei der Interpretation berücksichtigt werden. Der von den Lehrkräften initiierte Dropout kam maßgeblich durch gesundheitliche Gründe (Schwangerschaft oder Krankheit) zustande. Vereinzelt wurden kein Grund bzw. der Zeitaufwand für das Projekt als Abbruchgrund genannt. Eine Gefährdung der internen Validität kommt daher eher bei dem umfassenderen Ausschluss von Proband*innen durch die Versuchsleiterin in Betracht. Dieser wurde durchgeführt, um eine Aussage über die Umsetzungstreue und damit auch mögliche Effekte der *Intervention* treffen zu können (s. Kap. 4.3.3). Die Lehrkräfte, die mindestens 80 % der möglichen Datenpunkte erhoben hatten und deren Daten bei den finalen Berechnungen berücksichtigt wurden, könnten sich jedoch durch eine höhere Motivation und bessere Umsetzung der Intervention auszeichnen. Wäre dies der Fall, so könnte die Wirksamkeit der Intervention überschätzt werden. Dieser Aspekt muss ebenfalls bei der Interpretation berücksichtigt werden.

Eine weitere Einschränkung der internen Validität wird als *mangelnde instrumentelle Reliabilität* beschrieben: „Im Verlaufe der Zeit kann sich die Qualität eines Messinstrumentes verändern, so dass Messwertunterschiede entstehen, die fälschlich als kausale Treatmentwirkung interpretiert werden können“ (Bortz & Döring 2016, S. 103). Um dieser Gefährdung der internen Validität vorzubeugen, wurden in der vorliegenden Studie standardisierte Messinstrumente mit hoher Reliabilität eingesetzt. Die Einschätzungen erfolgten jedoch durch die Lehrkräfte selbst, sodass etwa Voreingenommenheit, Wunschenken bzw. erwünschte Antworten seitens der Beurteiler*innen nicht auszuschließen sind.

Während sich in der vorliegenden Studie eine größere Anzahl an vermuteten Gefährdungen der internen Validität identifizieren lassen, zeigen sich besondere Stärken im Bereich der externen Validität. Da die ungünstigen Einflüsse auf die externe Validität eher gering sind, können die Ergebnisse als ökologisch valide angesehen werden. Dies soll im Folgenden anhand der vier Aspekte externer Validität nach Shadish et al. (2002) genauer ausgeführt werden:

1. Personen: Da die Studie in einem natürlichen Setting stattfand, ist davon auszugehen, dass die Ergebnisse – die zuvor diskutierten Einschränkungen der internen Validität berücksichtigend – über die Untersuchungspersonen und -bedingungen hinaus generalisierbar sind und somit auf die Grundgesamtheit der Grundschullehrkräfte der Klassen eins bis vier als auch die der Grundschüler*innen im Bereich des FSP ESE in Deutschland übertragen werden können.
2. Treatment: Es ist nicht davon auszugehen, dass es starke Wechselwirkungen des Kausaleffektes mit den Treatmentbedingungen gab (Bortz & Döring 2016, S. 104). Dies betrifft neben dem *Treatment*¹⁸ auch die implementierte *Intervention*¹⁹. Der Einsatz der *84 Handlungsmöglichkeiten* beruhte auf dem Grundsatz evidenzbasierter Praxis (vgl. Kap. 2.1.3.3). Hierbei verlangt die „Umsetzung konkreter Maßnahmen im Unterricht [...] einen planenden und verantwortlich handelnden Praktiker, der die einzelne Situation mit den speziellen Bedingungen erfasst und berücksichtigt“ (Hartke et al. 2018, S. 10). Individuelle Vorgehensweisen, Varianten und Abweichungen von den konkreten Handlungsmöglichkeiten und -maßnahmen waren somit intendiert.
3. Abhängige Variablen (outcomes): Die Effekte des Professionalisierungskonzepts und seiner Elemente (vgl. Kap. 4.1) wurden auf unterschiedlichen Ebenen (vgl. Kap. 4.1) und anhand unterschiedlicher Operationali-

18 Treatment hier: Das Professionalisierungskonzept inklusive seiner Handlungsfelder und Professionalisierungselemente.

19 Intervention hier: Einsatz des Professionalisierungselements *84 Handlungsmöglichkeiten* (Hartke et al., 2018).

sierungen (vgl. Kap. 4.2.1) gemessen. Dies lässt den Schluss zu, dass die Ergebnisse für mehrere abhängige Variablen gültig sind (Bortz & Döring 2016, S. 95).

4. Untersuchungsbedingungen (settings): Das Untersuchungssetting wurde so gestaltet, dass es eine möglichst große Bandbreite an Beschulungsformen abbildet. Mit vier Schulen im städtischen und drei Schulen im ländlichen Gebiet, davon sechs öffentliche Schulen und eine in privater Trägerschaft sowie zwei Schulen mit reformpädagogischem Profil, bildet es eine diverse Schullandschaft ab. Die Ergebnisse können demnach auf unterschiedliche Grundschulsettings in Deutschland der Klassenstufen eins bis vier übertragen werden.

Neben der internen und externen Validität soll auch der Stichprobenumfang als Aspekt statistischer Validität diskutiert werden. Bortz und Döring (2016, S. 845) geben für die ANOVA mit Messwiederholungen, die im Rahmen des quasi-experimentellen Ein-Gruppen-Designs verwendet wurde um Merkmalssänderungen der Schüler*innen und Lehrkräfte über die Zeit zu erfassen, eine Stichprobengröße von mindestens 27 Personen an. Während die Untersuchungsgruppe Schüler*innen ($N = 27$) sich noch in dem vorgegebenen Rahmen befinden, unterschreitet die Untersuchungsgruppe Lehrkräfte ($N = 21$) diesen um sechs Personen. Da ein zu geringer Stichprobenumfang zu einer mangelnden Teststärke bei Signifikanztests führen kann, muss dies bei der Interpretation der Daten zu den Lehrkraftkognitionen berücksichtigt werden.

Handlungsleitend für die Interpretation der Studienergebnisse aus dem quasi-experimentellen Ein-Gruppen-Design kann festgehalten werden: Aussagen über Kausalzusammenhänge zwischen *Treatment/Intervention* und Lehrkräfte-/Schüler*innenoutcomes erlaubt das Studiendesign nicht. Dennoch können Merkmalsveränderungen über die Zeit auf einen kausalen Zusammenhang hindeuten, insbesondere, wenn diese auf mehreren Ebenen und anhand unterschiedlicher abhängiger Variablen sichtbar werden. Statistisch signifikante Effektgrößen können in diesem Zusammenhang daher als grober Hinweis auf Kausalzusammenhänge zwischen abhängigen und unabhängigen Variablen gedeutet werden. Die hohe externe Validität der Studie ermöglicht es außerdem, generalisierende Annahmen über die Untersuchungspersonen und -bedingungen hinaus zu treffen.

Die Wirkungsebene Schüler*innenoutcomes wurde neben statusdiagnostischen Verfahren auch mittels einer kontrollierten Einzelfallstudie im Multiple-Baseline-Design untersucht. Im Folgenden sollen deren interne sowie externe Validität kritisch diskutiert werden.

Bei dem umgesetzten AB-Design handelt es sich um ein quasi-experimentelles Einzelfallstudiendesign (Barlow et al., 2009). Da das Design nur aus einer A- und einer B-Phase besteht und keine der beiden Phasen umgekehrt durchgeführt wurde (wie in einem ABA- oder ABAB-Design), ist die interne Validität der Studie eingeschränkt. So können potentielle *Reifungs-*

prozesse und externe zeitliche Einflüsse nicht kontrolliert werden (Tate et al., 2016). Um die interne Validität zu erhöhen, wurde der AB-Versuchsplan um ein zweifachgestuftes Multiple Baseline Design über Personen ergänzt, bei dem die Grund- und Interventionsraten bei den Proband*innen gestaffelt einsetzen. In englischsprachigen Standardwerken zu Einzelfalldesigns (Barlow et al., 2009; Kazdin, 2011) werden mindestens drei Grundratenerhebungen empfohlen. Grundsätzlich genügen zwei Baselines, um die interne Validität eines AB-Plans deutlich zu erhöhen (Jain & Spieß 2012, S. 227).

Eine weitere mögliche Einschränkung bei kontrollierten Einzelfallstudien ist die (Nicht-)Verallgemeinerbarkeit der Ergebnisse. Die Auswirkungen eingesetzter Interventionen können grundsätzlich nur bedingt auf eine bestimmte Zielpopulation generalisiert werden (Maggin et al., 2018). Die Verallgemeinerbarkeit von kontrollierten Einzelfallstudien entsteht durch deren Replikation. Erst durch eine größere Anzahl kontrollierter Einzelfallstudien, deren Ergebnisse in einer Metaanalyse statistisch integriert werden, sind Rückschlüsse auf eine Zielpopulation zulässig (Tate et al., 2016). Eine Anzahl von 27 kontrollierten Einzelfallstudien, wie sie in der vorliegenden Studie gegeben ist, macht Rückschlüsse auf die Population der Grundschüler*innen mit Verhaltensauffälligkeiten und emotional-sozialen Entwicklungsstörungen möglich.

Ein anderer Kritikpunkt, der häufig im Kontext der Erhebung von Zeitreihen diskutiert wird, betrifft den der Autokorrelation der Daten (Lei et al., 2007). Daten aus Zeitreihenanalysen korrelieren häufig stark untereinander, weil „derselbe Fehler- oder Zufallseinfluss die abhängige Variable über mehrere aufeinanderfolgende Messgelegenheiten hinweg“ (Jain & Spieß 2012, S. 242) beeinflussen kann. Dies kann dazu führen, dass die Unabhängigkeitsannahme in Einzelfalluntersuchungen verletzt sowie die Reliabilität systematisch überschätzt wird (Rowley, 1989). Wenngleich dieser Punkt in der jüngeren Zeit verstärkt diskutiert wurde (Lei et al., 2007), gibt es derzeit noch keinen tragfähigen Vorschlag zur Lösung des Problems. Obwohl es nicht ratsam ist, bei Beobachtungsdaten derselben Personen automatisch von unkorrelierten Fehlern auszugehen, hängt das Ausmaß der Autokorrelation von der Art der Datenerhebung ab und ist bei Beobachtungsdaten von Einzelpersonen nach wie vor umstritten (Lei et al. 2007, S. 434). Der Herausforderung der Autokorrelation kann daher nicht begegnet werden. Es sei aber darauf hingewiesen, dass die Autokorrelationen in den meisten Fällen nicht besonders hoch waren.²⁰

Kritisch anzumerken ist weiter, dass die aus der kontrollierten Einzelfallstudie auf Gruppenebene ermittelten Befunde nicht zwangsläufig auf die Individualebene übertragbar sind. Die Daten lassen Aussagen über die Merkmalsveränderung auf Gruppenebene zu, die Anzahl der Non-Responder kann

20 Als hohe Werte wurden lag-1-Autokorrelationen von -0.63, 0.54, -0.50, 0.55, 0.50 angesehen.

dadurch allerdings unterschätzt werden. Dies muss bei der Interpretation der auf Gruppenebene generierten Ergebnisse berücksichtigt werden.

Zusammenfassend für die Interpretation der kontrollierten Einzelfallstudie im AB- Versuchsplan kann festgehalten werden, dass aufgrund des Multiple Baseline Designs über Personen und der Aggregation von 27 Fällen auf Gruppenebene neben grundsätzlichen Aussagen zu Merkmalsveränderungen über die Zeit auch Rückschlüsse über Kausalzusammenhänge zwischen der *Intervention* und den Schüler*innenoutcomes möglich sind.

Die Lehrkraftreaktionen wurden anhand tentativ qualitativer Befragungen erhoben. Die Zufriedenheits-, Nützlichkeits- und Praktikabilitätseinschätzungen der Lehrkräfte bezüglich der Professionalisierungsmaßnahme und ihrer -elemente wurde nach Beendigung des *Treatments* anhand von teilstandardisierten, qualitativen Fragebögen einmalig erfasst. Wie der Begriff „tentativ“ bereits andeutet, können anhand der Daten keine generalisierenden, über die Untersuchungspersonen und -bedingungen hinausreichenden Aussagen über Lehrkraftreaktionen gemacht werden. Die Ergebnisse lassen jedoch Rückschlüsse über die grundlegenden Reaktionen der an der Studie beteiligten Lehrkräfte zu und ermöglichen Aussagen über die Wahrscheinlichkeit des zukünftigen Einsatzes der eingeführten Professionalisierungselemente (vgl. Kap. 3.2). Anhand dieser Ergebnisse wiederum können hypothetische Ableitungen über die grundsätzliche Akzeptanz der Professionalisierungsmaßnahme bei Grundschullehrkräften getroffen werden.

Nach der kritischen Reflexion der Untersuchungsmethodik sollen im Folgenden die Ergebnisse diskutiert und in den fachlichen Gesamtzusammenhang eingeordnet werden. Das anhand theoretischer und empirischer Erkenntnisse (vgl. Kap. 2.3) entwickelte Professionalisierungskonzept zur Förderung der ESE von Schüler*innen durch Erweiterung spezifischer Lehrkraftkompetenzen (vgl. Kap. 3) soll dabei anhand der Ergebnisse auf den Wirkungsebenen Schüler*innenoutcomes, Lehrkrafthandeln, Lehrkraftkognitionen und Lehrkraftreaktionen (Lipowsky, 2010) kritisch begutachtet werden. Für jede Wirkungsebene erfolgt dabei eine zusammenfassende Darstellung der Ergebnisse, die Beantwortung der aufgestellten Fragestellungen und Hypothesen (vgl. Kap. 4.2.2), eine Wertung und Einschätzung in Bezug auf die Wirksamkeit der Professionalisierungsmaßnahme und ihrer -elemente, die Einordnung in den Gesamtzusammenhang der bisherigen Ergebnisse der aktuellen Fachliteratur sowie ggf. Hinweise auf mögliche Forschungsdesiderate.

6.2 Diskussion der Ergebnisse

6.2.1 Schüler*innenoutcomes

Spezifische Verhaltensauffälligkeiten und nichtregelkonformes Verhalten vor und während der Intervention

Hypothese 1a, wonach sich während des Einsatzes der *84 Handlungsmöglichkeiten* die spezifischen Verhaltensauffälligkeiten der Schüler*innen sprunghaft und kontinuierlich verringern, kann bestätigt und die Nullhypothese abgelehnt werden. Die Reduktion zeigt sich anhand der gemittelten Mediandifferenzen und Trends in Grund- und Interventionsraten (s. Tabelle 15), schwacher bis moderater Effekte bei den Überlappungsmaßen (s. Tabelle 16) sowie eines höchst signifikanten Level- und Slope-Effekts im mehrbenenanalytischen Regressionsmodell (s. Tabelle 18).

Hypothese 1b kann ebenfalls bestätigt werden. Auch das nichtregelkonforme Verhalten der 27 Schüler*innen hat sich während des Einsatzes der *84 Handlungsmöglichkeiten* reduziert. Dies zeigt sich an den gemittelten Mediandifferenzen sowie der Entwicklung der Trends und Standardabweichungen in Grund- und Interventionsraten (s. Tabelle 19), den moderaten Effekten bei drei Überlappungsmaßen (s. Tabelle 20) und einem höchst signifikanten Level- und Slope-Effekt im mehrbenenanalytischen Regressionsmodell (Tabelle 22).

Hauptergebnis 1: Während des Einsatzes der *84 Handlungsmöglichkeiten* verringern sich die spezifischen Verhaltensauffälligkeiten und das nichtregelkonforme Verhalten der Schüler*innen signifikant. Die Verringerung zeigt sich sowohl anhand eines unmittelbaren, sprunghaften als auch eines kontinuierlichen Abfalls des Verhaltensverlaufs.

Das Hauptergebnis 1 stützt die empirischen Ergebnisse von Hartke und Urban (2011), die im Rahmen einer Studie zur Vorgängerversion der *84 Handlungsmöglichkeiten* bei der Anwendung lerntheoretischer Ansätze wie der Verhaltensmodifikation mit Verstärkersystemen und Verhaltensverträgen einen deutlichen Rückgang von Verhaltensproblemen feststellen konnten. Wie bereits dargelegt wurde, können bei den vorliegenden Ergebnissen *Reifungsprozesse*, *zeitliche Einflüsse* und die *Autokorrelation* der Daten nicht gänzlich ausgeschlossen werden (Kap. 6.1). Die vorliegende Studie weist außerdem nach, dass während des Einsatzes von kognitionspsychologischen und schüler*innenzentrierten Handlungsmöglichkeiten ebenfalls signifikante Verhaltensveränderungen bei den Schüler*innen sichtbar werden. Da die Zusammenfassung aller 27 Einzelfälle und deren mehrbenenanalytische Auswertung verallgemeinernde Schlussfolgerungen über die Wirksamkeit von

Interventionen zulässt (Van den Noortgate & Onghena, 2003), sind zudem Annahmen über Kausalzusammenhänge möglich. Anhand der vorliegenden Studienergebnisse kann daher konstatiert werden, dass der Einsatz der *84 Handlungsmöglichkeiten* Schüler*innenverhalten positiv beeinflusst. Dies unterstützt die Bewertung des Konzepts als evidente Fördermaßnahme (Blumenthal et al. 2020, S. 64ff.; Hartke, 2017; Grosche, 2017) für die präventive unterrichtsintegrierte Förderung auf den Förderebenen I und II (Cheney et al., 2008; Gresham, 2005). Weitere kontrollierte Einzelfallstudien mit mehreren Grund- und Interventionsraten bzw. eines mindestens dreifach gestuften Multiple Baseline Designs und einer noch höheren Proband*innenzahl sollten die Ergebnisse jedoch ergänzen. Außerdem soll an dieser Stelle erneut darauf hingewiesen werden, dass die auf Gruppenebene ermittelten Befunde nicht zwangsläufig auf die Individualebene übertragbar sind (vgl. Kap. 6.1). In der vorliegenden Studie können bei immerhin gut einem Drittel (DBR-SIS) bzw. nahezu der Hälfte (DBR-MIS) der Schüler*innen kontinuierliche und/oder sprunghafte Verhaltensbesserungen festgestellt werden. Die Ergebnisse zeigen damit aber auch auf, dass in der schulischen Praxis und Anwendung wirksamer Fördermaßnahmen die sorgfältige Prüfung auf Einzelebene unabdingbar ist, um ggf. frühzeitig Modifikationen vornehmen zu können (Casale et al., 2015a). So kann ein evidenzbasierter Entscheidungsprozess auch zyklisch angeordnet sein, beispielsweise wenn die Maßnahmen sich als nicht wirksam erwiesen haben (Blumenthal & Mahlau, 2015). Ist dies der Fall, so beginnt der Prozess einer evidenzbasierten Entscheidung für ein angemessenes pädagogisches Vorgehen von neuem (vgl. Abbildung 3). Weitere Studien, die die individuelle Wirksamkeit der *84 Handlungsmöglichkeiten* in den Blick nehmen, sollten die aktuelle Forschungslage daher ergänzen.

Wirksamkeit der 84 Handlungsmöglichkeiten abhängig von deren Erklärungs- und Handlungsansatz

Hypothese 1c, wonach sich unterschiedliche Interaktionseffekte zwischen den Schüler*innenoutcomes und den *84 Handlungsmöglichkeiten* in Abhängigkeit zu deren Erklärungs- und Handlungsansatz zeigen, kann bestätigt werden. Bei den lerntheoretisch ausgerichteten Interventionen werden sowohl unmittelbare als auch kontinuierliche Interventionseffekte sichtbar (s. Tabelle 23 und Tabelle 25). Bei den kombinierten Interventionen (lerntheoretisch/kognitionspsychologisch/schüler*innenzentriert) zeigen sich in erster Linie kontinuierliche Interventionseffekte. Zwar reduzieren sich die spezifischen Verhaltensauffälligkeiten und das nichtregelkonforme Verhalten auch hier sprunghaft mit dem Einsetzen der *Intervention*, aber diese Level-Effekte sind nicht signifikant (s. Tabelle 24 und Tabelle 26).

Hauptergebnis 2: Nach dem Einsetzen lerntheoretisch ausgerichteter Handlungsmöglichkeiten kann sowohl eine sprunghafte als auch eine kontinuierliche Verringerung der spezifischen Verhaltensauffälligkeiten und des nichtregelkonformen Verhaltens beobachtet werden. Bei kombinierter Förderung zeigt sich maßgeblich eine kontinuierliche Reduzierung der Verhaltensauffälligkeiten.

Die Ergebnisse stützen die in der Forschung zu evidenzbasierten Maßnahmen im Bereich Verhalten generierten Erkenntnisse, wonach bei lerntheoretischen Interventionen Verhaltensänderungen schneller und sprunghafter eintreten als bei kognitionspsychologisch ausgerichteter Förderung, die ihre Wirkung eher langfristig entfaltet (Beelmann, 2006; Hillenbrand, 2015; Lösel & Beelmann, 2003; Wilson & Lipsey, 2007). Bemerkenswert ist, dass bei kognitionspsychologisch bzw. schüler*innenzentriert ausgerichteten Handlungsmöglichkeiten ebenso wie bei rein lerntheoretischen kontinuierliche Verringerungen des Problemverhaltens sichtbar werden. Reine Verhaltenstrainings zeigten hingegen in vielen Studien keine längerfristige Wirkung (Lösel & Beelmann, 2003). Die empirisch bewährten Handlungsansätze bei Verhaltensauffälligkeiten und emotional-sozialen Entwicklungsstörungen der Lerntheorie (Döpfner 2003, S. 263; Hulac & Briesch, 2017; Myschker & Stein 2014), der Kognitions- (Beelmann & Raabe, 2007; Crick & Dodge, 1994; Lemerise & Arsenio, 2000) und der Entwicklungspsychologie (Laucht et al., 1997) scheinen somit ihre positive Wirkung auch in der auf schulische Förderung hin methodisch-didaktisch aufgearbeiteten Form der *84 Handlungsmöglichkeiten* zu entfalten.

Verhaltensauffälligkeiten nach Beendigung des Treatments

In Hypothese 1d wurde vermutet, dass sich die Verringerung der Verhaltensauffälligkeiten auch nach Beendigung des *Treatments* zeigt und dass der positive Effekt vier Monate später noch anhält. Die Hypothese kann bestätigt werden. Die Verhaltensauffälligkeiten haben sich im Prä-/Post- und Prä-/Follow-up-Mittelwertvergleich signifikant reduziert und somit auch nachhaltig verändert. Dies gilt insbesondere für die Bereiche Verhaltensprobleme und Verhaltensprobleme mit Gleichaltrigen (s. Tabelle 27).

Hauptergebnis 3: Die Verhaltensauffälligkeiten reduzieren sich nach dem *Treatment* und der positive Effekt ist auch nach vier Monaten noch feststellbar.

Die Ergebnisse machen eine deutliche und anhaltende Reduzierung der Verhaltensauffälligkeiten der Schüler*innen offenbar. Von den bereits diskutierten Einschränkungen abgesehen, stützen die Ergebnisse des Prä-Post-Follow-up-Vergleichs demnach die Hauptergebnisse 1 und 2. Die Merkmalsveränderungen bei den Schüler*innen können damit mittels dreier Variablen (spezV,

nrV und V; vgl. Kap. 4.2.2) festgestellt werden und zeigen sich hier auf vielfältige Weise: abrupt und sprunghaft, kontinuierlich und anhaltend sowie auch noch vier Monate nach Beendigung der Intervention. Dies stärkt die These des kausalen Zusammenhangs zwischen der Anwendung der *84 Handlungsmöglichkeiten* und den positiven Schüler*innenoutcomes. Inwiefern die Lehrkräfte die *84 Handlungsmöglichkeiten* nach Beendigung der Intervention weiter unterrichtsintegriert genutzt haben, kann anhand der Daten jedoch nicht beurteilt werden.

Angenommen werden kann außerdem, dass die Professionalisierungsmaßnahme als solche einen positiven Einfluss auf die Schüler*innenoutcomes gehabt haben könnte. Angelehnt an die nach Durlak et al. (2011) identifizierten effektiven schulbasierten Präventionsmaßnahmen – die Etablierung positiver, sicherer Lernumgebungen und die gezielte Förderung sozial-emotionaler Fertigkeiten in der Schule – wurden in den Fortbildungen zusammen mit den Lehrkräften Maßnahmen auf universeller und selektiver Präventionsebene (Gresham et al., 2005) erarbeitet. Das auf diese Weise erworbene und erweiterte sonderpädagogisch-psychologische Spezialwissen für den inklusiven Unterricht im Bereich ESE (König et al., 2019; Gottfried et al., 2021) sollte durch die direkte Anwendung im Unterricht in routiniertes Lehrkrafthandeln überführt werden. Das Hauptergebnis 3 zu den Schüler*innenoutcomes deutet darauf hin, dass das auf diese Weise erworbene Handlungswissen, selbst nach Beendigung der Professionalisierungsmaßnahme, für die Lehrkräfte handlungsleitend gewesen sein könnte und dass die professionelle Handlungskompetenz, insbesondere im Bereich der situationsspezifischen Skills (Blömeke et al., 2015), erweitert werden konnte. Somit würden die Schüler*innenoutcomes indirekt auf ein verändertes Lehrkrafthandeln hindeuten – Effekte bezüglich eines positiven Transfers dieser Wirkungsebenen (Guskey, 2014; King, 2014; Lipowsky, 2010, 2014) konnten in Studien bereits nachgewiesen werden (Möller et al., 2006; Timperley et al., 2007). Um diese Annahme belegen zu können, wären allerdings in zukünftigen Studien Untersuchungen des konkreten Lehrkrafthandelns im Unterricht anzustreben, etwa anhand von Unterrichtsbeobachtungen und/oder -videographien.

Das Ergebnis stützt außerdem den Forschungsstand zur Dauer und Länge von Lehrkraftfortbildungen. Substantielle Veränderungen des unterrichtlichen Handelns (Darling-Hammond et al., 2017) und auf Schüler*innenebene (Timperley et al., 2007; Yoon et al., 2007) sind vorrangig bei Formaten zu konstatieren, die mindestens 14 Stunden umfassen und sich mindestens über ein halbes Schuljahr erstrecken. Bezüglich der Veränderung von Lehrkraft-Einstellungen geht man von der Notwendigkeit noch längerer Formate aus, die sich mitunter über mehrere Schuljahre erstrecken können (Rzejak & Lipowsky, 2019; Leidig, 2019). Die vorliegenden Ergebnisse unterstreichen die Bedeutung langfristiger Professionalisierungsformate, die insbesondere für inklusive Beschulung im FSP ESE äußerst relevant erscheinen.

Studien zeigen ferner, dass die Erweiterung des Lehrkraftwissens auch mit Veränderungen der Überzeugungen und subjektiven Theorien einhergehen kann (Copur-Gencturk & Thacker, 2020; Goldschmidt & Phelps, 2007). Speziell der Erwerb deklarativen Wissens über sonderpädagogische Förderschwerpunkte und die Erfahrung mit inklusiver Pädagogik können die Einstellungen von Lehrkräften positiv beeinflussen (de Boer et al., 2011; Qi & Ha, 2012). Es kann daher sein, dass sich durch das erweiterte pädagogisch-psychologische Handlungswissen im Bereich ESE der Blick der Lehrkräfte auf die Schüler*innen mit Verhaltensauffälligkeiten und emotional-sozialen Entwicklungsstörungen verändert hat. Dieser, aus Inklusionsorientierter und pädagogischer Sicht durchaus zu begrüßende Effekt (Hecht et al., 2016; Van Mieghem et al., 2020), könnte in der vorliegenden Studie das Ergebnis der Schüler*innenoutcomes dahingehend beeinflusst haben, dass die Lehrkräfte das Schüler*innenverhalten nach der Fortbildungsmaßnahme als „weniger auffällig“ einschätzten als sie dies zu deren Beginn taten. Um Merkmalsänderungen bei Schüler*innen verlässlich abbilden zu können und einer möglichen *mangelnden instrumentellen Reliabilität* vorzubeugen, ergibt sich für zukünftige Studien daher die Notwendigkeit einer externen Beurteilung der Schüler*innenoutcomes.

Prosoziales Verhalten nach Beendigung des Treatments

Hypothese 1e, welche davon ausging, dass sich das prosoziale Verhalten nach dem *Treatment* verbessert und der Effekt vier Monate später noch feststellbar ist, kann bestätigt werden. Das prosoziale Verhalten der Schüler*innen hat sich über die Zeit positiv verändert (s. Tabelle 28). Signifikante Mittelwertunterschiede zeigen sich bei einem Prä-Post- und Prä-Follow-up-Test-Vergleich. Das prosoziale Verhalten der Schüler*innen war somit nach der Intervention und vier Monate später stärker ausgeprägt als vor der Intervention.

Hauptergebnis 4: Das prosoziale Verhalten verbessert sich nach dem *Treatment* und der Effekt ist auch nach vier Monaten noch feststellbar.

Dieses Ergebnis stützt die vorherigen Hauptergebnisse 1, 2 und 3, wenn auch hier die zuvor formulierten Einschränkungen der internen Validität bei der Interpretation beachtet werden müssen. Es ist sehr erfreulich, dass neben der Reduzierung der Verhaltensauffälligkeiten auch die nachhaltige Stärkung des prosozialen Verhaltens der Schüler*innen sichtbar wird. Dies lässt vermuten, dass die eingesetzten Handlungsmöglichkeiten neben dem Abbau von Verhaltensauffälligkeiten auch die prosozialen Fähig- und Fertigkeiten der Schüler*innen fördern. Wie die Forschung zeigt, hängen Emotionsregulationsstrategien und prosoziales Verhalten eng miteinander zusammen. Schwierigkeiten in der Beziehung zu Gleichaltrigen haben sich dabei als hervorstechen-

der und wichtiger Prädiktor für die psychische Gesundheit und Verhaltensanpassung erwiesen (Koglin et al., 2013; Silver et al., 2010). Kinder mit gestörten Peer-Beziehungen haben ein erhöhtes Risiko, dass sich ihre Probleme in verschiedenen Lebensbereichen verstärken (Dodge & Pettit, 2003). Im Gegensatz dazu scheinen adaptive Peer-Beziehungen, etwa durch die Entwicklung positiver sozialer Beziehungen, einen Schutzfaktor für diese Kinder darzustellen (Castello, 2017). Insbesondere kognitionspsychologisch fundierte Handlungsansätze für den Bereich ESE sehen daher – neben der Differenzierung sozialer Wahrnehmung, dem Bewusstmachen von Emotionen und der Erweiterung des Repertoires zur emotionalen Steuerung – einen zentralen Förderschwerpunkt im Erlernen prosozialer Handlungsalternativen (Blumenthal et al. 2020, S. 29). Bei Kindern mit einem erhöhten Maß an Verhaltensproblemen sind Programme, die ausschließlich auf die Verringerung des problematischen Verhaltens abzielen und dabei die Entwicklung prosozialer Fähigkeiten unberücksichtigt lassen, als unzureichend zu bewerten (Andrade et al., 2014). Für die präventive universelle und selektive Förderung des Verhaltens im inklusiven Unterricht ergibt sich stattdessen der Auftrag, Übungen und Trainings zu prosozialem Verhalten noch mehr Beachtung zu schenken und die Etablierung positiver Peer-Beziehungen gezielt in den Fokus zu nehmen.

6.2.2 Lehrkrafthandeln

Einsatzhäufigkeit und Interaktionseffekte der 84 Handlungsmöglichkeiten

Die *84 Handlungsmöglichkeiten* wurden laut Rückmeldung der Lehrkräfte im Mittel an 67 % der Ratingtage eingesetzt, was bedeutet, dass der Einsatz oft erfolgte. Die Hypothese 2a kann somit bestätigt werden, wobei die Einsatzhäufigkeit die Annahme sogar noch um über 20 % übersteigt. Betrachtet man die Einzelfälle, so ist bei dem Großteil der Schüler*innen ($n = 11$) eine sehr hohe (80-100 %) Einsatzfrequenz zu verzeichnen, bei $n = 5$ Schüler*innen erfolgte der Einsatz oft (60-80 %), bei $n = 7$ Schüler*innen wurden die Handlungsmöglichkeiten gelegentlich (40-60 %) eingesetzt und bei wenigen Schüler*innen ($n = 4$) wurde eher selten (20-40 %) von der neuen Methode Gebrauch gemacht. Ein Wert < 20 % kam nicht vor (s. Abbildung 19).

Die Hypothese 2b, bei der angenommen wurde, dass sich ein positiver Interaktionseffekt zwischen der Einsatzhäufigkeit der *84 Handlungsmöglichkeiten* und den Schüler*innenoutcomes zeigt, kann bestätigt werden, bedarf indes einer differenzierten Betrachtung. Werden die Handlungsmöglichkeiten wenigstens an 40 % der Ratingtage eingesetzt, so wirkt sich dies positiv auf die Schüler*innenoutcomes aus. Finden die Handlungsmöglichkeiten an weniger als 40 % der Ratingtage Anwendung, wird demgegenüber keine positi-

ve Wirkung auf das Schüler*innenverhalten sichtbar (s. Tabelle 29 und Tabelle 30). Bei der genaueren Analyse der Interaktionseffekte ergeben sich jedoch keine stringent positiven Zusammenhänge: Es können zwar positive Interaktionseffekte bei sehr häufigem, häufigem und gelegentlichem Einsatz der Handlungsmöglichkeiten festgestellt werden, die größten Level-Effekte treten jedoch bei den letzten beiden Einsatzhäufigkeiten zu Tage. Werden die Handlungsmöglichkeiten selten eingesetzt, zeigt sich nahezu keine Levelbewegung (s. Tabelle 30) bzw. bei den spezifischen Verhaltensauffälligkeiten sogar eine Verschlechterung des Verhaltensverlaufs.

Hauptergebnis 5: Die Lehrkräfte setzen die Intervention im Mittel an 67 % der Interventionstage ein. Dabei zeigt sich ein grundsätzlich positiver Interventionseffekt zwischen der Einsatzhäufigkeit der *84 Handlungsmöglichkeiten* und den Schüler*innenoutcomes: Werden die Handlungsmöglichkeiten öfter eingesetzt, so wirkt sich dies positiv auf die Schüler*innenoutcomes aus. Finden die Handlungsmöglichkeiten seltener Anwendung, wird keine positive Veränderung auf das Schüler*innenverhalten sichtbar. Die größte Wirkung scheint sich bei einem gelegentlichen und häufigen Einsatz zu entfalten.

Die Interaktionseffekte zwischen der Einsatzhäufigkeit der *84 Handlungsmöglichkeiten* und den Schüler*innenoutcomes sind zunächst vor dem Hintergrund der Spezifität des Fördermanuals zu diskutieren. Die *84 Handlungsmöglichkeiten* enthalten lerntheoretisch und kognitionspsychologisch begründete Handlungsmöglichkeiten sowie solche, die sich am schüler*innenzentrierten Ansatz orientieren. In der vorliegenden Studie kamen bei der Mehrheit der Kinder ($n = 15$) ausschließlich lerntheoretische Handlungsmöglichkeiten zum Einsatz (s. Abbildung 19). Bei weiteren $n = 6$ Schüler*innen entschieden sich die Lehrkräfte für eine kombinierte Förderung mit lerntheoretischem und kognitionspsychologischem/schüler*innenzentriertem Erklärungs- und Handlungsansatz. Bei lediglich $n = 6$ Schüler*innen kamen ausschließlich kognitionspsychologische und/oder schüler*innenzentrierte Handlungsmöglichkeiten zum Einsatz (s. Abbildung 18 und Abbildung 19). Damit überwiegen in der Anwendung der *84 Handlungsmöglichkeiten* ganz klar die Interventionen mit lerntheoretischem Erklärungs- und Handlungsansatz. Gerade lerntheoretisch ausgerichtete Fördermethoden, die darauf abzielen, Verhalten durch äußere Reize der positiven und negativen Verstärkung zu verändern, sind auf Repetition und Regelmäßigkeit angewiesen (Kap. 2.1.3.2). Dies könnte eine Erklärung dafür sein, warum sich keine Verhaltensänderungen zeigten, wenn die Handlungsmöglichkeiten im Durchschnitt nicht an wenigstens 40 % der Interventionstage und damit tendenziell mindestens an jedem dritten Tag eingesetzt wurden. Eine weitere, etwas anders gelagerte Erklärung für die erkennbaren Interaktionseffekte zwischen Schüler*innenoutcomes und Einsatzhäufigkeit ist, dass der Interaktionseffekt auch etwas

über die Einsatzqualität aussagen könnte. Die seltene, an weniger als 40 % der Interventionstage stattgefundene, Anwendung könnte für eine inkonsequente oder ungünstige Umsetzung der jeweiligen Handlungsmöglichkeiten sprechen und damit die niedrigen bzw. nicht vorhandenen Level-Effekte erklären. Da die Qualität der Umsetzung in der vorliegenden Studie allerdings nicht untersucht wurde, sind dies lediglich Mutmaßungen. Betrachtet man die Ergebnisse der Einsatzhäufigkeit und der Schüler*innenoutcomes über die gesamte Gruppe, so kann geschlussfolgert werden, dass insbesondere die verhaltenstheoretisch ausgerichtete Förderung bei den *84 Handlungsmöglichkeiten* an mindestens 40 % der Interventionstage eingesetzt werden sollte, um ihre Wirksamkeit entfalten zu können.

Die gute Implementation eines schulischen Präventionsprogramms stellt einen zentralen Indikator für die Wirkung präventiver Förderung dar (Hennemann et al. 2017, S. 57ff.). Diese Schlussfolgerung, die anhand der Zusammenfassung aus unterschiedlichen Metaanalysen (Beelmann, 2008; Durlak et al., 2011; Sklad et al., 2012), systemischen Reviews (Brezinka, 2003; Casale et al., 2014; Heinrichs et al., 2002) und der Zusammenfassung zu den Kriterien effizienter SEL-Ansätze (Denham & Burton, 2003; Reicher & Jauk, 2012) gezogen werden kann, bestätigt sich in der vorliegenden Studie. In zukünftigen Studien sollten die konkreten Implementationsbedingungen, etwa die Einsatzhäufigkeit und -qualität, der unterschiedlich ausgerichteten Handlungsmöglichkeiten deutlich mehr beachtet werden.

6.2.3 Lehrkraftkognitionen

*Selbstwirksamkeitsempfinden der Lehrkräfte nach Beendigung des Treatments sowie Interaktionseffekte mit den Schüler*innenoutcomes*

Forschungshypothese 3a kann teilweise bestätigt werden. So hat sich das Selbstwirksamkeitsempfinden der Lehrkräfte nach dem Treatment signifikant verbessert, vier Monate später hält der Effekt allerdings nicht mehr an (s. Tabelle 31).

Betrachtet man die Subgruppen, wird deutlich, dass sich vor allem Lehrkräfte, die keine Beratung erhalten hatten, nach Beendigung der Professionalisierungsmaßnahme selbstwirksamer einschätzten. Bei den Lehrkräften, die eine Beratung erhalten haben, ist keine Veränderung der Mittelwerte über die Zeit zu verzeichnen. Die Hypothese 3b muss somit verworfen werden.

Hypothese 3c, die beinhaltete, dass sich ein positiver Interaktionseffekt zwischen dem Selbstwirksamkeitsempfinden der Lehrkräfte und den Schüler*innenoutcomes zeigt, kann nur teilweise bestätigt werden. Während beim Regelverhalten der Schüler*innen kein positiver Interaktionseffekt zwischen Selbstwirksamkeitsempfinden und Schüler*innenoutcomes erkennbar ist

(s. Tabelle 33), lässt sich dieser bei den spezifischen Verhaltensauffälligkeiten der Schüler*innen feststellen (s. Tabelle 32). Je höher hier das Selbstwirksamkeitsempfinden der Lehrkräfte ausfiel, desto größer waren auch die Verhaltensverbesserungen auf Schüler*innenseite.

Hauptergebnis 6: Über die Gesamtgruppe gesehen verbessert sich das Selbstwirksamkeitsempfinden der Lehrkräfte nach dem *Treatment*. Der positive Effekt ist allerdings nach vier Monaten nicht mehr feststellbar. Anders als vermutet zeigen gerade die Lehrkräfte, die eine kooperativ-systemische Fallberatung erhalten haben, unmittelbar nach dem *Treatment* keine höhere Selbstwirksamkeit.

Während sich zwischen dem Selbstwirksamkeitsempfinden der Lehrkräfte und dem Regelverhalten kein positiver Interaktionseffekt zeigt, wird dieser bei den spezifischen Verhaltensauffälligkeiten der Schüler*innen sichtbar. Ein höheres Selbstwirksamkeitsempfinden geht hier mit besseren Schüler*innenoutcomes einher.

Das erhöhte Selbstwirksamkeitsempfinden der Lehrkräfte unmittelbar nach Durchführung des Treatments kann als erfolgreiche Weiterentwicklung der spezifischen Lehrkraftkompetenzen im Bereich pädagogisch-psychologischen Wissens (Baumert & Kunter, 2006; s. Abbildung 7) und damit als positiver Effekt der Professionalisierungsmaßnahme und seiner Elemente gedeutet werden. Konkrete Störfaktoren (s. Kap. 6.2.1) könnten das Ergebnis jedoch verzerren und sollen hier betrachtet werden.

Obwohl die Lehrkräfte gebeten wurden, neben der Professionalisierungsmaßnahme keine weiteren Fortbildungen zu besuchen, können *zeitliche Einflüsse* dieser Art nicht kategorisch ausgeschlossen werden. Zieht man jedoch die Arbeitsbelastung der Pädagog*innen in die Betrachtung mit ein²¹, ist zumindest stark anzuzweifeln, dass diese neben dem Treatment weitere Professionalisierungsmaßnahmen wahrgenommen haben. Gänzlich ausgeschlossen werden kann außerdem nicht, dass die Lehrkräfte auch ohne Teilnahme an der Professionalisierungsmaßnahme selbstwirksamer geworden wären (*Reifungsprozesse*). Da es sich bei der Untersuchungsgruppe im Durchschnitt um sehr erfahrene Lehrkräfte handelt – das Durchschnittsalter lag bei 47 Jahren ($SD = 9.7$), das Durchschnittsschuldienstjahr bei 22 ($SD = 19.09$) – kann angenommen werden, dass sie sich bereits im Status der kompetent Handelnden bzw. Profilierten (Berliner, 2006; König, 2010; vgl. Kap. 2.2.2.2) befanden. Dies spricht dafür, dass die Lehrkräfte schon vor Beginn des Treatments ein hohes Maß an Expertise aufwiesen – anzunehmen sind etwa (sonder-)pädagogisches Basiswissen über effektive Klassenführung, die Orchestrierung des Lernprozesses, allgemeine Unterrichtsmethoden sowie

21 In Mecklenburg-Vorpommern liegt die Unterrichtsverpflichtung von Lehrkräften an Grundschulen bei 27,5 Stunden pro Woche (Ministerium für Bildung und Kindertagesförderung, 2022b).

verschiedene Aspekte der Diagnostik, Leistungsbeurteilung und Heterogenität (Baumert & Kunter, 2006; Gottfried et al., 2021; König et al., 2019) – und dieses somit nicht erst erwerben mussten. Gestützt wird die These durch das bereits vor Beginn des Treatments überdurchschnittlich hoch ausgeprägte Selbstwirksamkeitsempfinden (s. Kap. 5.1.3). Nicht ausgeschlossen werden kann ferner, dass es sich bei der Untersuchungsgruppe um besonders motivierte und engagierte Lehrkräfte handelte (*experimentelle Mortalität*), denen es gegebenenfalls besser gelang, ihr Selbstwirksamkeitsempfinden durch die Fortbildungsmaßnahmen zu stärken. Gleichwohl bleibt festzuhalten: Dass die Lehrkräfte nach der Teilnahme an der Professionalisierungsmaßnahme ihr Selbstwirksamkeitsempfinden (nochmals) steigern konnten (s. Tabelle 31) zeigt nicht nur, dass sich die affektiv-emotionalen Lehrkraftkognitionen (Lipowsky, 2010, 2014) über die Zeit verändert haben. Da man von einem engen Zusammenhang zwischen affektiv-emotionalen Kognitionen und deklarativ-prozeduralem Wissen ausgeht (Guskey, 2014; King, 2014; Lipowsky, 2010, 2014), kann außerdem angenommen werden, dass die Lehrkräfte spezifische Kompetenzen in Form (sonder-)pädagogischen Spezialwissens (Gottfried et al., 2021; König et al., 2019; Leko & Roberts, 2014; Osher et al., 2014; vgl. Kap. 2.2.2.1) erworben haben und diese im Sinne situationsspezifischer Skills (Blömeke et al., 2015) im Unterricht einzusetzen wussten. Da beide Aspekte das konkrete unterrichtliche Handeln, insbesondere im FSP ESE, stark beeinflussen (Van Mieghem et al. 2020, S. 6), stellt dieses Ergebnis ein wichtiges Indiz für die Effektivität des entwickelten Professionalisierungskonzepts dar.

Vier Monate nach Beendigung des Treatments hielt der positive Effekt auf das Selbstwirksamkeitsempfinden jedoch nicht mehr an. Geht man von einem positiven Zusammenhang zwischen Lehrkraftkognitionen und Schüler*innenoutcomes aus (Möller et al., 2006; Timperley et al., 2007), verwundert dieses Ergebnis, schätzten die Lehrkräfte doch auch sechs Monate nach Beendigung der Professionalisierungsmaßnahme das Schüler*innenverhalten als signifikant besser ein als zu Beginn der Studie. Als eine mögliche Erklärung könnten Deckeneffekte herangezogen werden. Dies würde bedeuten, dass die eingesetzte *Skala Lehrer-Selbstwirksamkeitserwartung* (Schwarzer & Schmitz, 1999) im oberen Leistungsbereich nicht ausreichend zwischen verschiedenen Selbstwirksamkeitsniveaus unterscheidet und dadurch extrem hohe Messwerte sich nicht mehr vergrößern können (Bortz & Döring, 2016), ein Umstand, der durch die bereits zu Beginn der Studie überdurchschnittlich hohe Selbstwirksamkeit der Lehrkräfte (s. Kap. 5.1.3 und Tabelle 31) begünstigt würde. Die Beendigung der Professionalisierungsmaßnahme könnte eine weitere Erklärung liefern. Gerade durch die universitäre Anbindung und Anleitung könnten die Lehrkräfte sich gut begleitet und unterstützt gefühlt haben. Studien weisen auf einen entsprechenden positiven Zusammenhang zwischen universitär angebundenen Fortbildungsformaten und dem affektiv-

motivationalen Befinden teilnehmender Lehrkräfte hin (Avramidis et al., 2000; Leidig et al., 2016). Der Rahmen eines Professionalisierungskonzepts könnte sie somit im besonderen Maße zu einer hohen Umsetzungstreue und damit dem konsequenten Einsatz der Fördermaßnahmen angeregt haben.

Betrachtet man die Subgruppen, so wird ersichtlich, dass gerade die Lehrkräfte, die eine Beratung erhalten haben, ihr Selbstwirksamkeitsempfinden nicht steigern konnten, wohingegen dies bei der Gruppe ohne Beratung signifikant der Fall war. Dieses Ergebnis kann auf unterschiedliche Weise interpretiert werden. Zum einen könnten die Lehrkräfte, die beraten wurden, sich systematisch von denen ohne Beratung unterscheiden, obgleich erstere zufällig ausgewählt worden sind (vgl. Kap. 3.2.3). Zum anderen könnte die Beratung selbst eine eher abschwächende Wirkung auf das Selbstwirksamkeitsempfinden der Lehrkräfte gehabt haben. Dies könnte sowohl an der Art und Weise der Beratung liegen als auch an deren organisatorischer Durchführung. Ursprünglich war geplant, wenigstens zwei Beratungsgespräche pro Lehrkraft durchzuführen. Dies war jedoch durch die Kürze des zur Verfügung stehenden Zeitraumes nicht möglich, wodurch mit allen ausgewählten Pädagog*innen nur ein Beratungsgespräch realisiert werden konnte. Die Reflexionen folgten weitestgehend der Struktur der Kooperativen Beratung nach Mutzeck (2008; vgl. Kap. 3.2.3.1), wobei die Beraterin als systemisch ausgebildete Therapeutin nicht nur weitere Gesprächsmethoden einbrachte, sondern auch den Blick auf den sozialen Kontext ermöglichte. In der überwiegenden Mehrheit der Beratungssitzungen ging es um Schüler*innen mit herausforderndem Verhalten. In zwei Gesprächen wurden aber auch persönliche Themen, wie private Ängste und Sorgen bearbeitet. Zwei Gespräche wurden vornehmlich zu anderen schulischen Themen geführt, wie der Stellung im Team, den Herausforderungen im Umgang mit der Leitung und der Identität als Lehrkraft. Die systemische Vorbildung und -erfahrung im therapeutischen Setting wurden dabei von der Beraterin als äußerst hilfreich wahrgenommen. Möglicherweise hat aber gerade dieser Umstand dazu geführt, dass auch private Themen thematisiert wurden. Eventuell hätten die Lehrkräfte die Gespräche genauso hilfreich empfunden, wenn nur kooperativ und mit ausschließlicher Fokussierung auf problematische Schulsituationen (Mutzeck, 2008) beraten worden wäre. Gegebenenfalls haben die außerunterrichtlichen bzw. gar privaten Themen eher zu einem Aufdecken tiefer liegender Probleme geführt, die im Rahmen einer Beratungssitzung nicht aufzufangen waren. Wirksame Fortbildungsmaßnahmen, die Feedback und Beratung durch Fortbildner*innen und/oder Kolleg*innen vorsehen, fokussieren hingegen auf die konkrete Anwendung der Fortbildungsinhalte und den daraus resultierenden Erfahrungen der Lehrpersonen (Leidig et al., 2016; Lipowsky & Rzejak, 2019; Saxe et al., 2001).

Eine weitere Annahme der Studie war, dass ein positiver Zusammenhang zwischen dem Lehrkraft-Selbstwirksamkeitsempfinden und den Schüler*in-

nenoutcomes besteht. Diese aus der Theorie hergeleitete Hypothese stützt sich auf zwei Perspektiven. Danach wird einerseits angenommen, dass eine erhöhte Selbstwirksamkeit Ausdruck von wahrgenommenen Verhaltensbesserungen der Schüler*innen sein kann (Lipowsky & Rzejak 2012, S. 57). Andererseits wird erhöhtes Selbstwirksamkeitsempfinden auch als ein Indiz für kompetentes Lehrkraft Handeln gesehen, welches wiederum positives Schüler*innenverhalten auslösen kann (Van Mieghem et al., 2020). Die vorliegenden Ergebnisse können diese Annahmen nur teilweise bestätigen. Eine Erklärung für die Diskrepanz zwischen vorhandenen und nicht vorhandenen Interaktionseffekten von Selbstwirksamkeitsempfinden und Schüler*innenoutcomes könnte die Spezifität und Art und Weise der Herleitung der Items sein. Während den Lehrkräften das Item zur Messung des Regelverhaltens vorgegeben wurde, waren sie bei der Operationalisierung der Items für spezifische Verhaltensauffälligkeiten aktiv mit eingebunden. Sie hatten damit ein genaues Bild davon, welches Verhalten sich hinter dem operationalisierten Item befindet. Dies führt nachweislich zu zuverlässigeren Einschätzungen durch Lehrkräfte (Conley et al., 2014). Daneben ist das Regelverhalten durch ein globales, übergreifendes Item – eine Single Item Scale – operationalisiert („Hält sich nicht an Absprachen und allgemeine Regeln“), welches bei der Beurteilung ein hohes Ausmaß an Schlussfolgerungen auf Seiten der Lehrkräfte erfordert (Lotz et al., 2013). Die spezifischen Verhaltensauffälligkeiten wurden hingegen durch drei konkrete, niedrig-inferente Items (z.B. „redet/ruft dazwischen“, „stört Mitschüler*innen beim Arbeiten“, „ist nicht bereit für den Unterricht“) operationalisiert. Beide Aspekte könnten eine Erklärung dafür bieten, warum nur zwischen den spezifischen Verhaltensauffälligkeiten und dem Selbstwirksamkeitsempfinden der Lehrkräfte Interaktionseffekte sichtbar werden: Die Lehrkräfte konnten die Verhaltensänderungen der Schüler*innen hier eventuell besser beurteilen bzw. bezogen die Veränderungen innerhalb dieser Verhaltensdimension konkreter auf ihr unterrichtliches Lehrkraft Handeln. Dieses Ergebnis kann einen Hinweis dafür bieten, dass durch Lehrkräfte operationalisierte niedrig-inferente DBR-Items anhand einer Multi-Item-Scale – neben der zuverlässigeren Messung des interessierenden Merkmals (Casale et al., 2017) – das Selbstwirksamkeitsempfinden von Lehrkräften moderieren und positiv beeinflussen. Zukünftige Studien könnten diesen Aspekt gezielt in den Blick nehmen.

Aufgrund ihres handlungsleitenden und -regulierenden Charakters wird angenommen, dass die inklusive Einstellung von Lehrpersonen die Implementationsqualität eines inklusiven Bildungssystems maßgeblich beeinflusst (de Boer et al., 2011). Zukünftige Studien im Feld der Professionalisierung von Lehrkräften sollten daher die Zusammenhänge zwischen pädagogisch-psychologischem Wissen und übergeordneten affektiv-motivationalen Aspekten und Dispositionen, wie beispielsweise den Einstellungen zur Inklusion und der Selbstwirksamkeitserwartung in inklusiven Kontexten, noch

genauer herausarbeiten (Gottfried et al., 2021; Hecht et al., 2016). Neben der Erfassung der Selbstwirksamkeitserwartung der Lehrkräfte (Schwarzer & Schmitz, 1999) wäre es in Effektstudien zur Professionalisierung von Lehrkräften vorteilhaft, noch dezidiertere Aussagen über die Wahrscheinlichkeit einer zukünftigen inklusiven Unterrichtsgestaltung treffen zu können. In der Inklusionsforschung hat sich hierfür etwa die *Theorie des geplanten Verhaltens* (Ajzen, 1985) als aussagekräftig erwiesen (Ahmmed et al., 2016; MacFarlane & Woolfson, 2013). Diese postuliert, dass ein Verhalten umso wahrscheinlicher verwirklicht wird, je größer die *Intention* ist, dieses Verhalten zu zeigen (Ajzen, 1985). Die Intention einer Person, ein Verhalten auszuführen, ist wiederum von drei Determinanten abhängig: der Einstellung zum Verhalten, der subjektiven Norm und der wahrgenommenen Verhaltenskontrolle. Letztere bezeichnet die wahrgenommene Überzeugung, ein Verhalten ausführen zu können – sie ist somit kompatibel mit dem Konzept der Selbstwirksamkeitserwartung von Bandura. Die Einstellung zum Verhalten bezieht sich auf das Ausmaß, in dem eine Person ein Verhalten und die damit einhergehenden Konsequenzen als positiv oder negativ beurteilt. Die subjektive Norm ist definiert als wahrgenommener sozialer Druck, ein Verhalten auszuführen. Sie bezieht sich auf die Wahrnehmung einer Person, wie bedeutsame andere Personen das gezeigte Verhalten bewerten und auf die Motivation, im Sinne dieser Bewertung zu handeln (Ajzen, 1987). Bezogen auf das geplante Verhalten von Lehrkräften kann so die Schulleitung eine bedeutsame Größe darstellen, die einen Einfluss auf die Einstellungen von Lehrkräften haben kann (Urton et al., 2014). Verschiedene Studien konnten die wahrgenommene Einstellung der Schulleitung, operationalisiert als subjektive Norm, als signifikanten Prädiktor für das inklusive Verhalten von Lehrkräften nachweisen (Ahmmed et al., 2014; MacFarlane & Woolfson, 2013). Nach Ajzen (1987) ist die *Intention*, ein Verhalten auszuführen, umso größer, je positiver die Einstellung, je stärker die subjektive Norm und je größer die wahrgenommene Verhaltenskontrolle sind. Für zukünftige Professionalisierungsbestrebungen im Kontext von Inklusion stellt dies eindeutig die Relevanz der Einbindung der Schulleitungen heraus.

6.2.4 Lehrkraftreaktionen

Zufriedenheit und Nützlichkeitseinschätzung bezüglich der Professionalisierungsmaßnahme und seiner Elemente

Die Forschungshypothese 4a kann bestätigt werden: Die Lehrkräfte haben sich während der gesamten Professionalisierungsmaßnahme gut betreut gefühlt und betrachten das Gesamtkonzept als nützlich (s. Tabelle 34). Dies zeigt sich neben dem geschlossenen auch im offenen Antwortteil, bei dem zum Mehrwert der Professionalisierungsmaßnahme fast doppelt so viele

Antworten eingingen wie zum Themenkomplex Verbesserungsvorschläge. Zu ersterem wurden neben allgemeinen Zufriedenheitsbekundungen und Nützlichkeitseinschätzungen auch die Erweiterung des pädagogischen Wissens, die Veränderung der eigenen Einstellungen und Überzeugungen, die Veränderung des unterrichtspraktischen Handelns sowie die Initiierung von Reflexionsprozessen als Mehrwert hervorgehoben. Ganz konkret wurde der Erhalt und Einsatz des Buches *84 Handlungsmöglichkeiten* und die stärkere Auseinandersetzung mit einzelnen Schüler*innen positiv gewürdigt. Bei den Verbesserungsvorschlägen gingen vor allem kritische Anmerkungen zur Durchführung des Verhaltensmonitorings ein, insbesondere der Wunsch nach einer Ausdehnung des Beurteilungszeitraumes wurde oft geäußert. Auch hätten die Lehrkräfte sich bei den Fortbildungen mehr Trainings- und Praxisbeispiele/Übungen gewünscht. Für eine zukünftige Durchführung wurde zudem angeregt, Fremdhospitationen zu organisieren und den Austausch zwischen Kolleg*innen systematisch zu initiieren und einzuplanen.

Forschungshypothese 4b kann teilweise bestätigt werden: Die Lehrkräfte sind mit der Durchführung der einzelnen Professionalisierungselemente mehrheitlich zufrieden und erachten diese als nützlich für ihre tägliche Arbeit. Die Fortbildungen schätzten die Lehrkräfte im Mittel als hilfreich und deren Länge als angemessen ein (s. Tabelle 35). Das Verhaltensmonitoring mittels DBR wurde etwas weniger positiv eingeschätzt. Obwohl die Nützlichkeitseinschätzung dieses Instruments sowie die Einschätzung zu dessen Durchführbar- und Praxistauglichkeit als befriedigend zu interpretieren sind, ist die Wahrscheinlichkeit, dass die Lehrkräfte es in Zukunft einsetzen, nur ausreichend hoch. Die Wahrscheinlichkeit, dass die Lehrkräfte auch in Zukunft freiwillige Beratungen in Anspruch nehmen werden, ist hingegen tendenziell hoch (s. Tabelle 36). Positive Effekte der Beratung auf die Emotions- und Stressregulierung sowie die Sicherheit in der Klasse schätzten die Lehrkräfte als befriedigend ein. Die Zufriedenheit mit der Beratungssitzung zeigte sich auch im offenen Antwortteil. Gut gefallen hat den Lehrkräften vor allem die Atmosphäre sowie der Stil während der Beratung. Als nützlich für die Zukunft empfanden sie das Kennenlernen von Methoden der kooperativ-systemischen Gesprächsführung, den Gewinn neuer Erkenntnisse und ein erhöhtes Selbstvertrauen. Für die Zukunft wünschen sich die Lehrkräfte mehr professionelle Unterstützung und die stärkere Einbindung von Eltern sowie für sich selbst eine höhere Professionalität.

Hauptergebnis 7: Die Lehrkräfte sind mit der gesamten Professionalisierungsmaßnahme und der Durchführung der einzelnen Professionalisierungselemente zufrieden und erachten diese als nützlich für ihre tägliche Arbeit. Besonders positiv wurde die Arbeit mit dem Manual *84 Handlungsmöglichkeiten* und die *psychosozial-horizontale Beratung* eingeschätzt. Der Einsatz von Verhaltensmonitoring mittels DBR wurde hingegen nur als zufriedenstellend bis ausreichend bewertet.

Obwohl die Ergebnisse aus der Befragung der Lehrkräfte als tentativ und nicht generalisierbar einzustufen sind, ermöglichen sie dennoch evaluative Schlussfolgerungen über die Akzeptanz und das Nützlichkeitsempfinden der Lehrkräfte und somit den Mehrwert des Professionalisierungskonzepts. Zugleich verweisen die Lehrkraftaussagen auf Defizite und notwendige Modifikationen und eröffnen damit Räume für die Weiterentwicklung der Professionalisierung im Bereich ESE.

Erfreulich ist zunächst, dass die Lehrkräfte die gesamte Professionalisierungsmaßnahme und die meisten der darin enthaltenen Professionalisierungselemente als sinnvoll und nützlich für ihre tägliche Arbeit einschätzen. Es kann angenommen werden, dass sich dieses grundsätzlich positive Ergebnis auf der Wirkungsebene der Lehrkraftreaktionen auch auf der Ebene der Lehrkraftkognitionen anhand des erhöhten Selbstwirksamkeitsempfindens widerspiegelt (s. Hauptergebnis 6). Die hohe Anzahl an Kommentaren zum Mehrwert der Professionalisierungsmaßnahme unterstreicht die Akzeptanz der Lehrkräfte, ohne die die Weiterentwicklung ihrer professionellen Handlungskompetenzen zumindest fraglich wäre (Lipowsky, 2010). Die hierbei identifizierbaren zentralen Elemente – die Erweiterung des eigenen pädagogischen Wissens, die Veränderung der eigenen Einstellungen und Überzeugungen, die Veränderung des unterrichtspraktischen Handelns sowie die Initiierung von Reflexionsprozessen (vgl. Kap. 5.1.4) – werden in der Forschung als zentrale Entwicklungsaufgaben professioneller Handlungskompetenz in Bezug auf die inklusive Beschulung im Bereich ESE angesehen (Gottfried et al., 2021; König et al., 2019; Leko & Roberts, 2014; Leidig et al., 2016; Osher et al., 2014; Van Mieghem et al., 2020). Der Umstand, dass die Lehrkräfte diese Punkte selbst benennen, ist bedeutsam und kann als ein weiterer Wirksamkeitshinweis der entwickelten Professionalisierungsmaßnahme interpretiert werden. Der bereits vermutete Zusammenhang zwischen der Erweiterung des Lehrkraftwissens und Veränderungen der Überzeugungen und subjektiven Theorien der Pädagog*innen (Copur-Gencturk & Thacker, 2020; Goldschmidt & Phelps, 2007) über Schüler*innen mit Verhaltensschwierigkeiten, wird ferner durch konkrete Lehrkraftaussagen gestützt: die Wahrnehmung einer positiven Veränderung der eigenen Einstellungen/Überzeugungen, intensivere Beschäftigung mit einzelnen Schüler*innen, neuer Blick auf sogenanntes „Problemverhalten“, mehr Verständnis für und weniger Vorurteile gegenüber Kinder/n mit Verhaltensauffälligkeiten sowie eine spürbare Stärkung des eigenen Selbstvertrauens und der Selbstwirksamkeit.

Grundlegende Vorschläge zur Umgestaltung der Professionalisierungsmaßnahme konzentrieren sich auf den Einsatz des Verhaltensmonitorings, die inhaltlich-didaktische Ausrichtung der Fortbildungen sowie die Erweiterung des Konzepts durch Elemente der Fremdhospitation und der Förderung des Austauschs zwischen Kolleg*innen. Da detailliertere Lehrkraftaussagen zu den Professionalisierungselementen Fortbildung und Verhaltensmonitoring

(vgl. Kap. 3.2) vorliegen und diese separat thematisiert werden, erfolgt zunächst die Diskussion der zuletzt aufgeworfenen Punkte. Der Wunsch einiger Lehrkräfte nach (weiterführenden) Fremdhospitationen und der Etablierung einer Gesprächs- und Austauschkultur in der Schule kann ein Hinweis dafür sein, dass das Bedürfnis nach weitergehender Professionalisierung für die schulische Förderung im Bereich ESE bei einigen Lehrkräften nach dem Durchlaufen der Professionalisierungsmaßnahme geweckt wurde. Die Kritik bietet außerdem einen Anhaltspunkt für die Diskussion der Nachhaltigkeit des Konzepts. Wie auf der Ebene der Lehrkraftkognitionen sichtbar wurde, bewegt sich das Selbstwirksamkeitsempfinden vier Monate nach Beendigung der Professionalisierungsmaßnahme fast wieder auf dem, absolut gesehen zwar hohen, aber dennoch niedrigeren Ausgangsniveau im Vergleich zur Prätestmessung. Ein signifikant höherer Wert, wie unmittelbar nach Beendigung des Treatments, wurde nicht mehr erreicht. Der Wunsch nach einer zusätzlichen Begleitung des eigenen schulpraktischen Tuns, sei es durch Unterrichtsbeobachtungen und/oder den organisierten fachlichen Austausch zwischen Kolleg*innen, deutet darauf hin, dass die etablierten Professionalisierungselemente während der *Treatment*phase zwar sehr intensiv genutzt (vgl. Kap. 5.1.2), nach deren Beendigung jedoch nicht konsequent weitergeführt worden ist. Dies bedeutet, dass für eine Modifikation des vorliegenden Professionalisierungskonzepts die nachhaltige Verankerung der Maßnahmen noch stärker in den Blick genommen werden sollte. Studien belegen ebenfalls, dass für die Veränderung von Einstellungen anhand tiefergehender Reflexionen längere, mitunter sich über mehrere Schuljahre erstreckende Formate vonnöten sein könnten (Leidig, 2019; Rzejak & Lipowsky, 2019).

Aufgrund der Ausrichtung des vorliegenden Professionalisierungskonzepts auf die individuelle Weiterentwicklung der professionellen Handlungskompetenz der Lehrkräfte auf der Ebene des Unterrichts, standen außerunterrichtliche Schulentwicklungsprozesse nicht im Fokus. Dennoch erscheint es sinnvoll, neben der Eruiierung einer nachhaltigen Etablierung der eingeführten Methoden und Materialien, weitere zielgerichtete und selbstreflexive Entwicklungsprozesse hin zu einer Professionalisierung der schulischen Prozesse systematisch einzuführen und zu konstituieren. Denkbar wären etwa Formate der kollegialen (Fall-)Beratung, die, eingebettet in ein Professionalisierungskonzept, sukzessive an der Schule etabliert werden könnten – konzeptionelle Überlegungen für alle drei Phasen der Lehramtsausbildung existieren bereits (Baumann et al., 2022; Liebner et al., 2022; Lindmeier, 2022). Auch das Modell der Kooperativen Beratung (Mutzeck, 2008) bietet ein niederschwelliges, aber zielgerichtetes, ressourcennutzendes sowie lösungs- und transferbezogenes Angebot der Hilfe zur Selbsthilfe. Die Methoden der kooperativ-systemischen Gesprächsführung identifizierten die Lehrkräfte bereits als nützlich für ihr zukünftiges Arbeiten (s. Tabelle 36). In diesem Zusammenhang sollten auch die Potentiale und Chancen multi-professio-

nellen Arbeitens mit Sonderpädagog*innen, Sozialarbeiter*innen, Psycholog*innen und weiteren pädagogischen Fachkräften an den Schulen konzeptionell für die gemeinsame Arbeit genutzt werden (Huber, 2015b; Urban, 2022).

Wendet man sich den weiteren Professionalisierungselementen zu, lässt sich resümieren, dass die Lehrkräfte mit der Durchführung und inhaltlich-methodischen Ausrichtung der Fortbildungsveranstaltungen vorwiegend zufrieden waren. Einige wünschten sich jedoch mehr Übungen und eine noch weiterreichende Wissensvertiefung. Für die Modifizierung der Professionalisierungsmaßnahme würde dies implizieren, sowohl die Vermittlung theoretischen Wissens im Handlungsfeld „Erweiterung des Lehrkraftwissens“ als auch den Praxisbezug im Handlungsfeld „Veränderung des unterrichtlichen Lehrkrafthandelns“ zu intensivieren sowie Input-, Erprobungs- und Reflexionsphasen noch enger zu verbinden, wie dies Studien zu wirksamen Fortbildungsmaßnahmen bereits beschreiben (Darling-Hammond et al., 2017; Leidig, 2019; Rzejak et al., 2020; Rzejak & Lipowsky, 2019). Die Verzahnung der Professionalisierungselemente könnte etwa durch gemeinsame Hospitationen und/oder die Methode der Videographie realisiert werden, die wiederum weiterführende Reflexionsmöglichkeiten gestatten (Faix et al., 2019).

Das Professionalisierungselement „Verhaltensverlaufsdiagnostik“ wurde bezüglich seiner Nützlichkeit zwar als befriedigend eingeschätzt, allerdings machen die nur geringe Wahrscheinlichkeit der zukünftigen Nutzung sowie die konkreten Verbesserungsvorschläge deutlich, dass vor allem der Aspekt der Praktikabilität skeptisch beurteilt worden ist. Die Wünsche der Lehrkräfte, wie die Ausdehnung des Beurteilungszeitraumes auf einen gesamten Tag, der Einsatz des DBR bei vielen Schüler*innen und die Reduktion der Skalenbreite stehen jedoch im Kontrast zu den Güte-Empfehlungen des Messinstruments (vgl. Kap. 2.1.3.4). Der Kritik der Lehrkräfte könnte deshalb nur bedingt nachgekommen werden. Dies wirft die Frage auf, inwieweit sich DBR grundsätzlich für Verhaltensmonitoring in der schulischen Praxis eignet. Denn für Lehrkräfte stellt es eine Herausforderung dar, wenn sie – wie in der vorliegenden Studie – ihre Schüler*innen auf RTI-Ebene I bzw. II fördern und unterrichtsbegleitend ein Verhaltensmonitoring bei mehreren Schüler*innen durchführen. Unterschiedliche Studien konnten zwar zeigen, dass DBR im schulischen Kontext von den praktizierenden Lehrkräften angewendet werden kann und Lehrkräfte die bevorzugte Anwender*innengruppe für diese Methode zu sein scheinen (Huber & Rietz, 2015), dabei kommt es aber maßgeblich auf die konkrete Umsetzung an (Mahlau & Erdmann, 2020). Danach muss das Messinstrument für den schulischen Gebrauch angepasst und dessen Anwendung so effizient und praktikabel wie möglich gestaltet werden. Denkbar und bereits in der Entwicklung sind hierbei etwa digitale Formate, bei denen die Beurteilungsdaten sowohl elektronisch erfasst als auch ausgewertet werden (Scheer 2021, S. 132ff.). Die größtmögliche Auto-

matisierung der Messabläufe würde zu einer Entlastung der Lehrkräfte führen, was vermutlich eine häufigere Anwendung der Methode zur Folge hätte. Für die Interpretation der Lehrkraftaussagen zum Verhaltensmonitoring muss allerdings auch berücksichtigt werden, dass DBR in der vorliegenden Studie maßgeblich für die Untersuchung der Effektivität der *84 Handlungsmöglichkeiten* zum Einsatz kam. Doch erst die Nutzung als Förderinstrument schöpft dessen Potential vollumfänglich aus. So können die Daten eine unmittelbare Rückmeldung zum Fördererfolg geben und Hinweise zur eventuell notwendigen Anpassung der Förderung liefern (vgl. Kap. 2.1.3.4) – das von den Lehrkräften geäußerte Anliegen einer zwischenzeitlichen Auswertung der Beurteilungen greift eben diesen Punkt auf. Darüber hinaus können die Ergebnisse bei der multiprofessionellen Arbeit von Lehrkräften, Sonderpädagog*innen, Schulpsycholog*innen und weiteren pädagogischen Fachkräften systematisch genutzt sowie bei der Beratung von Eltern und Schüler*innen eingesetzt werden (Casale et al., 2019). Da diese Fördermöglichkeiten in der vorliegenden Studie nicht realisiert wurden und der Einsatz von DBR als reines Diagnostikum erfolgte, müssen die Lehrkraftaussagen zu deren (zukünftiger) Anwendung relativiert werden. Im Hinblick auf die Etablierung mehrstufiger Fördersysteme an Schulen und der Notwendigkeit geeigneter Diagnostik- und Förderinstrumente (vgl. Kap. 2.1.3.4) sollten jedoch aussagekräftige Untersuchungen die konkrete Anwendung und Praktikabilität des DBR in der Schule noch gezielter berücksichtigen.

6.3 Schlussfolgerungen für die Konzeption zukünftiger Professionalisierungsmaßnahmen

Nach der Diskussion der Ergebnisse auf den Wirkungsebenen Lehrkraftreaktionen, Lehrkraftkognitionen, Lehrkrafthandeln und Schüler*innenoutcomes sollen im Folgenden allgemeinere Schlussfolgerungen gezogen werden. Zunächst kann resümiert werden: Nach der Durchführung der Professionalisierungsmaßnahme zeigen sich auf allen der vier untersuchten Wirkungsebenen (Lipowsky, 2010) bei den status- sowie verlaufsdiagnostischen Ergebnissen positive Merkmalsveränderungen bei Schüler*innen und Lehrkräften. Besonders bemerkenswert ist, dass auf der Ebene der Schüler*innenoutcomes in den intendierten Bereichen sozial-emotionaler Kompetenzen anhand unterschiedlicher abhängiger Variablen sowohl die nachhaltige Verringerung von Verhaltensauffälligkeiten als auch die Stärkung des prosozialen Verhaltens festgestellt werden können. Die Ergebnisse zur Einsatzhäufigkeit der *84 Handlungsmöglichkeiten* verweisen darauf, dass die Lehrkräfte im Durchschnitt an mehr als jedem zweiten Tag eine spezifischere unterrichtsintegrierte Förderung bei den Schüler*innen der Förderebene II (Kap. 2.1.3.4 und

Kap. 3.1) durchgeführt haben. Dies zeigt ein verändertes Lehrkrafthandeln, welches durchaus im Zusammenhang mit den Merkmalsveränderungen auf den anderen Wirkungsebenen gedeutet werden kann (Garet et al., 2001; Gärtner, 2007; Stipek et al., 2001; Vehmeyer et al., 2007). Neben den Schüler*innenoutcomes und dem Lehrkrafthandeln zeigen sich eine tendenziell günstige Entwicklung bei den Lehrkraftkognitionen und maßgeblich zufriedene Lehrkraftreaktionen. Die Ergebnisse zum Selbstwirksamkeitsempfinden sowie zur Zufriedenheit und Akzeptanz verweisen aber auch auf zukünftige Entwicklungsbedarfe. Die Vorüberlegungen zu den Stärken und Schwächen der Konzeption, die in Kapitel 3.3 ausgeführt worden sind, sollen an dieser Stelle wieder aufgegriffen werden.

Angesichts der Bewertung des Professionalisierungskonzepts und seiner -elemente durch die Lehrkräfte ist davon auszugehen, dass die didaktisch-methodische Gestaltung und die inhaltliche Schwerpunktsetzung des Professionalisierungskonzepts grundsätzlich geeignet sind, die Bedürfnisse der Zielgruppe hinsichtlich der Erweiterung der professionellen Handlungskompetenz im Umgang mit Schüler*innen mit Verhaltensproblemen zu erfüllen. Viele der vermuteten Stärken der Konzeption – die Vermittlung von im beruflichen Alltag umsetzbaren und evidenzbasierten Strategien (Hillenbrand, 2015; Leidig et al. 2016, S. 66), das langfristig angelegte Format (Kurniawati et al. 2014, S. 321; Lipowsky, 2014), prozessbezogene Begleitung und Reflexion bei der Umsetzung der Konzeption (Kurniawati et al. 2014, S. 321; Lipowsky, 2014), das fallbezogene Arbeiten (Hillenbrand 2015, S. 175; Leidig 2019, S. 55), das Verhaltensmonitoring (Leidig et al., 2016) und die universitäre Anbindung (Avramidis et al., 2000; Leidig et al. 2016, S. 65) – sind auf der Ebene der Lehrkraftreaktionen aufgegriffen und positiv hervorgehoben worden. Vor dem Hintergrund der Ergebnisse auf den anderen Wirkungsebenen können die genannten Kriterien gleichsam als wirksame Elemente von Lehrkraftprofessionalisierung für schulische Inklusion im Bereich ESE gedeutet werden. Die weiteren vermuteten Stärken der Konzeption wie die Einbeziehung der Schulleitung (Leidig et al., 2016; Urton et al., 2014), die Erfassung der Outcomes auf Schüler*innenebene (Leidig, 2016; Lipowsky, 2010; Timperley et al., 2007) und die Untersuchung der längerfristigen Wirkung der Fortbildungen und Interventionen (Amrhein & Badstieber, 2013) können sich ebenfalls positiv ausgewirkt haben. Wie elementar es ist, insbesondere im Grundschulbereich, die Schulleitungen bei Fortbildungsprogrammen mit einzubeziehen, zeigte sich anhand der guten Resonanz und dem relativ geringen Dropout, dem meist reibungslosen Studienablauf und der engmaschigen Kommunikations- und Absprachemöglichkeiten. An fünf von sieben teilnehmenden Schulen partizipierten die Schulleitungen an den Informations-, Fortbildungs- und Abschlussveranstaltungen (s. Abbildung 14). Vier Schulleitungen nahmen außerdem aktiv an der Studie teil, weil sie auch eine Klassenleitungsfunktion innehatten. Bemerkenswerterweise verließ

gleich zu Anfang diejenige Schule das Projekt, bei der der Schulleitungsposten unbesetzt war. Interessant in diesem Zusammenhang ist auch, dass von den bis zur Beendigung der Maßnahme teilnehmenden Schulen jeweils nahezu das gesamte Kollegium involviert war. Auch dieser Umstand scheint den Verlauf der Studie positiv beeinflusst zu haben.

Eine weitere Stärke des Professionalisierungskonzepts wurde in der Erfassung der Outcomes auf Schüler*innenebene vermutet, da diese in der Vergangenheit bei Fortbildungsevaluationen nicht entschieden genug berücksichtigt wurden (Leidig, 2016; Lipowsky, 2010; Timperley et al., 2007), obwohl sie wichtige Rückschlüsse über die Wirkungsintensität und die Reichweite von Professionalisierungsmaßnahmen ermöglichen können. In der vorliegenden Studie konnte gezeigt werden, dass insbesondere bei den Schüler*innen Merkmalsveränderungen stattgefunden haben. Die Untersuchung der längerfristigen Wirkung der Fortbildungsmaßnahmen wiederum erlaubte außerdem Rückschlüsse über die nachhaltige Wirkung der eingesetzten Elemente. Es wurde deutlich, dass die positiven Schüler*innenoutcomes auch noch vier Monate nach Beendigung der Professionalisierungsmaßnahme sichtbar waren.

Allerdings bestätigten die Lehrkraftaussagen zugleich die vermuteten Schwächen der Konzeption. Einige Pädagog*innen hätten sich umfangreichere Inputphasen, eine engere Begleitung der Erprobungsphasen und eine Erhöhung der Reflexionstätigkeiten gewünscht (vgl. Leidig et al. 2016, S. 58f.; Van Mieghem et al., 2020). Eine noch engere Verzahnung von Input-, Erprobungs- und Reflexionsphasen (Lipowsky 2014, S. 522), eine der weiteren vermuteten Schwächen, würde die drei genannten Aspekte gut zusammenführen können. Darüber hinaus wäre bei einer Modifikation des Professionalisierungskonzeptes der Einsatz der Verhaltensverlaufsdiagnostik DBR als Förderinstrument (Casale et al., 2019) anzustreben, damit die pädagogische Verzahnung von Diagnostik und Förderung tatsächlich realisiert werden kann. Außerdem wurde angeregt, den Umfang der gesamten Professionalisierungsmaßnahme auszuweiten und die dann intensiveren und umfangreicheren Professionalisierungselemente und -phasen über einen noch längeren Zeitraum zu strecken. Insbesondere, wenn eine systematische Beratungskultur an Schulen etabliert werden soll, muss eine längere professionelle Begleitung von mindestens einem Schuljahr eingeplant werden (Kurniawati et al., 2014).

Auf den Grundlagen des wissenschaftlichen Erkenntnisstands zu schulischer Inklusion im Bereich ESE und Professionalisierungsangeboten sowie aufbauend auf den Ergebnissen der vorliegenden Evaluationsstudie werden in Abbildung 20 die zentralen Schlüsselfaktoren für die Veränderung pädagogischen Handelns und in der Folge für den Schulerfolg der Schüler*innen auf sozial-emotionaler Ebene zusammengetragen und damit ein Ausblick auf zukünftige Professionalisierungsformate für die schulische Inklusion im Bereich ESE gegeben.

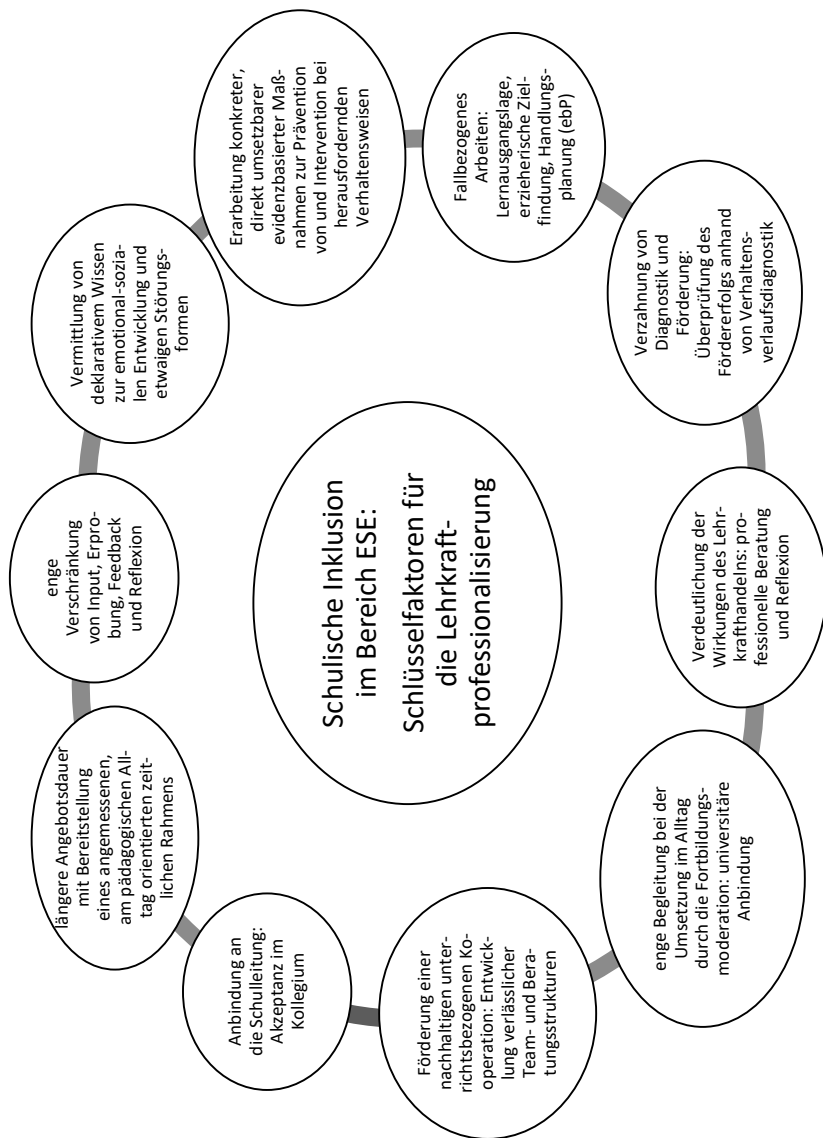


Abbildung 20: Zentrale Schlüsselfaktoren für die Professionalisierung von Lehrkräften für die schulische Inklusion im Bereich ESE (eigene Darstellung)

Anmerkungen. EbP = Evidenzbasierte Praxis, ESE = Emotional-soziale Entwicklung.

6.4 Ausblick für Forschung, Lehre und Praxis

Aus dem Auftrag, eine inklusive Bildung nach dem Verständnis der UNESCO umzusetzen, leitet sich auch der Anspruch ab, besonders wirksame, effektive Verfahren vor allem bei Kindern und Jugendlichen mit Risikokonstellationen einzusetzen (Blumenthal & Mahlau, 2015; Hillenbrand, 2015). Die inklusive Beschulung von Kindern mit Verhaltensauffälligkeiten und emotional-sozialen Entwicklungsstörungen stellt Forschung, Lehre und Praxis vor besondere Aufgaben. In diesem Zusammenhang stellt die evidenzbasierte Praxis (Blumenthal & Mahlau, 2015; Grosche, 2017) ein interdisziplinär und international anerkanntes Konzept dar, das – wie die vorliegende Studie gezeigt hat – für die Praxis anspruchsvoll, aber umsetzbar ist. Im Folgenden sollen aus den bereits diskutierten Ergebnissen und dargelegten Schlussfolgerungen weitere Forschungsdesiderate identifiziert sowie Anregungen für Maßnahmen und Methoden abgeleitet werden, die künftig in Forschung, Lehre und Praxis umgesetzt werden können.

Eine wissenschaftlich ausgerichtete (Sonder-)Pädagogik bedarf qualitativ hochwertiger Studien, die Nachweise über die Effektivität von in der Schulpraxis einsetzbaren Verfahren erbringen. Gleichzeitig hat sie den Auftrag, aus wissenschaftlichen Theorien und Erkenntnissen Förderkonzepte und -maßnahmen für die pädagogische Praxis zu entwickeln. Angelehnt an schulische Präventionsmodelle wie das RTI-Modell, die Förderung auf mehreren Ebenen vorsehen, sollten für kontinuierlich spezifischer werdende Förderungsszenarien wirksame Schul-, Unterrichts- und Förderkonzepte sowie konkrete Verfahren und Fördermaßnahmen entwickelt werden, die praktikabel an Schulen eingesetzt werden können.

Für die Förderebene I gilt es, Classroom-Management-Strategien als präventives Setting stärker auszubauen und effiziente wie wirkungsvolle Programme zur praktischen Umsetzung zu entwickeln. Hier bieten die *84 Handlungsmöglichkeiten* eine gute Anleitung für Pädagog*innen. In einem separaten Kapitel werden neun proaktive und zwei reaktive Prinzipien (Evertson & Emmer, 2009; Helmke, 2010) zunächst theoretisch dargestellt. Anhand einer Checkliste (S. 134f.) kann dann die eigene Unterrichtsführung mit konkreten Maßnahmen guten Classroom-Managements abgeglichen werden. Neben Strukturierungs- und Steuerungsaspekten beinhaltet ein gutes Classroom-Management auch das Training von sozial-emotionalen und prosozialen Fertigkeiten (Durlak et al., 2011). Diese sollten daher bei der unterrichtsintegrierten Förderung auf den Präventionsebenen I und II systematisch eingesetzt werden. Im Bereich der Einzelförderung auf Ebene III werden ggf. spezifischere und therapeutische Förder- und Diagnostikmethoden benötigt. Für den Grundschulbereich existieren bereits eine Reihe evidenzbasierter Verfahren, die nachweislich das Sozialverhalten fördern und auf der Ebene

der universellen und indizierten Förderung eingesetzt werden können, wie etwa das *KlasseKinderSpiel* (Hillenbrand & Pütz, 2008), das *Verhaltenstraining in der Grundschule* (Petermann et al., 2013) oder das *Training im Problemlösen* (Lösel et al., 2008). Bei vielen anderen entwickelten Förderprogrammen fehlt es allerdings an hochwertigen empirischen Studien und damit (noch) an Beweiskraft (vgl. Grüne Liste Prävention, 2022). Für die inklusive Beschulung im FSP ESE gilt es, weitere evidenzbasierte Verfahren zu entwickeln sowie bestehende Förderprogramme auf deren Wirksamkeit hin zu untersuchen.

Präventive Förderung bedeutet auch, betroffene Kinder so früh wie möglich zu identifizieren und die Ursache der Problematik festzustellen. Im FSP ESE treten sehr häufig komorbide Störungen auf, insbesondere die Wechselwirkungen zwischen ESE und Lernen sind hoch (Klauer & Lauth, 1997; Visser et al., 2019). Häufig tritt ESE dabei als Sekundärstörung auf. Mahlau und Blumenthal (2014) konnten in einer Studie zeigen, dass bei Kindern mit Sprachentwicklungsstörungen bereits sehr früh – zum Ende der 2. Klasse – Sekundärprobleme im Bereich Lernen und ESE diagnostiziert worden sind, ohne dass die Regelschulpädagog*innen die eigentliche Primärproblematik erkannten. Zu einer präventiven Förderung gehört demnach auch, eine geeignete Diagnostik durchzuführen, um die Primärproblematik gut erkennen zu können. Dies spricht für schulische Eingangsscreenings, die die für schulisches Lernen relevanten Vorläuferfähigkeiten testen und bei unterdurchschnittlichen Leistungen, etwa im Bereich der Sprachentwicklung, eine weiterführende Diagnostik einleiten. Auf diese Weise könnte der Entstehung von Sekundärproblematiken im Bereich ESE entgegengewirkt werden. Für die Wissenschaft ergibt sich der Auftrag, geeignete Screeningverfahren zu entwickeln. Bildungspolitik und Schuladministration müssen wiederum die Bedingungen schaffen, dass solche Verfahren im Unterricht auch eingesetzt werden (können).

In der vorliegenden Studie wurde ein vornehmlich kindzentrierter Ansatz verfolgt, der interpersonelle Risiko- und Schutzfaktoren des Kindes zu beeinflussen versucht (Beelmann & Raabe, 2007). Da Ursachen und Problemlagen im FSP ESE multifaktoriell bedingt sind (Blumenthal et al., 2022) und sich Problemlagen situations- und ortsunabhängig zeigen, werden zudem Förderkonzepte benötigt, die über die Förderung im Unterricht hinausgehen. Neben Maßnahmen auf individueller und Klassen-Ebene sollten pädagogische Präventionsangebote demnach auch auf Schul- und Systemebene etabliert werden. Hennemann et al. (2017, S. 88) stellen folgende Maßnahmen auf der Ebene der Schule heraus:

- Pädagogische Geschlossenheit, die einem im gesamten Kollegium anerkannten Leitbild folgt
- Transparenz und Einigkeit bezüglich der Abläufe und Vorgänge im Schul- bzw. Institutionssystem

- Regelmäßige Treffen und Lehrer*innenkonferenzen
- Wertschätzende Haltung gegenüber Kolleg*innen, Schüler*innen sowie den Erziehungsberechtigten
- Eine angenehme räumliche Gestaltung
- Klare Schulregeln und Konsequenzen
- Schulweite Umsetzung von Präventionsprogrammen
- Aktive Pausengestaltung
- Meditationsangebote
- Schulische Erziehungshilfe (bspw. Schulsozialarbeit)
- Beratungs- und Supervisionsangebote

Das in der vorliegenden Studie evaluierte Professionalisierungskonzept könnte anhand der Kriterien erweitert werden. In regelmäßig stattfindenden Förderplammentreffen im multiprofessionellen Team könnte ein für jedes Kind zwar individuelles, aber dennoch unter den Pädagog*innen einheitliches Vorgehen bei der Umsetzung präventiver Maßnahmen geplant und vorbereitet werden. Der Einsatz der *84 Handlungsmöglichkeiten* könnte im Team besprochen und ggf. nicht nur bei einer Lehrkraft in einer bestimmten Unterrichtssituation, sondern auch in anderen Unterrichts- und Schulsituationen angewendet werden. Begleitend sollten der Erfolg der präventiven Förderung systematisch anhand von DBR überprüft, die Ergebnisse ins Pädagog*innen-Team zurückgeführt und bei der Weiterführung des Förderplans berücksichtigt werden. Ein außerschulisches Netzwerk der Erziehungshilfe kann den innerschulischen Präventionsprozess außerdem unterstützen und anreichern. Zu nennen sind hier sozialpädagogische Angebote, psychologische Dienste, Beratungsstellen, Ärzte und Vereine usw. (Hennemann et al. 2017, S. 89). Auch diese Akteur*innen sind ggf. in eine multiprofessionelle Beratungskultur an Schulen mit einzubeziehen.

Neben kindzentrierten haben sich zwei weitere psychologisch-pädagogische Ansätze zur Prävention von emotional-sozialen Entwicklungsstörungen bewährt (Beelmann & Raabe, 2007). Während der umfeldzentrierte Ansatz maßgeblich versucht, das Erziehungsverhalten der Erziehungsberechtigten zu beeinflussen – Beispiele für gut evaluierte Trainingsprogramme sind etwa das *Positive Parenting Program* (Naumann et al., 2007; Sanders, 1999), das *Kompetenztraining für Eltern sozial auffälliger Kinder* (KES) (Lauth & Heubeck, 2006) und *EFFEKT-E* (Lösel et al., 2008) – nehmen multimodale Ansätze sowohl die Fertigkeiten des Kindes selbst als auch sein Umfeld in den Blick (Heinrichs et al., 2013). Letztere sind durch ihre systemische Ausrichtung oftmals schwerer zu realisieren. Ein zielführendes Beispiel für solch einen multimodalen Ansatz ist das Förderkonzept *Familienklassenzimmer* (Wuntke et al., 2022), deutschlandweit auch verbreitet unter den Namen *Familienklasse* oder *Familie in Schule* (Dawson et al., 2020). Das multimodale Förderkonzept für die Regelschule sieht vor, dass Schüler*innen an einem Tag in der Woche zusammen mit ihren Eltern und anderen Familien

von systemisch ausgebildeten Fachkräften unterrichtet werden. Möglichen Schulversionen und Verhaltensauffälligkeiten soll durch das Verbessern der Beziehung zwischen Eltern, Kind und Lehrkraft als System entgegengewirkt werden, um dem/der Schüler*in einen gelingenden Umgang mit Schul- und Alltagssituationen zu ermöglichen. Daneben sorgen verhaltenstheoretische Elemente wie Verhaltensverträge, Punkteabzugsprogramme (Response-Cost-Verfahren) und Token-Programme – wie sie auch die *84 Handlungsmöglichkeiten* beschreiben (Hartke et al. 2018, S. 50ff.) – für eine kontinuierliche Vernetzung zwischen Unterricht, Familienklassenzimmer und Elternhaus (Behme-Matthiessen et al., 2012). Neben der Etablierung einer multimethodalen Diagnostik an inklusiven Schulen (Voß et al., 2020) ist demnach auch die Verbindung unterschiedlicher Förderansätze im Sinne einer multimodalen Förderung im Bereich ESE sehr zu begrüßen. Die in der vorliegenden Studie evaluierten *84 Handlungsmöglichkeiten* könnten gewinnbringend in ein multimodales Förderszenario integriert werden.

Über den Bedarf an nachweislich wirksamen Unterrichts- und Förderverfahren hinaus sind im Sinne des präventiven Beschulungsansatzes nach dem RTI-Modell datengeleitete Förderentscheidungen zu treffen (Blumenthal et al., 2014; Cheney et al., 2008; Gresham et al., 2005). Für die dafür notwendige prozessbegleitende Diagnostik wurde mit dem *Direct Behavior Rating* (Christ et al., 2009) ein Diagnostikum entwickelt, das Verhalten reliabel und valide misst (Casale et al., 2021; Chafouleas, 2011; Matta et al., 2020; Riley-Tillman et al., 2011). Nun gilt es, das Verfahren in der unterrichtlichen und schulischen Praxis zu verankern. Wichtige Schritte auf diesem Weg, wie etwa eine theoretisch fundierte und gleichzeitig praxisorientierte Handlungsanleitung zur Planung, Umsetzung und Auswertung von DBR (Casale et al., 2019) sowie digitale Formate, die die effiziente Eingabe und Auswertung von Schüler*innendaten ermöglichen (Scheer 2021, S. 132ff.), wurden und werden bereits realisiert. In der vorliegenden Studie stellte allerdings gerade der praktische Einsatz des Messinstruments die Lehrkräfte vor Herausforderungen. Es bedarf daher Studien, die gezielt die Praktikabilität und Anwendbarkeit des DBR in den Blick nehmen und, ausgehend von den Studienergebnissen, Anregungen für die schulische Etablierung entwickeln.

Jenseits der Entwicklung effektiver Förderprogramme und -maßnahmen für die pädagogische Praxis sollten weiterführende Erkenntnisse zur Verbesserung der schulischen Förderung von Kindern mit Verhaltensauffälligkeiten und emotional-sozialen Entwicklungsstörungen gewonnen werden. In diesem Zusammenhang sind Studien gefragt, die sich der Theoriefundierung und Evidenzbasierung von entwickelten Förderprogrammen widmen. Hilfreich ist dabei, wenn Präventionsprogramme ihre Annahmen über Wirkungszusammenhänge eindeutig formulieren, damit sie explizit auf ihre Wirksamkeit hin geprüft werden können (Veermann & van Yperen, 2007). Neben der Untersuchung externer Evidenzen sollten insbesondere auch die interne sowie

soziale Evidenz (Blumenthal & Mahlau, 2015; Grosche, 2017; Hillenbrand, 2015) betrachtet werden. Aus diesem Grund müssen entsprechende Interventionsstudien unterschiedliche Designs verwenden. Neben experimentellen Einzelfallstudien auf der Grundlage mehrfach gestufter Multiple Baseline Designs und randomisierten (Warte)-Kontroll-Gruppenstudien mit einer möglichst großen Stichprobengröße, sollten umfassende Evaluationsstudien konzipiert werden, die es ermöglichen, Effekte auf mehreren Wirkungsebenen (Guskey, 2014; King, 2014; Lipowsky, 2010, 2014) zu erfassen. Quantitative Designs sollten durch qualitative Methoden ergänzt werden – die Unterrepräsentanz von Studien im Mixed-Method-Design in der Professionalisierungsforschung im Bereich ESE deutet bereits auf dieses Forschungsdesiderat hin (Leidig et al., 2016; vgl. Tabelle 9). Mit den Ergebnissen der vorliegenden Studie spricht vieles dafür, den gewählten Ansatz modifiziert weiter zu verfolgen. Die Untersuchung der Lehrkraft- und Schüler*innenmerkmale auf den einzelnen Wirkungsebenen sollte dabei anhand der in der Diskussion aufgeworfenen Kritikpunkte verändert werden. Aus den gewonnenen Erkenntnissen könnte dann ein verbessertes Treatment entwickelt werden. Sinnvoll wäre etwa, nur eine Auswahl der aufgeführten Interventionsmöglichkeiten (unterrichtsintegrierte, kindzentrierte Handlungsmöglichkeiten/kooperative Beratung/Verhaltensverlaufsdiagnostik) umzusetzen bzw. bei dem Einsatz mehrerer Interventionen die einzelnen besser zu kontrollieren, um feststellen zu können, welche Maßnahme tatsächlich wie wirkt. Um ein solches Treatment zu evaluieren, sollte aufgrund des Standes der Theorieentwicklung eine hypothesenprüfende Interventionsstudie mit einem Zweigruppenversuchsplan (Interventions- und Wartekontrollgruppe) über mindestens ein Schuljahr mit Prä-Post-Follow-Up-Design durchgeführt werden. Umsetzung und Überprüfung der Wirksamkeit des Treatments bzw. der Intervention sollten anhand eines hochwertigen Begleitstudiendesigns und multipler Erhebungsmöglichkeiten (Screenings, Fragebögen, Verlaufsmessungen, Interviews, Videoanalysen/Hospitationen) abgesichert sein.

Mit der Ratifizierung der UN-BRK im Jahre 2009 hat sich das Professionalisierungsprofil von Regelschullehrkräften maßgeblich verändert (Melzer et al., 2015). Die notwendigen Reformprozesse in der Aus- und Weiterbildung von Lehrkräften finden jedoch noch nicht flächendeckend und konsequent genug statt (Döbert & Weishaupt, 2013). Zum einen gilt es, bereits formulierte Professionalisierungsprofile (Gottfried et al., 2021; König et al., 2019; Melzer et al., 2015) weiter zu schärfen. Zum anderen müssen Konzepte entwickelt werden, wie diese sowohl in der ersten als auch in der zweiten und dritten Phase der Lehrer*innenbildung in die Praxis überführt werden können. Für die Forschung ergibt sich konkret der Auftrag, wissenschaftsbasierte Professionalisierungsformate für alle drei Phasen der Lehrer*innenbildung zu konzipieren und durch Begleitstudien auf ihre Wirksamkeit und Effektivität hin zu überprüfen. Einige Institute und Forschungsprojekte haben sich hier

bereits auf den Weg gemacht. Im deutschsprachigen Raum sind etwa die *Inklusionsorientierten Lehr-Lern-Bausteine* der Humboldt-Universität zu Berlin (Brodeser et al., 2020), die *Inklusionsdidaktischen Lehrbausteine* der Universität München (Riesch et al., 2022) sowie die *Lernpfade in die inklusive Pädagogik* der Universität Bremen (Müller, 2018) zu nennen. Bei den oftmals interdisziplinär entwickelten Professionalisierungskonzepten und -materialien handelt es sich einerseits um didaktisch aufbereitete Lehr-Elemente, die in der Hochschullehre eingesetzt werden können. Andererseits bieten sie Lernplattformen ohne tutorielle Begleitung, die darauf abzielen, sowohl Lehramtsstudierende als auch Lehrkräfte aller Fächer und Schularten auf ein inklusives Schulsystem vorzubereiten. International lassen sich ebenfalls derartige Beispiele für die Aus- und Weiterbildung im Zuge einer inklusiver werdenden Lehrer*innenbildung finden. Das vom *Office of Special Education Programs* (OSEP) des US-Bildungsministeriums finanzierte *PROGRESS Center* ([Promoting Rigorous Outcomes and Growth by Redesigning Educational Services for Students With Disabilities Center]; American Institutes for Research, 2022) etwa bietet umfangreiche, größtenteils kostenfreie Informationen, Materialien, Online-Fortbildungen und -Trainings zu Fachwissen in den Bereichen evidenzbasierter Praktiken, Sonder- und Inklusionspädagogik sowie multiprofessionelles Arbeiten und Netzwerken an. Ziel ist die Entwicklung und Umsetzung hochwertiger Bildungsprogramme für Schüler*innen mit Beeinträchtigungen. Der Vorteil dieser online gestützten Formate ist deren große Reichweite und ein niederschwelliger Zugang. Um professionelle Handlungskompetenzen im Umgang mit einer heterogener werdenden Schülerschaft zu erwerben und auszubauen, reichen diese vorrangig Input-basierten, instruktiven und wissensvermittelnden Professionalisierungsangebote und -methoden allerdings nicht aus, selbst wenn sie eine gewinnbringende Ergänzung zu anderen Lehr-Lern-Formaten bieten mögen. Zur nachhaltigen (Weiter-)Entwicklung professioneller Handlungskompetenzen für den inklusiven Unterricht bzw. auch die Veränderung unterrichtspraktischen Handelns gilt es neben der Wissenserweiterung auch die Lehrkraftkognitionen und die affektiv-motivationale Entwicklung der Lehrkräfte zu fördern (Lipowsky, 2010; Van Mieghem et al., 2020). Neben der Bereitstellung digitaler Möglichkeiten lautet der Auftrag an Wissenschaft, Praxis und Politik vor allem, die praktische Aus- und Fortbildungslandschaft (zukünftiger) Lehrkräfte konzeptionell weiterzuentwickeln. Für die dritte Phase der Lehrer*innenbildung bietet die Professionalisierungsforschung bereits konkrete Hinweise für die Planung und Durchführung von Fortbildungsmaßnahmen (Ahmed et al., 2014; Darling-Hammond et al., 2017; Kurniawati et al., 2014; MacFarlane & Woolfson, 2013; Rzejak et al. 2020, S. 21ff.; Rzejak & Lipowsky 2019, S. 20ff; Van Mieghem et al., 2020). Für den Bereich ESE sind insbesondere die Vermittlung konkreter, direkt umsetzbarer evidenzbasierter Maßnahmen und das fallgestützte Arbeiten essen-

tiell. Daneben spielen Beratung, Feedback und Coaching sowie Kooperation und Teamarbeit bei der Umsetzung der erlernten Handlungsstrategien eine tragende Rolle, wie in Abbildung 19 dargestellt. Diese Kompetenzen können nur durch intensive und interaktive Formen der Aus- und Weiterbildung gewährleistet werden.

Auch für die Lehre und Praxis können wichtige Schlussfolgerungen gezogen werden. So könnten die Erkenntnisse der Studie Gegenstand der universitären Lehre und verbindlicher Qualifizierungsmaßnahmen vor dem Hintergrund einer inklusiven Schule sein. Für die Ausbildung von Regelschullehrkräften muss zunächst gefordert werden, dass obligatorische sonderpädagogische Lehreinheiten und Module flächendeckend in das Studium integriert werden. Im Rahmen inklusionsorientierter Lehre könnten Elemente des vorliegenden Professionalisierungskonzepts realisiert werden, wie etwa die *84 Handlungsmöglichkeiten* als eine Möglichkeit unterrichtsintegrierter, kindzentrierter oder auch systemisch-multimodaler Förderung sowie – neben statusdiagnostischen Verfahren – das *Direct Behavior Rating* (Christ et al., 2009) als Instrument zur Verhaltensverlaufsmessung und zum Verhaltensmonitoring. Angesichts des Überforderungsempfindens und tendenziell negativer Einstellungen vieler im Schuldienst befindlicher Regelschullehrkräfte und Regelschullehramtsstudierender schulischer Inklusion gegenüber (Avramidis et al., 2000; Ruberg & Porsch, 2017; Urbutt, 2001; Van Mieghem et al., 2020) ist neben der Wissensvermittlung zu evidenzbasierten Verfahren auch der Erwerb von Handlungswissen und professionellen Handlungskompetenzen im Sinne der Erweiterung um situative Skills und inklusive Kompetenzaspekte (Baumert & Kunter, 2006; Blömeke et al., 2018; Gottfried et al., 2021; König et al., 2019) essentiell. In ihrem systematischen Review deutschsprachiger Forschungsarbeiten über Einstellungen zur schulischen Inklusion von (angehenden) Lehrkräften schlussfolgern Ruberg und Porsch (2017), dass zwar eine positive Entwicklungstendenz der Einstellungen zur schulischen Inklusion zu verzeichnen ist, diese aber maßgeblich mit dem „inklusionsorientierten Kontext“ korrespondiert: Mit einer „fortschreitenden Implementierung inklusiver Strukturen und der damit erweiterten Erfahrungsräume“ (S. 410) könnten sich die Einstellungen zukünftig weiter verbessern. Auch diese Erkenntnis unterstreicht die Notwendigkeit der Neuformierung des Regelschullehramtes.

Es sollten außerdem zunehmend Ausbildungskonzepte realisiert werden, die eine stärkere Verzahnung von Theorie und Praxis anstreben. Der Einsatz von Lehrer*innentrainings etwa, der von einigen schon lange gefordert wird (Havers & Toepell, 2002; Jürgens & Krause, 2010; Toepell, 2010), könnten einen Beitrag zur Überwindung der beklagten Kluft zwischen Theorie und Praxis in der Ausbildung von Lehrkräften leisten. Die Inklusionsforschung stellt die Relevanz einer Kooperations- und Beratungskultur an inklusiven Schulen heraus, die für Eltern- und Schülerarbeit, aber ebenso im kollegialen

Austausch und zur eigenen Professionalisierung hilfreich sein können (Melzer et al., 2015). Um die Beratungskompetenz der Lehrkräfte zu schulen, sollten Beratungsformate wie die Kooperative Beratung (Mutzeck, 2008) und/oder die Kollegiale Supervision (Schlee & Mutzeck, 1996) in die Ausbildung integriert werden. Die Universität Rostock beispielsweise hat deutschlandweit ein einzigartiges Modell entworfen, bei dem grundständige Sonderpädagogik-Lehramtsstudierende (Studiengangsspezifische Prüfungs- und Studienordnung für den Studiengang Lehramt für Sonderpädagogik, 2012) sowie Masterstudierende des Weiterbildungsstudiengangs "Sonder- und Inklusionspädagogik" (Studiengangsspezifische Prüfungs- und Studienordnung für den Masterstudiengang Sonder- und Inklusionspädagogik der Universität Rostock, 2016) umfangreiche Beratungskompetenzen erwerben.

Neben der praktischen Anwendung in Lehr-Lern-Simulationen, sollte der frühe Praxisbezug im Studium systematisch etabliert werden. Einige Universitäten setzen bereits entsprechende Formate um, wie etwa einen Praxistag pro Woche von Anfang des Studiums an (vgl. Prüfungs- und Studienordnung für den Studiengang Grundschullehramt an der Universität Greifswald 2020). In Anbetracht einer sich fortlaufend und dynamisch entwickelnden Schullandschaft sind auch Konzepte zur gemeinsamen Professionalisierung aller drei Phasen der Lehramtsausbildung denkbar. Die Neuformierung der Lehramtsausbildung an den Universitäten sollte ferner für die zweite und dritte Phase der Lehrer*innenbildung fruchtbar gemacht werden.

Schließlich ist erstrebenswert, dass die Akteur*innen im Kontext Schule verpflichtende und regelmäßig stattfindende hochwertige Weiter- und Fortbildungsangebote zum Erwerb von inklusiv ausgerichteten Handlungskompetenzen wahrnehmen können. Zum einen muss dafür die Fortbildungslandschaft qualitativ und strukturell weiterentwickelt werden. Zum anderen gilt es, die Weiterqualifizierung durch Bildungsministerien, Schulämter und Lehrkräfte-Fortbildungsinstitute organisatorisch zu ermöglichen. Neben der Fortbildung von Regelschullehrkräften, die in der vorliegenden Arbeit besonders im Fokus stehen, kommt der Professionalisierung von Schulleitungen als sogenannte „change agents“ (Stamann, 2015) eine ebenso wichtige, wenn nicht sogar ausschlaggebende Rolle zu. Ihnen obliegt es, Rahmenbedingungen zu schaffen, die ein Hervorbringen und die Verbreitung von Innovationen ermöglichen bzw. unterstützen (Huber 2010, S. 214). Nachweislich können Schulleitungen zudem die inklusive Einstellung von Lehrkräften beeinflussen (Ahmed et al., 2014; MacFarlane & Woolfson, 2013; Urton et al., 2014). Die Konzipierung eines inklusiven und präventiven Schulkonzepts ist vor dem Hintergrund knapper Ressourcen jedoch eine herausfordernde Aufgabe. Zahlreiche Stellungnahmen von Lehrkraftverbänden und -gewerkschaften weisen darauf hin, dass dafür notwendige Ressourcen und Rahmenbedingungen nicht konsequent und systematisch genug zur Verfügung gestellt bzw. etabliert werden (Lehrergewerkschaften Hamburg, 2022; Verband Bildung

und Erziehung, 2022; Weidemann, 2019). Auf der personellen Ebene etwa sind Möglichkeiten von Team-Teaching-Formaten durch Doppelbesetzung, von Differenzierung und Individualisierung des Unterrichts auf Förderebene I, II und III, der Durchführung begleitender schulischer Diagnostik, der additiven oder parallelen Förderung, der Bereitstellung von Sonderpädagog*innen, Schulpsycholog*innen, Schulsozialarbeiter*innen sowie der professionellen Beratung anzustreben. Trotz der Neuformierung von Professionalisierungs- und Ausbildungsprofilen sollte dabei die hohe Qualität der sonderpädagogischen Ausbildung gesichert sein. Schon jetzt müssten eigentlich mehr Sonderpädagog*innen ausgebildet werden, um den Personalbedarf decken zu können (Ahrbeck 2017, S. 10). Auf der sächlich-räumlichen Ebene müssten Schulen für die gelingende inklusive Förderung gut ausgestattete und geräumige Unterrichts- und Förderräume sowie zugängliche Förder- und Diagnostikmaterialien bereitstellen können. Weiter gilt es die berufsbegleitende Professionalisierung der Schulakteur*innen finanziell zu unterstützen. Eine Möglichkeit für die prozessbegleitende und gleichzeitig relativ ressourcenschonende Verbreitung von Fortbildungsinhalten bieten Train-the-Trainer-Modelle oder Multiplikatoren*innenschulungen (Lipowsky 2014, S. 531). Idealerweise realisiert unter dem Einbezug von Universitäten, werden pädagogische Akteur*innen durch professionelle Trainer*innen bzw. Fortbildner*innen dazu befähigt, selbst auch Trainer*in bzw. Fortbildner*in zu werden. Diese können dann wiederum weitere Fachkräfte qualifizieren. Wird die Qualifizierung prozessbegleitend angeboten, bietet sich die Möglichkeit, die Maßnahme dialogisch weiterzuentwickeln sowie Coaching- und Beratungselemente zu realisieren. Besonderes Augenmerk sollte bei der Durchführung von Multiplikator*innenformaten auf der langfristigen Qualitätssicherung der Fortbildungen liegen. Dementsprechend sollte die Maßnahme im Prozess evaluiert, bei Bedarf frühzeitig modifiziert und auf diesem Wege professionell weiterentwickelt werden (Behr et al., 2020).

7 Fazit

Infolge der Unterzeichnung der Behindertenrechtskonvention durch Deutschland im Jahr 2009 wurde die Verpflichtung zur Umsetzung eines inklusiven Schulsystems Bundesrecht. Die ländergesetzlichen Verankerungen und die konkrete Umsetzung in den Bildungsinstitutionen zeigt, dass die Bundesländer vor dem Hintergrund der Kulturhoheit den dafür notwendigen Transformationsprozess auf unterschiedliche Weise vollziehen (Blanck, 2015). Trotz länderspezifischer Ausformungen können zwei zentrale Aufgaben dieses Reformprozesses festgehalten werden: die größtmögliche gemeinsame Beschulung von Kindern mit und ohne sonderpädagogischen Förderbedarf im Regelschulsystem (Klemm, 2018) sowie die „Erfüllung der Bedürfnisse aller Lernenden“ (Hillenbrand 2013, S. 366). Der schulischen Förderung von Kindern mit Verhaltensauffälligkeiten und emotional-sozialen Entwicklungsstörungen sollte im Zuge der Inklusionsbestrebungen aus mehreren Gründen besondere Aufmerksamkeit gewidmet werden. Zum einen weisen Schüler*innen mit dem FSP ESE eine der höchsten Quoten gemeinsamer Beschulung auf (KMK, 2022) und wurden schon vor der Ratifizierung der UN-BRK mehrheitlich an Regelschulen unterrichtet (Stein & Müller 2015, S. 39f.). Zum anderen stellen diese wie kaum eine andere Schüler*innengruppe die Unterrichtenden vor sehr herausfordernde pädagogische Aufgaben. Lebensgeschichtlich weisen Kinder mit Verhaltensauffälligkeiten und emotional-sozialen Entwicklungsstörungen oftmals erhebliche Belastungen auf, die zu „Bindungsstörungen, unzureichend entwickelten psychischen Strukturen, gravierenden inneren Konflikten und nicht selten Traumatisierungen geführt haben“ können (Ahrbeck 2017, S. 3). Mit der Diagnose SF ESE gehen häufig persistierende und kumulierende Risikokonstellationen einher (Beelmann & Raabe, 2007; Ghandour et al., 2019; Salmanian et al., 2019; Visser et al., 2019), die auf Seiten der Kinder mitunter zu einem langen Leidensweg mit ungünstigen Entwicklungsprognosen führen können (Ihle & Esser, 2008). Angesichts der komplexen Problemlagen verwundert es nicht, dass Lehrkräfte diese Schüler*innen als besonders herausfordernd erleben (de Boer et al., 2011). Sie bringen den Kindern zusätzlich negative Einstellungen entgegen (Avramidis & Norwich, 2002; Lübke et al., 2016; Sermier Dessemonet et al., 2011). Es wird deutlich: Um die erforderlichen (sonder-)pädagogischen Aufgaben bewältigen zu können, bedarf es einer qualifizierten und professionellen Unterstützung. Zentral ist demnach die Kompetenzentwicklung für die pädagogische Arbeit in Bezug auf Schüler*innen mit Verhaltensproblemen.

Das Ziel der vorliegenden Dissertation war die Entwicklung und Evaluation eines prozessbegleitenden Professionalisierungskonzepts zur Förderung der emotional-sozialen Entwicklung von Schüler*innen im inklusiven Unterricht. Das entwickelte Konzept, welches auf dem Forschungsstand zur schulischen Förderung der ESE von Kindern sowie Merkmalen wirksamer Lehrkraftprofessionalisierung basiert, zielte damit auf die Erweiterung spezifischer Lehrkraftkompetenzen für herausfordernde Lehr-Lernsituationen ab. Die Ergebnisse der Evaluationsstudie sollten zum einen Hinweise dazu liefern, inwieweit die entwickelte Konzeption einen generellen Orientierungsrahmen bei der Professionalisierung von Lehrkräften für den inklusiven Unterricht im FSP ESE bieten kann. Zum anderen sollte untersucht werden, ob sich die spezifischen Professionalisierungselemente *84 Handlungsmöglichkeiten* (Hartke et al., 2018) und *psychosozial-horizontale Beratung* (Mutzeck, 2008) als potentiell wirkungsvoll und praktikabel für die schulische Förderung im Bereich ESE erweisen.

Die vorliegenden Ergebnisse zeigen das Potential der Fortbildungskonzeption im Kontext der Lehrkraftprofessionalisierung im Bereich ESE auf. Sie liefern außerdem konkrete Hinweise dafür, dass die Erweiterung des Lehrkraftwissens anhand von Fortbildungen, die Veränderung unterrichtspraktischen Handelns anhand einer reflektierten Einbindung pädagogischer (evidenzbasierter) Strategien sowie die psychosozial-horizontale Beratung als hilfreiche Elemente zur Entwicklung schulischer Inklusion eingeschätzt werden können. Die Konzeption kann damit sowohl auf inhaltlicher als auch auf didaktisch-methodischer wie struktureller Ebene einen Orientierungsrahmen für die Entwicklung von prozessbegleitenden Professionalisierungsmaßnahmen im Kontext Inklusion im Bereich ESE bieten. Die eingesetzten Professionalisierungselemente *84 Handlungsmöglichkeiten* und *psychosoziale Beratung* stellen dabei vielversprechende, von den Lehrkräften als nützlich erachtete Instrumente dar, um die (Weiter-)Entwicklung ihrer professionellen Handlungskompetenzen strukturiert und kleinschrittig zu unterstützen. Das Buch *84 Handlungsmöglichkeiten* (Hartke et al., 2018) bietet ein theoretisch fundiertes Gesamtkonzept zur Diagnostik, Förderplanung und Evaluation, welches sich sehr gut in ein prozessbegleitendes Professionalisierungskonzept einbinden lässt. Die eingesetzten Handlungsmöglichkeiten haben sich bei den Schüler*innen in der vorliegenden Studie als wirksam erwiesen. Die kooperativ-systemische Beratung wurde von den Lehrkräften zwar als nützlich eingeschätzt, zeigte auf der Ebene der Lehrkraftkognitionen aber keine positiven Auswirkungen. Dies könnte durch eine Ausweitung der Beratungselemente und deren systematische Implementation in den schulischen Alltag sehr wahrscheinlich gesteigert werden.

Als eine der aktuell größten Herausforderungen an Schulen in Deutschland geben Lehrkräfte im Schuldienst in repräsentativen Umfragen das Verhalten der Schüler*innen und die schulische Inklusion an, ein Umstand, der

sich mit der Corona-Pandemie verstärkt hat (Robert Bosch Stiftung, 2019; 2022). Es ist demnach essentiell, im Schuldienst befindliche Lehrkräfte, die in ihrem Studium keine inklusionsorientierten Anteile absolviert haben, für ihre stetig heterogener werdende Schülerschaft, insbesondere in der Kompetenzdomäne pädagogisch-psychologischen Wissens (Baumert & Kunter, 2006), weiter zu professionalisieren. Das Paradigma, Verfahren mit hoher Evidenz einzusetzen, ist in der Schulforschung, auch im Bereich der allgemeinen und der Sonderpädagogik, angekommen. Dieses sollte systematisch auch auf die Professionalisierungsforschung ausgeweitet werden. Beides gilt es in die Schul- und Professionalisierungspraxis zu überführen. Die vorliegende Studie konnte dazu beitragen, wichtige Erkenntnisse zur Professionalisierung von Lehrkräften zu generieren und damit dem Verständnis einer inklusiven Schule für Schüler*innen im Bereich ESE zu entsprechen. Sie bietet des Weiteren einen konzeptionellen Rahmen für die Gestaltung zukünftiger Professionalisierungsmaßnahmen auf dem Weg zu einem inklusiven Bildungssystem.

Literaturverzeichnis

- Adey, P. (2004). *The professional development of teachers: Practice and theory*. Dordrecht, Netherlands: Kluwer Academic Publishers.
- Ahmed, M., Sharma, U. & Deppeler, J. (2014). Variables affecting teachers' intentions to include students with disabilities in regular primary schools in Bangladesh. *Disability & Society*, 29(2), 317–331.
- Ahrbeck, B. (2011). *Der Umgang mit Behinderung*. Stuttgart: Kohlhammer.
- Ahrbeck, B. (2017). *Welchen Förderbedarf haben Kinder mit emotional-sozialen Entwicklungsstörungen? EXPERTISE*. Verband Bildung und Erziehung.
- Ajzen, I. (1985). From Intentions to Actions: A Theory of Planned Behavior. In J. Kuhl & J. Beckmann (Hrsg.), *Action Control. From Cognition to Behavior* (S. 11-39). Heidelberg: Springer.
- Ajzen, I. (1987). Attitudes, Traits, and Actions: Dispositional Prediction of Behavior in Personality and Social Psychology. In L. Berkowitz (Hrsg.), *Advances in Experimental Social Psychology. Advances in Experimental Social Psychology* (S. 1-63). New York: Elsevier.
- Alresheed, F., Hott, B. L. & Bano, C. (2013). Single Subject Research: A Synthesis of Analytic Methods. *The Journal of Special Education Apprenticeship*, 2, 1.
- American Institutes for Research (2022). *PROGRESS Centre*. Zugriff am 30.11.2022 auf <https://promotingprogress.org/contact-us>.
- Amrhein, B. & Badstieber, B. (2013). *Lehrerfortbildungen zu Inklusion – eine Trendanalyse*. Gütersloh: Bertelsmann-Stiftung.
- Andersson, C. & Palm, T. (2017). The impact of formative assessment on student achievement. A study of the effects of changes to classroom practice after a comprehensive professional development programme. *Learning and Instruction*, 49, 92-102.
- Andrade, B. F., Browne, D. T. & Tannock, R. (2014). Prosocial skills may be necessary for better peer functioning in children with symptoms of disruptive behavior disorders. *PeerJ*, 2. <https://doi.org/10.7717/peerj.487>.
- Avramidis, E., Bayliss, P. & Burden, R. (2000). A Survey into Mainstream Teachers' Attitudes Towards the Inclusion of Children with Special Educational Needs in the Ordinary School in one Local Education Authority. *Educational Psychology*, 20(2), 191-211.
- Avramidis, E. & Norwich, B. (2002). Teachers' attitudes towards integration/inclusion: A review of the literature. *Journal of Special Needs Education*, 17(2), 129-147.
- Barker, C., Pistrang, N. & Elliott, R. (2015). *Research Methods in Clinical Psychology: An Introduction for Students and Practitioners*. Hoboken: Wiley-Blackwell.
- Barlow, D. H., Nock, M. K. & Hersen, M. (2009). *Single case experimental designs: Strategies for studying behavior change* (3rd ed.). New York: Pergamon Press.

- Baumann, R., Grüning, M. & Martschinke, S. (2022). Kollegiale Fallberatung in der zweiten Ausbildungsphase zum Umgang mit Herausforderungen in inklusiven Settings. *PFLB – PraxisForschungLehrer*innenBildung*, 4(3), 279-295. <https://doi.org/10.11576/pflb-5225>.
- Baumert, J. & Kunter, M. (2006). Stichwort: Professionelle Kompetenz von Lehrkräften. *Zeitschrift für Erziehungswissenschaft*, 9, 469-520.
- Baumert, J. & Kunter, M. (2011). Das Kompetenzmodell von COACTIV. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss & M. Neubrand (Hrsg.), *Professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV* (S. 29-53). Münster: Waxmann.
- Beelmann, A. (2006). Wirksamkeit von Präventionsmaßnahmen bei Kindern und Jugendlichen: Ergebnisse und Implikationen der integrativen Erfolgsforschung. *Zeitschrift für klinische Psychologie und Psychotherapie*, 35, 151-162.
- Beelmann, A. (2008). Prävention im Schulalter. In B. Gasteiger-Klicpera & J. Klicpera (Hrsg.), *Sonderpädagogik der sozialen und emotionalen Entwicklung* (S. 442-464). Göttingen: Hogrefe.
- Beelmann, A. & Raabe, T. (2007). *Dissoziales Verhalten von Kindern und Jugendlichen. Erscheinungsformen, Entwicklung, Prävention und Intervention*. Göttingen: Hogrefe.
- Behme-Matthiessen, U., Pletsch, T., Bock, K. & Nykamp, A. (2012). *Handbuch Familienklassen. Multifamiliencoaching im Unterricht*. Aachen: Shaker Verlag.
- Behr, J., Leidig, T. & Hennemann, T. (2020). Train-the-Trainer. Entwicklung und Umsetzung einer prozessbegleitenden Fortbildung zur Erweiterung von Handlungskompetenzen und Förderung des Selbstwirksamkeitserlebens im inklusiven Kontext. *Journal für Psychologie*, 27(2), 6-28. <https://doi.org/10.30820/0942-2285-2019-2-6>.
- Berk, L. (2011). *Entwicklungspsychologie*. München: Pearson.
- Berkemeyer, N., Järvinen, H. & Otto, J. (2012). Schulische Netzwerke als Professionalisierungsstrategie. Erfahrungen und Forschungsbefunde aus dem Projekt „Schulen im Team“. In M. Kobarg, C. Fischer, I. M. Dalehefte, F. Trepke & M. Menk (Hrsg.), *Lehrerprofessionalisierung wissenschaftlich begleiten. Strategien und Methoden* (S. 117-130). Münster: Waxmann.
- Berliner, D. C. (2004). Describing the behavior and documenting the accomplishments of expert teachers. *Bulletin of Science, Technology & Society*, 24, 200-212.
- Blanck, J. M. (2015). Schulische Integration und Inklusion in Deutschland. Eine vergleichende Betrachtung der 16 Bundesländer. In P. Kuhl, P. Stanat, B. Lütje-Klose, C. Gresch, H. A. Pant & M. Prenzel (Hrsg.), *Inklusion von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf in Schulleistungserhebungen* (S. 153-177). Wiesbaden: Springer.
- Blömeke, S. (2002). *Universität und Lehrerbildung*. Bad Heilbrunn: Klinkhardt.
- Blömeke, S., Gustafsson, J.-E. & Shavelson, R. J. (2015). Beyond dichotomies. Competence viewed as a continuum. *Zeitschrift für Psychologie*, 223(1), 3-13.
- Blumenthal, S., Blumenthal, Y. & Mahlau, K. (2022). *Kinder mit Lern- und emotional-sozialen Entwicklungsauffälligkeiten in der Schule. Diagnostik – Prävention – Förderung*. Stuttgart: Kohlhammer.

- Blumenthal, Y., Casale, G., Hartke, B., Hennemann, T., Hillenbrand, C. & Vierbuchen, M. C. (2020). *Kinder mit Verhaltensauffälligkeiten und emotional-sozialen Entwicklungsstörungen – Förderung in inklusiven Schulklassen*. Stuttgart: Kohlhammer.
- Blumenthal, Y., Kuhlmann, K. & Hartke, B. (2014). Diagnostik und Prävention von Lernschwierigkeiten im Aptitude Treatment Interaction- (ATI-) und Response to Intervention- (RTI-)Ansatz. In M. Hasselhorn, W. Schneider & U. Trautwein, U. (Hrsg.), *Tests & Trends: Formative Leistungsdiagnostik* (NF Bd. 12, S. 61-81). Göttingen: Hogrefe.
- Blumenthal, Y. & Mahlau, K. (2015). Effektiv fördern – Wie wähle ich aus? Ein Plädoyer für die Evidenzbasierte Praxis in der schulischen Sonderpädagogik. *Zeitschrift für Heilpädagogik*, 66(9), 408-421.
- Bonnet, A. & Hericks, U. (2014). Professionalisierung und Deprofessionalisierung im Lehrer/innenberuf. Ansätze und Befunde aktueller empirischer Forschung. *Zeitschrift für interpretative Schul- und Unterrichtsforschung*, 3, 3-13.
- Bortz, J. & Döring, N. (2016). *Forschungsmethoden und Evaluation in den Sozialwissenschaften* (5. Auflage). Berlin: Springer.
- Boylan, M., Coldwell, M., Maxwell, B. & Jordan, J. (2018). Rethinking models of professional learning as tools: a conceptual analysis to inform research and practice. *Professional Development in Education*, 44(1), 120-139.
- Brezinka, V. (2003). Zur Evaluation von Präventivinterventionen für Kinder mit Verhaltensstörungen. *Kindheit und Entwicklung*, 12(2), 71-83.
- Brodesser, E., Frohn, J., Welskop, N., Liebsch, A.-C., Moser, V. & Pech, D. (2020). *Inklusionsorientierte Lehr-Lern-Bausteine für die Hochschullehre. Ein Konzept zur Professionalisierung zukünftiger Lehrkräfte*. Bad Heilbrunn: Verlag Julius Klinkhardt.
- Bromme, R. (1992). *Der Lehrer als Experte. Zur Psychologie des professionellen Wissens*. Bern: Huber.
- Bromme, R. (2001). *Teacher expertise. International encyclopedia of the social & behavioral sciences*. Amsterdam: Smelser and Baltes.
- Brossart, D. F., Laird, V. C. & Armstrong, T. W. (2018). Interpreting Kendall's Tau and Tau- U for single-case experimental designs. *Cogent Psychology*, 5(1), 1-26. <https://doi.org/10.1080/23311908.2018.1518687>.
- Brunner, M., Kunter, M., Krauss, S., Baumert, J., Blum, W., Dubberke, T., Jordan, A., Klusmann, U., Tsai, Y.-M. & Neubrand, M. (2006). Welche Zusammenhänge bestehen zwischen dem fachspezifischen Professionswissen von Mathematiklehrkräften und ihrer Ausbildung sowie beruflichen Fortbildung? *Zeitschrift für Erziehungswissenschaft*, 9(4), 521-544.
- Brunstein, J. & Julius, H. (2014). Evaluationen von Interventionen durch Einzelfallstudien. In G. Lauth, M. Grünke & J. Brunstein (Hrsg.), *Interventionen bei Lernstörungen* (S. 119-138). Göttingen: Hogrefe.
- Campos, J. J., Frankel, C. B. & Camras, L. (2004). On the nature of emotion regulation. *Child Development*, 75, 377-394.
- Casale, G., Grosche, M., Volpe, R. J. & Hennemann, T. (2017). Zuverlässigkeit von Verhaltensverlaufsdiagnostik über Rater und Messzeitpunkte bei Schülern mit externalisierenden Verhaltensproblemen. *Empirische Sonderpädagogik*, 9(2), 143-164.

- Casale, G., Hennemann, T., & Grosche, M. (2015a). Zum Beitrag der Verlaufsdagnostik für eine evidenzbasierte sonderpädagogische Praxis am Beispiel des Förderschwerpunkts der emotionalen und sozialen Entwicklung. *Zeitschrift für Heilpädagogik*, 66(7), 325-334.
- Casale, G., Hennemann, T. & Hövel, D. (2014). Systematischer Überblick über deutschsprachige schulbasierte Maßnahmen zur Prävention von Verhaltensstörungen in der Sekundarstufe I. *Empirische Sonderpädagogik*, 6(1), 33-58.
- Casale, G., Hennemann, T., Huber, C., & Grosche, M. (2015b). Testgütekriterien der Verlaufsdagnostik von Schülerverhalten im Förderschwerpunkt Emotionale und soziale Entwicklung. *Heilpädagogische Forschung*, 41(1), 37-54.
- Casale, G., Huber, C., Grosche, M. & Hennemann, T. (2019). *Verhaltensverlaufsdagnostik in der Schule – die Methode der Direkten Verhaltensbeurteilung in der Praxis*. München: Ernst Reinhardt.
- Casale, G., Volpe, R. J., Briesch, A. M., Hennemann, T., & Grosche, M. (2021). Dependability of Direct Behavior Rating Single- and Multi-Item Scales Across Raters and Occasions in Two School Subjects. *Assessment for Effective Intervention*, 46(2), 143-154.
- Castello, A. (2017). *Schulische Inklusion bei psychischen Auffälligkeiten*. Stuttgart: Kohlhammer.
- Chafouleas, S. M. (2011). Direct Behavior Rating: A review of the issues and research in its development. *Education & Treatment of Children*, 34(4), 575-591.
- Chafouleas, S. M., Jaffery, R., Riley-Tillman, T. C., Christ, T. J. & Sen, R. (2013). The Impact of Target, Wording, and Duration on Rating Accuracy for Direct Behavior Rating. *Assessment for Effective Intervention*, 39(1), 39-53.
- Chafouleas, S. M., Sanetti, L. M. H., Kilgus, S. P. & Maggin, D. M. (2012). Evaluating Sensitivity to Behavioral Change Using Direct Behavior Rating Single-Item Scales. *Exceptional Children*, 78(4), 491-505.
- Cheney, D., Flower, A. & Templeton, T. (2008). Applying response to intervention metrics in the social domain for students at risk of developing emotional or behavioral disorders. *The Journal of Special Education*, 42(2), 108-126.
- Christ, T. J., Riley-Tillman, T. C. & Chafouleas, S. M. (2009). Foundation for the development and use of direct behavior rating (DBR) to assess and evaluate student behavior. *Assessment for Effective Intervention*, 34(4), 201-213.
- Clarke, D. & Hollingsworth, H. (2002). Elaborating a model of teacher professional growth. *Teaching and teacher education*, 18(8), 947-976.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2. Aufl.). Hillsdale, NJ: Erlbaum.
- Conley, L., Marchant, M. & Caldarella, P. (2014). A Comparison of Teacher Perceptions and Research-Based Categories of Student Behavior Difficulties. *Education*, 134, 439-451.
- Copur-Gencturk, Y. & Thacker, I. (2020). A comparison of perceived and observed learning from professional development: Relationships among self-reports, direct assessments, and teacher characteristics. *Journal of Teacher Education*, 82(3). <https://doi.org/10.1177/0022487119899101>.
- Correnti, R. (2007). An empirical investigation of professional development effects on literacy instruction using daily logs. *Educational Evaluation and Policy Analysis*, 29(4), 262-295.

- Crick, N. R. & Dodge, K. A. (1994). A review and reformulation of social information-processing mechanisms in children's social adjustment. *Psychological bulletin*, 115(1), 74.
- Daniels, B., Volpe, R. J., Briesch, A. M. & Gadow, K. D. (2017). Dependability and Treatment Sensitivity of Multi-Item Direct Behavior Rating Scales for Interpersonal Peer Conflict. *Assessment for Effective Intervention*, 1-12.
- Darling-Hammond, L., Hyler, M. E. & Gardner, M. (2017). *Effective teacher professional development*. Palo Alto, CA: Learning Policy Institute.
- Dawson, N., McHugh, B. & Asen, E. (2020). *Die Familienklasse. Multifamiliengruppenarbeit an Schulen*. Dortmund: verlag modernes lernen.
- de Boer, A., Pijl, S. J. & Minnaert, A. (2011). Regular primary schoolteachers' attitudes towards inclusive education: A review of the literature. *International Journal of Inclusive Education*, 15(3), 331-353. <https://doi.org/10.1080/13603110903030089>.
- Deci, E. L. & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. New York: Plenum Publishing Co.
- Denham, S. A. (1998). *Emotional development in young children*. New York: Guilford.
- Denham, S. A. & Burton, R. (2003). *Social and emotional prevention and intervention programming for preschoolers*. New York: Kluwer Academic/Plenum.
- Desimone, L. M. (2009). Improving impact studies of teachers' professional development: Toward better conceptualizations and measures. *Educational Researcher*, 38(3), 181-199.
- Desimone, L. M., Porter, A. C., Garet, M., Suk Yoon, K. & Birman, B. (2002). Effects of professional development on teachers' instruction: Results from a three-year study. *Educational Evaluation and Policy Analysis*, 24(2), 81-112.
- Dodge, K. A., & Pettit, G. S. (2003). A biopsychosocial model of the development of chronic conduct problems in adolescence. *Developmental Psychology*, 39(2), 349-371. <https://doi.org/10.1037/0012-1649.39.2.349>.
- Döbert, H. & Weishaupt, H. (2013). *Inklusive Bildung professionell gestalten. Situationsanalyse und Handlungsempfehlungen*. Münster: Waxmann.
- Döpfner, M. (1989). Soziale Informationsverarbeitung – ein Beitrag zur Differenzierung sozialer Inkompetenzen. *Zeitschrift für pädagogische Psychologie*, 3, 1-8.
- Döpfner, M. (2003). Wie wirksam ist Kinder- und Jugendlichenpsychotherapie? *Psychotherapeutenjournal*, 1, 258-266.
- Dresing, T. & Pehl, T. (2018). *Praxisbuch Interview, Transkription & Analyse. Anleitungen und Regelsysteme für qualitativ Forschende*. Marburg: Eigenverlag.
- Durksen, T., Klassen, R. M. & Daniels, L. M. (2017). Motivation and collaboration: The keys to a developmental framework for teachers' professional learning. *Teaching and Teacher Education*, 67, 53-66.
- Durlak, J. A. & DuPre, E. P. (2008). Implementation Matters: A Review of Research on the Influence of Implementation on Program Outcomes and the Factors Affecting Implementation. *American Journal of Community Psychology*, 41, 327-350.
- Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D. & Schellinger, K. B. (2011). The impact of enhancing student's social and emotional learning: a meta-analysis of school-based universal interventions. *Child Development*, 82(1), 405-432.

- Dyson, A. (2010). Developing inclusive schools: three perspectives from England. *DDS – Die Deutsche Schule*, 102(2), 115-129.
- Ellinger, S. & Stein, R. (2012). Effekte inklusiver Beschulung: Forschungsstand im Förderschwerpunkt emotionale und soziale Entwicklung, *Empirische Sonderpädagogik*, 4(2), 85-109.
- Evans, L. (2014). Leadership for professional development and learning: enhancing our understanding of how teachers develop. *Cambridge Journal of Education*, 44(2), 179-198.
- Evertson, C. M. & Emmer, E. T. (2009). *Classroom management for elementary teachers*. New Jersey: Pearson Education.
- Faix, A.-C., Wild, E., Lütje-Klose, B. & Textor, A. (2019). Professionalisierung für inklusiven Unterricht im Rahmen interdisziplinärer und videogestützter Lehrveranstaltungen. *Journal für Psychologie*, 27(2), 71–94. <https://doi.org/10.30820/0942-2285-2019-2-71>.
- Fischer, N. (2006). *Motivationsförderung in der Schule. Konzeption und Evaluation einer Fortbildungsmaßnahme für Mathematiklehrkräfte*. Hamburg: Kovac.
- Fishman, B. J., Marx, R. W., Best, S. & Tal, R. T. (2003). Linking teacher and student learning to improve professional development in systemic reform. *Teaching and Teacher Education*, 19, 643-658.
- Frey, A. (2014). Kompetenzmodelle und Standards in der Lehrerbildung und im Lehrerberuf. In E. Terhart, H. Bennewitz & M. Rothland (Hrsg.), *Handbuch der Forschung zum Lehrerberuf* (2., überarbeitete und erweiterte Aufl.) (S. 712-744). Münster: Waxmann.
- Fussangel, K., Rürup, M. & Gräsel, C. (2016). Lehrerfortbildung als Unterstützungssystem. In H. Altrichter & K. Maag Merki (Hrsg.), *Handbuch Neue Steuerung im Schulsystem* (S. 361-384). Wiesbaden: Springer.
- Garet, M. S., Porter, A. C., Desimone, L., Birman, B. F. & Yoon, K. S. (2001). What makes professional development effective? Results from a national sample of teachers. *American Educational Research*, 38(4), 915-945.
- Gärtner, H. (2007). *Unterrichtsmonitoring*. Münster: Waxmann.
- Georgiades, K., Duncan, L., Wang, L., Comeau, J., Boyle, M. H. & 2014 Ontario Child Health Study Team. (2019). Six-month prevalence of mental disorders and service contacts among children and youth in Ontario: Evidence from the 2014 Ontario Child Health Study. *The Canadian Journal of Psychiatry*, 64(4), 246-255.
- Ghandour, R. M., Sherman, L. J., Vladutiu, C. J., Ali, M. M., Lynch, S. E., Bitsko, R. H. & Blumberg, S. J. (2019). Prevalence and treatment of depression, anxiety, and conduct problems in US children. *The Journal of pediatrics*, 206, 256-267.
- Göbel, K., Baumgarten, F., Kuntz, B., Hölling, H. & Schlack, R. (2018). ADHS bei Kindern und Jugendlichen in Deutschland – Querschnittergebnisse aus KiGGS-Welle 2 und Trends. *Journal of Health Monitoring*, 3(3), 46-53.
- Goldschmidt, P. & Phelps, G. (2007). *Does teacher professional development affect content and pedagogical knowledge: How much and for how long?* National Center for Research on Evaluation, Standards, and Student Testing (CRESST), Center for the Study of Evaluation (CSE), Graduate School of Education and Information Studies University of California, Los Angeles.
- Goodman, R. (2005). *Fragebogen zu Stärken und Schwächen (SDQ-Deu)*. Zugriff am 20.09.2022 auf <https://www.sdqinfo.org/py/sdqinfo/b3.py?language=German>.

- Gottfredson, M. R. & Hirschi, T. (1990). *A general theory of crime*. Stanford University Press.
- Gottfried, K., Casale, G., Hennemann, T., Huber, C., Kaspar, K., Spilles, M., Strauss, S. & König, J. (2021). Adaptiver Umgang mit externalisierenden Verhaltensproblemen: Pädagogisches Wissen zu inklusivem Unterricht mit Fokus emotionale und soziale Entwicklung. *Schweizerische Zeitschrift für Bildungswissenschaften*, 43(2), 260-272.
- Gräsel, C. & Trempler, K. (2017). Einleitung. In C. Gräsel & K. Trempler (Hrsg.), *Entwicklung von Professionalität pädagogischen Personals. Interdisziplinäre Betrachtungen, Befunde und Perspektiven* (S. 1-16). Wiesbaden: Springer.
- Gresham, F. M., VanDerHeyden, A. & Witt, J. C. (2005). *Response to Intervention in the Identification of Learning Disabilities: Empirical Support and Future Challenges*. Zugriff am 01.08.2024 auf <http://www.joewitt.org/Downloads/Response%20to%20Intervention%20MS%20Gresham%20%20Vanderheyden%20Witt.pdf>.
- Grob, A., Meyer, C. & Hagmann von Arx, P. (2009). *Intelligence and Development Scales*. Bern: Hans Huber, Hogrefe.
- Grüne Liste Prävention (2022). *Die Empfehlungsliste evaluierter Präventionsprogramme*. Zugriff am 17.11.2022 auf <https://www.gruene-liste-praevention.de/nano.cms/datenbank/alle>.
- Guskey, T. R. (2002). Professional development and teacher change. *Teachers and Teaching: Theory and Practice*, 8(3/4), 381-391.
- Guskey, T. (2014). Measuring the effectiveness of educators' professional development. In L. E. Martin, S. Kragler, D. J. Quatroche & K. L. Bauserman (Hrsg.), *Handbook of professional development in education: Successful models and practices, PreK-12* (S. 447-465). New York: The Guilford Press.
- Guthrie, J. T., Wigfield, A., Barbosa, P., Perencevich, K. C., Taboada, A., Davis, M. H., Scaffidi, N. T. & Tonks, S. (2004). Increasing reading comprehension and engagement through conceptoriented reading instruction. *Journal of Educational Psychology*, 96(3), 403-423.
- Hartke, B. (2017). *Handlungsmöglichkeiten Schulische Inklusion. Das Rügener Modell kompakt*. Stuttgart: Kohlhammer.
- Hartke, B., Blumenthal, Y., Carnein, O. & Vrban, R. (2016). *Schwierige Schüler – Sekundarstufe: 64 Handlungsmöglichkeiten bei Verhaltensauffälligkeiten (5.-10. Klasse)* (3. Auflage). Hamburg: Persen.
- Hartke, B., Blumenthal, Y., Carnein, O. & Vrban, R. (2018). *Schwierige Schüler – Förderschule: 84 Handlungsmöglichkeiten bei Verhaltensauffälligkeiten und sonderpädagogischem Förderbedarf (1. bis 10. Klasse)*. Hamburg: Persen.
- Hartke, B. & Vrban, R. (2011). *Schwierige Schüler – 49 Handlungsmöglichkeiten bei Verhaltensauffälligkeiten* (6. Aufl.). Buxtehude: Persen.
- Hattie, J. (2018). *Lernen sichtbar machen für Lehrpersonen. Überarbeitete deutschsprachige Ausgabe von „Visible Learning for Teachers“*. Hohengehren: Schneider.
- Havers, N. & Toepell, S. (2002). Trainingsverfahren für die Lehrerbildung im deutschen Sprachraum. *Zeitschrift für Pädagogik*, 48(2), 174-193.

- Hecht, P., Niedermair, C. & Feyerer, E. (2016). Einstellungen und inklusionsbezogene Selbstwirksamkeitsüberzeugungen von Lehramtsstudierenden und Lehrpersonen im Berufseinstieg. Messverfahren und Befunde aus einem Mixed-Methods-Design. *Empirische Sonderpädagogik*, 8(1), 86-102.
- Heinrichs, N., Döpfner, M. & Petermann, F. (2013). Prävention psychischer Störungen. In F. Petermann (Hrsg.), *Lehrbuch der Klinischen Kinderpsychologie* (S. 721-738). Hogrefe: Göttingen.
- Heinrichs, N., Saßmann, H., Halhweg, K. & Perrez, M. (2002). Prävention kindlicher Verhaltensstörungen. *Psychologische Rundschau*, 53(4), 170-183.
- Helmke, A. (2010). *Unterrichtsqualität und Lehrerprofessionalität. Diagnose, Evaluation und Verbesserung der Unterrichtsqualität*. Seelze: Kallmeyer.
- Hennemann, T., Hövel, D., Casale, G., Hagen, T. & Fitting-Dahlmann, K. (2017). *Schulische Prävention im Bereich Verhalten*. Fördern lernen – Prävention, Band 19. Stuttgart: Kohlhammer.
- Hennemann, T., Ricking, H. & Huber, C. (2015). Organisationsformen inklusiver Förderung im Bereich emotional-sozialer Entwicklung. In R. Stein & T. Müller (Hrsg.), *Inklusion im Förderschwerpunkt emotionale und soziale Entwicklung* (S. 110-144). Stuttgart: Kohlhammer.
- Heran-Dörr, E. (2006). *Entwicklung und Evaluation einer Lehrerfortbildung zur Förderung der physikdidaktischen Kompetenz von Sachunterrichtslehrkräften. Eine explorative Studie*. Berlin: Logos.
- Herzmann, P. & König, J. (2016). *Lehrerberuf und Lehrerbildung*. Bad Heilbrunn: Klinkhardt.
- Herzog, W. & Makarova, E. (2012). Anforderungen an und Leitbilder für den Lehrerberuf. In E. Terhart, H. Bennewitz & M. Rothland (Hrsg.), *Handbuch der Forschung zum Lehrerberuf* (S. 83-102). Münster: Waxmann.
- Heyvaert, M., & Onghena, P. (2014). Analysis of single-case data: Randomization tests for measures of effect size. *Neuropsychological Rehabilitation*, 24, 507-527. <https://doi.org/10.1080/09602011.2013.818564>.
- Hillenbrand, C. (2008). *Einführung in die Pädagogik der Verhaltensstörungen*. München: Reinhardt.
- Hillenbrand, C. (2013). Inklusive Bildung in der Schule: Probleme und Perspektiven für die Bildungsberichterstattung. *Zeitschrift für Heilpädagogik*, 64(9), 359-369.
- Hillenbrand, C. (2015). Evidenzbasierte Praxis im Förderschwerpunkt emotional-soziale Entwicklung. In R. Stein & T. Müller (Hrsg.), *Inklusion im Förderschwerpunkt emotionale und soziale Entwicklung* (S. 170-216). Stuttgart: Kohlhammer.
- Hillenbrand, C. & Hennemann, T. (2006). Präventive Erziehungshilfe in der Grundschulstufe. *Zeitschrift für Heilpädagogik*, 57(2), 42-51.
- Hillenbrand, C. & Pütz, K. (2008). *KlasseKinderSpiel. Spielerisch Verhaltensregeln lernen*. Hamburg: edition Körber Stiftung.
- Hochholdinger, S. & Schaper, N. (2007). Trainingsevaluation und Transfersicherung. In H. Schuler & K. Sonntag (Hrsg.), *Handbuch der Arbeits- und Organisationspsychologie* (S. 625-633). Göttingen: Hogrefe.
- Hövel, D. (2014). *Adaption und Evaluation des Präventionsprogrammes „Lubo aus dem All!“ für Kinder mit hohen Risikobelastungen*. Dissertation: Universität Oldenburg.

- Holodynski, M. & Friedlmeier, W. (2006). *Emotionen – Entwicklung und Regulation*. Heidelberg: Springer.
- Holodynski, M., Hermann, S. & Kromm, H. (2013). Entwicklungspsychologische Grundlagen der Emotionsregulation. *Psychologische Rundschau*, 64(4), 196-207.
- Horn, K.-P. (2016). Profession, Professionalisierung, Professionalität, Professionalismus – Historische und systematische Anmerkungen am Beispiel der deutschen Lehrerbildung. *Zeitschrift für Pädagogik und Theologie*, 68(2), 153-164.
- Huber, C. (2015a). Verhaltensverlaufsdiagnostik. In K. Seifried, S. Drewes & M. Hasselhorn (Hrsg.), *Handbuch Schulpsychologie: Psychologie für die Schule* (S. 150-160). Stuttgart: Kohlhammer Verlag.
- Huber, C. (2015b). Verhaltensprobleme gemeinsam lösen! Wie sich multiprofessionelle Teams nachdem RTI-Modell effektiv organisieren lassen. *Lernen und Lernstörungen*, 4(4), 283-291.
- Huber, C. & Rietz, C. (2015). Direct Behavior Rating (DBR) als Methode zur Verhaltensverlaufsdiagnostik in der Schule: Ein systematisches Review von Methodenstudien. *Empirische Sonderpädagogik*, 7(2), 75-98.
- Huber, S. G. (2010). Schulleitung international. In T. Bohl, W. Helsper, H. G. Holtappels & C. Schelle (Hrsg.), *Handbuch Schulentwicklung. Theorie – Forschungsbefunde – Entwicklungsprozesse – Methodenrepertoire* (S. 213-221). Bad Heilbrunn: Klinkhardt.
- Hugger, K. U., Dinkelaker, J., Idel, T.-S., Schütz, A. & Thünemann, S. (2021). *Professionalität und Professionalisierung in pädagogischen Handlungsfeldern: Schule, Medienpädagogik, Erwachsenenbildung*. Opladen & Toronto: Barbara Budrich.
- Huitema, B. E. & Mckean, J. W. (2000). Design Specification Issues in Time-Series Intervention Models. *Educational and Psychological Measurement*, 60(1), 38-58. <https://doi.org/10.1177/00131640021970358>.
- Hulac, D. M. & Briesch, A. M. (2017). *Evidence-based strategies for effective classroom management*. New York: The Guilford Press.
- Ihle, W. & Esser, G. (2008). Epidemiologie psychischer Störungen des Kindes- und Jugendalters. In H. Julius & B. Gasteiger-Klicpera (Hrsg.), *Sonderpädagogik der sozialen und emotionalen Entwicklung. Band 3 Handbuch Sonderpädagogik* (S. 49-62). Göttingen: Hogrefe.
- Jäger, R. & Bodensohn, R. (2007). Die Situation der Lehrerfortbildung im Fach Mathematik aus Sicht der Lehrkräfte. Ergebnisse einer Befragung von Mathematiklehrern. Bonn: Deutsche Telekomstiftung.
- Jahnukainen, M. & Itkonen, T. (2015). Tiered intervention: history and trends in Finland and the United States. *European Journal of Special Needs Education*. <https://doi.org/10.1080/08856257.2015.1108042>.
- Jain, A. & Spieß, R. (2012). Versuchspläne der experimentellen Einzelfallforschung. *Empirische Sonderpädagogik*, 4(3), 211-245.
- Janke, B. (2007). Entwicklung von Emotionen. In M. H. Schneider (Hrsg.), *Handbuch der Entwicklungspsychologie* (S. 347-358). Göttingen: Hogrefe.
- Janke, B. (2008). Emotionswissen und Sozialkompetenz von Kindern im Alter von drei bis zehn Jahren. *Empirische Pädagogik*, 22, 127-144.
- Jürgens, B. & Krause, G. (2013). *Professionalisierung durch Trainings*. Aachen: Shaker.

- Kaiser, G., Blömeke, S., König, J., Busse, A., Döhrmann, M. & Hoth, J. (2017). Professional competencies of (prospective) mathematics teachers – cognitive versus situated approaches. *Educational Studies in Mathematics*, 94(2), 161-182.
- Kanning, U. P. (2009). *Diagnostik sozialer Kompetenzen*. Göttingen: Hogrefe.
- Kany, W. & Schöler, H. (2009). *Diagnostik schulischer Lern- und Leistungsschwierigkeiten. Ein Leitfaden*. Stuttgart: Kohlhammer.
- Kazdin, A. E. (2011). *Single-case research designs: Methods for clinical and applied settings*. Oxford University Press.
- Kazdin, A. E. (2021). Single-case experimental designs: Characteristics, changes and challenges. *Journal of Experimental Analysis of Behavior*, 115(1), 56-85.
- Kiemer, K., Gröschner, A., Pehmer, A.-K. & Seidel, T. (2015). Effects of a classroom discourse intervention on teachers' practice and students' motivation to learn mathematics and science. *Learning and Instruction*, 35, 94-103.
- King, F. (2014). Evaluating the impact of teacher professional development: an evidence-based framework. *Professional Development in Education*, 40(1), 89-111.
- Klasen, F., Reiß, F., Otto, C., Haller, A.-C., Meyrose, A.-K., Barthel, D. & Ravens-Sieberer, U. (2017). Die BELLA-Studie – das Modul zur psychischen Gesundheit in KiGGS Welle 2. *Journal of Health Monitoring*, 2, 55-65.
- Klauer, K. & Lauth, G. (1997). Lernbehinderungen und Leistungsschwierigkeiten bei Schülern. In F. E. Weinert (Hrsg.), *Psychologie des Unterrichts und der Schule* (S. 701-738). Göttingen: Hogrefe.
- Kleickmann, T. & Möller, C. (2007). Haben Lehrerfortbildungen einen Effekt auf Leistungszuwächse bei Schülerinnen und Schülern? In D. Höttecke (Hrsg.), *Naturwissenschaftlicher Unterricht im internationalen Vergleich* (S. 506-508). Münster: LTI Verlag.
- Klemm, K. (2018). *Unterwegs zur inklusiven Schule. Lagebericht 2018 aus bildungsstatistischer Perspektive*. Bertelsmann-Stiftung. Zugriff am 01.08.2024 auf https://www.bertelsmann-stiftung.de/fileadmin/files/BS/Publikationen/GrauePublikationen/Studie_IB_Unterwegs-zur-inklusive-Schule_2018.pdf.
- Koglin, U. & Petermann, F. (2008). Kindergarten- und Grundschulzeit: Entwicklungsrisiken und -abweichungen. In F. Petermann (Hrsg.), *Lehrbuch der Klinischen Kinderpsychologie* (S. 81-98) Göttingen: Hogrefe.
- Koglin, U., Petermann, F., Jaščenoka, J., Petermann, U. & Kullik, A. (2013). Emotionsregulation und aggressives Verhalten im Jugendalter. *Kindheit und Entwicklung*, 22, 155-164.
- König, J. (2010). Lehrerprofessionalität – Konzepte und Ergebnisse der internationalen und deutschen Forschung am Beispiel fachübergreifender, pädagogischer Kompetenzen. In J. König & B. Hofmann (Hrsg.), *Professionalität von Lehrkräften – Was sollen Lehrkräfte im Lese- und Schreibunterricht wissen und können?* (S. 40-105). Berlin: DGLS.
- König, J., Gerhard, K., Kaspar, K. & Melzer, C., (2019). Professionelles Wissen von Lehrkräften zur Inklusion: Überlegungen zur Modellierung und Erfassung mithilfe standardisierter Testinstrumente. *Pädagogische Rundschau*, 73(1), 43-64.
- Kratochwill, T. R., Hitchcock, J., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. & Shadish, W. R. M. (2010). *Single case designs technical documentation*. Zugriff am 20.09.2022 auf <https://files.eric.ed.gov/fulltext/ED510743.pdf>.
- Kullik, A. & Petermann, F. (2012). *Emotionsregulation im Kindesalter*. Göttingen: Hogrefe.

- Kultusministerkonferenz (2000). *Empfehlungen zum Förderschwerpunkt emotionale und soziale Entwicklung*. Zugriff am 01.08.2024 auf https://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2000/2000_03_10-FS-Emotionale-soziale-Entw.pdf.
- Kultusministerkonferenz (2022). *Statistische Veröffentlichungen der Kultusministerkonferenz. Dokumentation Nr. 231 – Januar 2022. Sonderpädagogische Förderung in Schulen 2011 bis 2020*. Zugriff am 01.08.2024 auf https://www.kmk.org/fileadmin/Dateien/pdf/Statistik/Dokumentationen/Dok231_SoPaeFoe_2020.pdf.
- Kultusministerkonferenz (2018). *Statistische Veröffentlichungen der Kultusministerkonferenz. Dokumentation Nr. 223 – Februar 2020. Sonderpädagogische Förderung in Schulen 2009 bis 2018*. Zugriff am 01.08.2024 auf https://www.kmk.org/fileadmin/Dateien/pdf/Statistik/Dokumentationen/Dok223_SoPae_2018.pdf.
- 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.
- Kunter, M., Baumert, J., Blum, W., Klusmann, U., Krauss, S. & Neubrand, M. (2011). *Professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV*. Münster: Waxmann.
- Kunter, M., Klusmann, U. & Baumert, J. (2009). Professionelle Kompetenz von Mathematiklehrkräften: Das COACTIV-Modell. In O. Zlatkin-Troitschanskaia, K. Beck, D. Sembill, R. Nickolaus & R. Mulder (Hrsg.), *Lehrerprofessionalität – Bedingungen, Genese, Wirkungen und ihre Messung* (S. 153-165). Weinheim: Beltz.
- Kurniawati, F., de Boer, A., Minnaert, A. E. M. G. & Mangunsong, F. (2014). Characteristics of primary teacher training programmes on inclusion: a literature focus. *Educational Research*, 56(3), 310-326.
- Kwakman, K. (2003). Factors influencing teachers' participation in professional learning activities. *Teaching and Teacher Education*, 19, 149-170.
- Laucht, M., Esser, G. & Schmidt, M. H. (1997). Wovon schützen Schutzfaktoren? Anmerkungen zu einem populären Konzept der modernen Gesundheitsforschung. *Zeitschrift für Entwicklungspsychologie und pädagogische Psychologie*, 29(3), 260-270.
- Lauth, G. W. & Heubeck, B. (2006). *Kompetenztraining für Eltern sozial auffälliger Kinder (KES)*. Göttingen: Hogrefe.
- Lehrergewerkschaften Hamburg (2022, 24. Februar). *Hamburgs Bildungsqualität und Chancengleichheit gefährdet. Lehrkräfte- und Personalmangel viel dramatischer als befürchtet! VBE Hamburg zur Veröffentlichung der Klemm-Studie 02.2022*. Zugriff am 01.08.2024 auf <https://www.lehrergewerkschaften-hamburg.de/aktuelles/news/hamburgs-bildungsqualitaet-und-chancengleichheit-gefaehrdet/>.
- Lei, P.-W., Smith, M. & Suen, H. K. (2007). The use of generalizability theory to estimate data reliability in single-subject observational research. *Psychology in the Schools*, 44, 433-439. <https://doi.org/10.1002/pits.20235>
- Leidig, T. (2019). *Wie kann es gelingen? – Professionalisierung von Lehrkräften auf dem Weg zum inklusiven Schulsystem unter besonderer Berücksichtigung prozessbegleitender Fortbildungsangebote*. Dissertation, Universität zu Köln.

- Leidig, T., Hennemann, T., Casale, G. & Hillenbrand, C. (2016). Wirksamkeit von Lehrerfortbildungen zur inklusiven Beschulung im Förderschwerpunkt Emotionale und soziale Entwicklung – ein systematisches Review empirischer Studien. *Heilpädagogische Forschung*, 42(2), 54-70.
- Leko, M. M. & Roberts, C. A. (2014). How Does Professional Development Improve Teacher Practice in Inclusive Schools? In J. McLeskey, N.L. Waldron, F. Spooner & B. Algozzine (Hrsg.), *Handbook of effective inclusive schools: Research and practice* (S. 43-54). New York: Routledge.
- Lemerise, E. A. & Arsenio, W. F. (2000). An integrated model of emotion processes and cognition in social information processing. *Child Development*, 71(1), 107-118.
- Leinhard, W. (2021). *Erleben, Lernen und Verhalten von Kindern und Jugendlichen: Wie Schule mit Auffälligkeiten umgehen kann*. Stuttgart: Kohlhammer Verlag.
- Liebner, S., Wagner, S., Kobs, S., Rogge, F., Krauskopf, K. & Knigge, M. (2022). Kollegiale Fallberatung in der inklusionsorientierten Lehrkräftebildung – Ein Instrument zur Förderung der Selbstwirksamkeit angehender Lehrkräfte? In A. Schröter, M. Kortmann, S. Schulze, K. Kempfer, S. Anderson, G. Sevdiren, J. Bartz & C. Kreutchen (Hrsg.), *Inklusion in der Lehramtsausbildung – Lerngegenstände, Interaktionen und Prozesse* (S. 131-146). Münster: Waxmann.
- Linderkamp, F. (2022). Lern- und Verhaltensstörungen: Bedingungsgefüge und Implikationen für die Sonderpädagogische Diagnostik. In Y. Blumenthal, S. Blumenthal & K. Mahlau (Hrsg.), *Kinder mit Lern- und emotional-sozialen Entwicklungsauffälligkeiten in der Schule. Diagnostik – Prävention – Förderung* (S. 31-40). Stuttgart: Kohlhammer.
- Lindmeier, C. (2022). Sonderpädagogische Beratung in der inklusiven Schule – Ein phasenübergreifendes Beratungscurriculum für sonderpädagogische Lehrkräfte als Basis gelingender Teamkooperation. In B. Serke & B. Streese (Hrsg.), *Wege der Kooperation im Kontext inklusiver Bildung* (S. 173-182). Bad Heilbrunn: Klinkhardt.
- Lipowsky, F. (2014). Theoretische Perspektiven und empirische Befunde zur Wirksamkeit von Lehrerfort- und -weiterbildung. In E. Terhart, H. Bannewitz & M. Rothland (Hrsg.), *Handbuch der Forschung im Lehrerberuf* (2. Aufl.) (S. 511-541). Münster/New York: Waxmann.
- Lipowsky, F. (2010). Lernen im Beruf. Empirische Befunde zur Wirksamkeit von Lehrerfortbildung. In F. H. Müller (Hrsg.), *Lehrerinnen und Lehrer lernen. Konzepte und Befunde zur Lehrerfortbildung* (S. 51-70.). Münster: Waxmann.
- Lipowsky, F., Kuntze, S., Ratzka, N., Klieme, E. & Reiss, K. (2006). *Unterricht entwickeln und verbessern – Was leisten videobasierte Lehrerfortbildungen?* Abschlussbericht an die Robert Bosch Stiftung.
- Lipowsky, F. & Rzejak, D. (2015). Das Lernen von Lehrpersonen und Schülern/-innen im Fokus. Was zeichnet wirksame Lehrerfortbildungen aus? *Forum Lehrerfortbildung*, 46, 11-50.
- Lipowsky, F. & Rzejak, D. (2019). Was macht Fortbildungen für Lehrkräfte erfolgreich? – Ein Update. In B. Groot-Wilken & R. Koerber (Hrsg.), *Nachhaltige Professionalisierung für Lehrerinnen und Lehrer. Ideen, Entwicklungen, Konzepte*. wbv Media: Bielefeld.

- Loeber, R. (1990). Development and risk factors of juvenile antisocial behavior and delinquency. *Clinical Psychology Review*, 10(1), 1-41. [https://doi.org/10.1016/0272-7358\(90\)90105-J](https://doi.org/10.1016/0272-7358(90)90105-J).
- Lösel, F. & Beelmann, A. (2003). Effects of Child Skills Training in Preventing Antisocial Behavior: A Systematic Review of Randomized Evaluations. *The ANNALS of the American Academy of Political and Social Science*, 587(1), 84-109. <https://doi.org/10.1177/0002716202250793>.
- Lösel, F. & Bender, D. (2008). Von generellen Schutzfaktoren zu spezifischen protektiven Prozessen: Konzeptuelle Grundlagen und Ergebnisse der Resilienzforschung. In G. Opp & M. Fingerle (Hrsg.), *Was Kinder stärkt. Erziehung zwischen Risiko und Resilienz* (S. 57-78). München: Ernst Reinhardt.
- Lösel, F., Runkel, D., Beelmann, A., Jaurisch, S., Stemmler, M. (2008). Das Präventionsprogramm EFFEKT: Entwicklungsförderung in Familien: Eltern- und Kinder-Training (EFFEKT). In Bundesministerium des Innern (Hrsg.), *Theorie und Praxis des gesellschaftlichen Zusammenhalts – Aktuelle Aspekte der Präventionsdiskussion um Gewalt und Extremismus* (S. 199-219.). Berlin: BMI.
- Lohbeck, A., Schultheiß, J., Petermann, F. & Petermann, U. (2015). Die deutsche Selbstbeurteilungsversion des Strengths and Difficulties Questionnaire (SDQ-Deu-S). Psychometrische Eigenschaften, Faktorenstruktur und Grenzwerte. *Diagnostica*, 61, 222-235.
- Lotz, M., Gabriel, K. & Lipowsky, F. (2013). Niedrig und hoch inferente Verfahren der Unterrichtsbeobachtung. Analysen zu deren gegenseitiger Validierung. *Zeitschrift für Pädagogik*, 59(3), 357-380.
- Lübke, L., Meyer, J.; Christiansen, H. (2016). Effekte von Einstellungen und subjektiven Erwartungen von Lehrkräften: Die Theorie des geplanten Verhaltens im Rahmen schulischer Inklusion. *Empirische Sonderpädagogik*, 8(3), 225-238.
- MacFarlane, K. & Woolfson, L. M. (2013). Teacher attitudes and behavior toward the inclusion of children with social, emotional and behavioral difficulties in mainstream schools: An application of the theory of planned behavior. *Teaching & Teacher Education*, 29, 46-52.
- Maggin, D. M., Cook, B. G. & Cook, L. (2018). Using Single-Case Research Designs to Examine the Effects of Interventions in Special Education. *Learning Disabilities Research & Practice*, 33(4), 182-191. <https://doi.org/10.1111/ldrp.12184>
- Mahlau, K. (2019). Diagnostik im Förderschwerpunkt Emotionale und soziale Entwicklung – Überlegungen zum praktischen Einsatz von Direct Behavior Ratings im inklusiven Unterricht. *Zeitschrift für Heilpädagogik*, 70(11), 581-591.
- Mahlau, K. & Blumenthal, Y. (2014). Unterrichtliche Förderung von Kindern mit eingeschränktem Sprachverständnis. Eine vergleichende Studie zwischen Kindern des Rügener Inklusionsmodells (RIM) und des herkömmlichen Unterrichts. *Logos*, 84-95.
- Mahlau, K., Blumenthal, Y. & Hartke, B. (2016). Prävention und Inklusion in den Förderschwerpunkten emotional-soziale Entwicklung, Lernen und Sprache im Rügener Inklusionsmodell (RIM). *Zeitschrift für Heilpädagogik*, 67(5), 104-118.
- Mahlau, K. & Erdmann, C. (2020). Gelingensbedingungen für die Implementation von Verhaltensverlaufsmeasurements – Eine quasi-experimentelle Feldstudie mit Schülern mit aggressivem Verhalten und ihren Lehrkräften. *Zeitschrift für Heilpädagogik*, 71(3), 113-124.

- Manolov, R., Jamieson, M., Evans, J. J., & Sierra, V. (2016). A discussion of alternatives for establishing empirical benchmarks for interpreting single-case effect sizes. *Psicológica*, 37, 209-234.
- Manolov, R., & Moeyaert, M. (2017). Recommendations for choosing single-case data analytical techniques. *Behavior Therapy*, 48(1), 97-114. <https://doi.org/10.1016/j.beth.2016.04.008>.
- Masten, A. S. & Powell, J. (2003). A resilience framework for research, policy and practice. In S. Luthar (Hrsg.), *Resilience and vulnerability. Adaption in the context of childhood adversities* (S. 1-25). Cambridge: Cambridge University Press.
- Matta, M., Volpe, R. J., Briesch, A. M. & Owens, J. S. (2020). Five direct behavior rating multi-item scales: Sensitivity to the effects of classroom interventions. *Journal of School Psychology*, 81, 28-46.
- Mayer, H., Heim, P. & Scheithauer, H. (2007). *Papilio. Ein Programm für Kindergärten zur Primärprävention von Verhaltensproblemen und zur Förderung sozial-emotionaler Kompetenz. Ein Beitrag zur Sucht- und Gewaltprävention. Theorie und Grundlagen* (2., aktualis. & erw. Aufl.). Augsburg: beta Institutsverlag.
- McDowall, S., Cameron, M., Dingle, R., Gilmore, A. & MacGibbon, L. (2007). *Evaluation of the Literacy Professional Development Project*. Ministry of Education, New Zealand.
- MacFarlane, K., & Woolfson, L. M. (2013). Teacher attitudes and behavior toward the inclusion of children with social, emotional and behavioral difficulties in mainstream schools: An application of the theory of planned behavior. *Teaching and Teacher Education*, (29), 46-52. <https://doi.org/10.1016/j.tate.2012.08.006>
- Melzer, C. & Hillenbrand, C. (2015). Aufgabenprofile. Welche Aufgaben bewältigen sonderpädagogische Lehrkräfte in verschiedenen schulischen Tätigkeitsfeldern? *Zeitschrift für Heilpädagogik*, 66(5), 230-242.
- Melzer, C., Hillenbrand, C., Sprenger, D., Hennemann, T. (2015). Aufgaben von Lehrkräften in inklusiven Bildungssystemen – Review internationaler Studien. *Erziehungswissenschaft*, 26/51, 61-80.
- Methner, A., Popp, K. & Seebach, B. (2017). *Verhaltensprobleme in der Sekundarstufe. Unterricht – Förderung – Intervention*. Stuttgart: Kohlhammer.
- Meyer, H. (2011). *Was ist guter Unterricht?* Berlin: Cornelsen-Scriptor.
- Ministerium für Bildung und Kindertagesförderung (2022a). *Einführung eines inklusiven Schulsystems in Mecklenburg-Vorpommern bis 2028*. Zugriff am 01.08.2024 auf https://www.bildung-mv.de/export/sites/bildungsserver/downloads/publikationen/10092019_Entschleunigung-im-Schulgesetz_-Zeitstrahl.pdf.
- Ministerium für Bildung und Kindertagesförderung (2022b). *Schuldienst*. Zugriff am 01.08.2024 auf <https://www.regierung-mv.de/Landesregierung/bm/Schule/Schuldienst/>.
- Möller, K., Hardy, I., Jonen, A., Kleickmann, T. & Blumberg, E. (2006). Naturwissenschaften in der Primarstufe. Zur Förderung konzeptuellen Verständnisses durch Unterricht und zur Wirksamkeit von Lehrerfortbildungen. In M. Prenzel & L. Allolio-Näcke (Hrsg.), *Untersuchungen zur Bildungsqualität von Schule* (S. 161-193). Münster: Waxmann.
- Müller, F. J. (2018). *Blick zurück nach vorn – WegbereiterInnen der Inklusion*. Band 2. Gießen: Psychosozial-Verlag.
- Mutzeck, W. (2008). *Kooperative Beratung. Grundlagen, Methoden, Training, Effektivität*. 6., überarbeitete und erweiterte Auflage. Weinheim: Beltz.

- Myschker, N. (2009). *Verhaltensstörungen bei Kindern und Jugendlichen. Erscheinungsformen – Ursachen – Hilfreiche Maßnahmen*. Stuttgart: Kohlhammer.
- Myschker, N. & Stein, R. (2014). *Verhaltensstörungen bei Kindern und Jugendlichen. Erscheinungsformen – Ursachen – Hilfreiche Maßnahmen*. Stuttgart: Kohlhammer.
- Naumann, S., Kuschel, A., Bertram, H., Heinrichs, N. & Hahlweg, K. (2007). Förderung der Elternkompetenz durch Triple P-Elterntrainings. *Praxis der Kinderpsychologie und Kinderpsychiatrie*, 56(8), 676-690.
- Nir, A. & Bogler, R. (2008). The antecedents of teacher satisfaction with professional development programs. *Teaching and Teacher Education*, 24, 377-386.
- Novins, D. K., Green, A. E., Legha, R. K. & Aarons, G. A. (2013). Dissemination and Implementation of Evidence-Based Practices for Child and Adolescent Mental Health: A Systematic Review. *Journal of the American Academy of Child & Adolescent Psychiatry*, 52(10), 1009-1025.
- Opfer, D. & Pedder, D. (2011). Conceptualizing teacher professional learning. *Review of Educational Research*, 81(3), 376-407.
- Opp, G. (2003). Symptomatik, Ätiologie und Diagnostik bei Gefühls- und Verhaltensstörungen. In A. Leonardt & F. Wember (Hrsg.), *Bildung, Erziehung, Behinderung. Grundlagen und Methoden der pädagogischen Rehabilitation* (S. 504-517). Weinheim: Beltz.
- Osher, D., Kendziora, K., Spier, E. & Garibaldi, M. L. (2014). School Influences on Child and Youth Development. In Z. Sloboda & H. Petras (Hrsg.), *Defining prevention science* (S. 151-169). Heidelberg: Springer.
- Otto, B. (2007). *SELVES – Schüler-, Eltern- und Lehrertrainings zur Vermittlung effektiver Selbstregulation*. Berlin: Logos.
- Otto, C., Reiss, F., Voss, C., Wüstner, A., Meyrose, A., Hölling, H. & Ravens-Sieberger, U. (2021). Mental health and well-being from childhood to adulthood: design, methods and results of the 11-year follow-up of the BELLA study. *European Child & Adolescent Psychiatry*, 30, 1559-1577.
- Parker, R. I., Hagan-Burke, S. & Vannest, K. J. (2007) Percent of all nonoverlapping data PAND: An alternative to PND. *Journal of Special Education*, 40, 194-204.
- Parker, R. I. & Vannest, K. J. (2009). An improved effect size for single-case research: Nonoverlap of all pairs. *Behavior Therapy*, 40(4), 357-367. <https://doi.org/10.1016/j.beth.2008.10.006>.
- Parker, R. I., Vannest, K. J. & Davis, J. L. (2011). Effect size in single-case research: A review of nine nonoverlap techniques. *Behavior Modification*, 35, 303-322.
- Parr, J., Timperley, H., Reddish, P., Jesson, R. & Adams, R. (2007). *Literacy professional development project: Identifying effective teaching and professional development practices for enhanced student learning*. Report to the Ministry of Education. Wellington: Ministry of Education.
- Petermann, F. & Barnow, S. (2013). Emotionsregulation. *Psychologische Rundschau*, 64(4), 193-195.
- Petermann, F., Koglin, U., Natzke, H. & Marées, N. von (2013). *Verhaltenstraining in der Grundschule: Ein Programm zur Förderung emotionaler und sozialer Kompetenzen* (2., überarb. Aufl.). Göttingen: Hogrefe.
- Petermann, F. & Wiedebusch, S. (2008). *Emotionale Kompetenz bei Kindern*. Göttingen: Hogrefe.

- Piwowar, V., Thiel, F. & Ophardt, D. (2013). Training inservice teachers' competencies in classroom management. A quasi-experimental study with teachers of secondary schools. *Teaching and Teacher Education*, 30, 1-12.
- Pfeifer, W. (1993). *Etymologisches Wörterbuch des Deutschen, digitalisierte und von Wolfgang Pfeifer überarbeitete Version im Digitalen Wörterbuch der deutschen Sprache*. Zugriff am 20.09.2022 auf <https://www.dwds.de/d/wb-etymwb>.
- Pons, F., Harris, P. L. & de Rosnay, M. (2004). Emotions comprehension between 3 and 11 years: Development periods and hierarchical organization. *European Journal of Development Psychology*, 1, 127-152.
- Prüfungs- und Studienordnung für den Studiengang Grundschullehramt an der Universität Greifswald vom 25. Mai 2020. Zugriff am 01.08.2024 auf https://www.uni-greifswald.de/storages/uni-greifswald/2_Studium/2.4_Rund_um_die_Pruefungen/2.4.1_Pruefungs_und_Studienordnungen/Lehramt_modularisiert/Grundschullehramt/PSO_Grundschullehramt_2020_Lesefassung-2023.pdf.
- Qi, J. & Ha, A. S. (2012). Inclusion in Physical Education: A review of literature. *International Journal of Disability, Development and Education*, 59(3), 257-281.
- Reicher, H. & Jauk, M. (2012). Programme zur Förderung sozialer Kompetenz im schulischen Setting. In M. Fingerle & M. Grumm (Hrsg.), *Prävention von Verhaltensauffälligkeiten bei Kindern und Jugendlichen. Programme auf dem Prüfstand* (S. 29-48). München: Ernst-Reinhardt.
- Rheinberg, F. & Krug, S. (1999). *Motivationsförderung im Schulalltag. Psychologische Grundlagen und praktische Durchführung*. Göttingen: Hogrefe.
- Richardson, V. (1990). Significant and worthwhile change in teaching practice. *Educational Researcher*, 19(7), 10-18.
- Richter, D. (2016). Lehrerinnen und Lehrer lernen: Fort- und Weiterbildung im Lehrerberuf. In M. Rothland (Hrsg.), *Beruf Lehrer/Lehrerin. Ein Studienbuch* (S. 245-260). Münster: Waxmann.
- Riesch, M., Eiperle, J. & Gloe, M. (2022). *Inklusionsdidaktische Lehrbausteine (!DL)*. Zugriff am 28.11.2022 auf <https://www.idl.lehrerbildung-at-lmu.mzl.uni-muenchen.de/index.html>.
- Riley-Tillman, T. C., Chafouleas, S. M., Christ, T., Briesch, A. M. & LeBel, T. J. (2009). The impact of item wording and behavioral specificity on the accuracy of direct behavior ratings (DBRs). *School Psychology Quarterly*, 24, 1-12.
- Riley-Tillman, T. C., Chafouleas, S. M., Sassu, K. A., Chanese, J. A. M. & Glazer, A. D. (2008). Examining the Agreement of Direct Behavior Ratings and Systematic Direct Observation Data for On-Task and Disruptive Behavior. *Journal of Positive Behavior Interventions*, 10(2), 136-143.
- Riley-Tillman, T. C., Christ, T. J., Chafouleas, S. M., Boice-Mallach, C. H. & Briesch, A. (2011). The impact of observation duration on the accuracy of data obtained from direct behavior rating (DBR). *Journal of Positive Behavior Interventions*, 13(2), 119-128.
- Rindskopf, D. M., & Ferron, J. M. (2014). Using multilevel models to analyze single-case design data. In D. M. Rindskopf, J. M. Ferron, T. R. Kratochwill & J. R. Levin (Hrsg.), *Single-case intervention research: Methodological and statistical advances*. American Psychological Association.
- Röhrig, B., du Prel, J.-B., Wachtlin, D. & Blettner, M. (2009). Studientypen in der medizinischen Forschung. *Deutsches Ärzteblatt*, 106(15), 262-268.

- Rzejak, D. & Lipowsky, F. (2019). Feedback durch Coaching – eine zentrale Komponente wirksamer Lehrkräftefortbildungen. In M.-C. Vierbuchen & F. Bartels (Hrsg.), *Feedback in der Unterrichtspraxis. Schülerinnen und Schüler beim Lernen wirksam unterstützen* (S. 161-177). Stuttgart: Kohlhammer Verlag.
- Rzejak, D., Lipowsky, F. & Bleck, V. (2020). Synopse bedeutsamer Merkmale von Lehrkräftefortbildungen. Wirkungsebenen und konzeptionelle Aspekte. *Journal für LehrerInnenbildung*, 20(4), S. 18-30.
- Robert Bosch Stiftung (2021). *Das Deutsche Schulbarometer Spezial: Zweite Folgebefragung. Ergebnisse einer Befragung von Lehrerinnen und Lehrern an allgemeinbildenden Schulen im Auftrag der Robert Bosch Stiftung in Kooperation mit der ZEIT. Durchgeführt von forsa Politik- und Sozialforschung GmbH*. Stuttgart: Robert Bosch Stiftung.
- Robert Bosch Stiftung (2022). *Das Deutsche Schulbarometer. Eine repräsentative Befragung von Lehrkräften zu den aktuellen Herausforderungen der Schulen. Durchgeführt von Forsa*. Stuttgart: Robert Bosch Stiftung.
- Roberts, J. & Simpson, K. (2016). A review of research into stakeholder perspectives on inclusion of students with autism in mainstream schools. *International Journal of Inclusive Education*, 20(10), 1084-1096. <https://doi.org/10.1080/13603116.2016.1145267>.
- Ruberg, C. & Porsch, R. (2017). Einstellungen von Lehramtsstudierenden und Lehrkräften zur schulischen Inklusion. Ein systematisches Review deutschsprachiger Forschungsarbeiten. *Zeitschrift für Pädagogik*, 63(4), 393-415.
- Saarni, C. (1999). *The development of emotional competence*. New York: Guilford.
- Saile, H. (2007). Psychometrische Befunde zur Lehrerversion des “Strengths and Difficulties Questionnaire“ (SDQ-L). *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 39, 25-32.
- Salas, E. & Cannon-Bowers, J. A. (2001). The science of training: A decade of progress. *Annual Review of Psychology*, 52, 471-499.
- Salmanian, M., Mohammadi, M. R., Hooshyari, Z., Mostafavi, S. A., Zarafshan, H., Khaleghi, A., Ahmadi, A., Alavi, S. S., Shakiba, A., Rahgozar, M., Safavi, P., Arman, S., Delisheh, A., Mohammadzadeh, S., Hosseini, S. H., Ostovar, R., Hoojjat, S. K., Armani, A., Talepasand, S. & Amiri, S. (2019). Prevalence, comorbidities, and sociodemographic predictors of conduct disorder: the national epidemiology of Iranian children and adolescents' psychiatric disorders (IRCAP). *European child & adolescent psychiatry*, 1-15.
- Sanders, M. R. (1999). The Triple P-Positive Parenting Program: Towards an empirically validated multi-level parenting and family support strategy for the prevention and treatment of child behavior and emotional problems. *Child and Family Psychology Review*, (2), 71-90.
- Saxe, G., Gearhart, M. & Nasir, N. S. (2001). Enhancing students' understanding of mathematics: A study of three contrasting approaches to professional support. *Journal of Mathematics Teacher Education*, 4(1), 55-79.
- Scheer, D. (2021). *Toolbox Diagnostik: Hilfen für die (sonder-)pädagogische Praxis*. Stuttgart: Verlag W. Kohlhammer.
- Scheffelt, E. & Teichmann, A. (2018). Evaluieren Schritt für Schritt – Prozesse und Instrumente am Praxisbeispiel. Beiträge zur Weiterbildungsdiskussion. *Weiter gelernt*, (11), 4-30.

- Schellenbach-Zell, J. (2009). *Motivation und Volition von Lehrkräften in Schulinnovationsprojekten*. Elektronische Dissertation, Bergische Universität Wuppertal.
- Schellenbach-Zell, J. & Gräsel, C. (2007). *Warum beteiligen sich Lehrkräfte an Innovationsprojekten?* Lüneburg: Poster auf der 70. Tagung der Arbeitsgruppe Empirische Pädagogische Forschung (AEPF), Leuphana Universität.
- Schlee, J. & Mutzeck, W. (1996). *Kollegiale Supervision: Modelle zur Selbsthilfe für Lehrerinnen und Lehrer*. Heidelberg: Winter.
- Schlippe, A. von & Schweitzer, J. (2012). *Lehrbuch der systemischen Therapie und Beratung I. Das Grundlagenwissen*. Göttingen: Vandenhoeck & Ruprecht.
- Schreier, M. (2014). Varianten qualitativer Inhaltsanalyse: Ein Wegweiser im Dickicht der Begrifflichkeiten. *Forum Qualitative Sozialforschung*, 15(1), Art. 18. Zugriff am 01.08.2024 auf <https://www.ssoar.info/ssoar/handle/document/37708>.
- Schwarzer, R., & Schmitz, G.S. (1999). Skala Lehrer-Selbstwirksamkeitserwartung. In R. Schwarzer & M. Jerusalem (Hrsg.), *Skalen zur Erfassung von Lehrer- und Schülermerkmalen* (S. 60-61). Berlin.
- Schwarzer, R. & Schmitz, G.S. (2002). WirkLehr. Skala Lehrer-Selbstwirksamkeit [Verfahrensdokumentation, Autorenbeschreibung und Fragebogen]. In Leibniz-Institut für Psychologie (ZPID) (Hrsg.), *Open Test Archive*. Trier: ZPID.
- Seitz-Stein, K., Schumann-Hengsteler, R., Zoelch, C., Grube, D., Mähler, C. & Hasselhorn, M. (2012). Diagnostik der Funktionstüchtigkeit des Arbeitsgedächtnisses bei Kindern zwischen 5 und 12 Jahren: die Arbeitsgedächtnistestbatterie AGTB 5-12. In M. Hasselhorn & C. Zoelch (Hrsg.), *Funktionsdiagnostik des Arbeitsgedächtnisses* (S. 1-22). Göttingen: Hogrefe.
- Sermier Dessemontet, R., Benoit, V. & Bless, G. (2011). Schulische Integration von Kindern mit einer geistigen Behinderung. Untersuchung der Entwicklung der Schulleistungen und der adaptiven Fähigkeiten, der Wirkung auf die Lernentwicklung der Mitschüler sowie der Lehrereinstellungen zur Integration. *Empirische Sonderpädagogik*, 3(4), 291-307.
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Houghton, Mifflin and Company.
- Siegler, R., de Loache, J. & Eisenberg, N. (2008). *Entwicklungspsychologie im Kindes- und Jugendalter*. Heidelberg: Springer.
- Silver, R. B., Measelle, J. R., Armstrong, J. M. & Essex, M. J. (2010). The impact of parents, child care providers, teachers, and peers on early externalizing trajectories. *Journal of School Psychology*, 48, 555-583. <https://doi.org/10.1016/j.jsp.2010.08.003>.
- Shayer, M. & Adhami, M. (2007). Fostering cognitive development through the context of Mathematics: Results of the CAME project. *Educational studies in Mathematics*, 64(3), 265-291.
- Shulman, L. S. (1987). Knowledge and teaching: Foundations of the New Reform. *Harvard Educational Review*, 57(1), 1-23.
- Shulman, L. S. (1998). Theory, practice, and the education of professionals. *The Elementary School Journal*, 98(5), 511-526.
- Shulman, L. S. (2004). *The wisdom of practice. Essays on teaching, learning, and learning to teach*. San Francisco: Jossey-Bass.

- Sklad, M., Diekstra, R., Ritter, M. d. & Ben, J. (2012). Effectiveness of school-based universal social, emotional, and behavioral programs: do they enhance students' development in the area of skill, behavior, and adjustment? *Psychology in the Schools*, 49(9), 892-909.
- Smith, C. & Gillespie, M. (2007). Research on professional development and teacher change: Implications for Adult Basic Education. *Review of Adult Learning and Literacy*, 7, 205-244.
- Smith, R. L., Eklund, K. & Kilgus, S. P. (2018). Concurrent validity and sensitivity to change of Direct Behavior Rating Single-Item Scales (DBR-SIS) within an elementary sample. *School Psychology Quarterly*, 33(1), 83-93.
- Speck, O. (1990). *Verhaltensstörungen, Psychopathologie und Erziehung – Grundlagen zu einer Verhaltensgestörtenpädagogik*. Berlin: Marhold.
- Stamann, C. (2015). Personalentwicklung als Innovation – Schulleitung als ‚change agent‘ im sozialen System Schule. In K. Kansteiner & C. Stamann (Hrsg.), *Personalentwicklung in der Schule zwischen Fremdsteuerung und Selbstbestimmung*. Bad Heilbrunn: Verlag Julius Klinkhardt.
- Statistisches Bundesamt (2019). *Statistisches Jahrbuch Deutschland und Internationales 2019*. Zugriff am 01.08.2024 auf <https://www.destatis.de/DE/Themen/Querschnitt/Jahrbuch/statistisches-jahrbuch-2019-dl.html>.
- Stein, R. & Müller, T. (2015). Verhaltensstörungen und emotional-soziale Entwicklung: zum Gegenstand. In R. Stein & T. Müller (Hrsg.), *Inklusion im Förderschwerpunkt emotionale und soziale Entwicklung* (S. 19-44). Stuttgart: Kohlhammer.
- Stipek, D. J., Givvin, K. B., Salmon, J. M. & MacGyvers, V. L. (2001). Teachers' beliefs and practices related to mathematics instruction. *Teaching and Teacher Education*, 17, 213-226.
- Studiengangsspezifische Prüfungs- und Studienordnung für den Masterstudiengang Sonder- und Inklusionspädagogik der Universität Rostock vom 3. Juni 2016. Zugriff am 01.08.2024 auf https://www.uni-rostock.de/storages/uni-rostock/UniHome/Gremien/Rechtsgrundlagen/Amtliche_Bekanntmachungen/2017/NR_6_2017.pdf.
- Studiengangsspezifische Prüfungs- und Studienordnung für den Studiengang Lehramt an Grundschulen der Universität Rostock vom 8. März 2021. Zugriff am 01.08.2024 auf https://www.uni-rostock.de/storages/uni-rostock/UniHome/Gremien/Rechtsgrundlagen/Amtliche_Bekanntmachungen/2021/NR_26_2021.pdf.
- Studiengangsspezifische Prüfungs- und Studienordnung für den Studiengang Lehramt für Sonderpädagogik inkl. Fachanhänge vom 9. Oktober 2012 (SPSO Sonder 2012). Zugriff am 01.08.2024 auf https://www.uni-rostock.de/storages/uni-rostock/UniHome/Gremien/Rechtsgrundlagen/Amtliche_Bekanntmachungen/2013/NR_9_2013.pdf.
- Tate, R. L., Perdices, M., Rosenkoetter, U., McDonald, S., Togher, L., Shadish, W., Kratochwill, T., Kazdin, A., Sgamseer, L. & Vohra, S. (2016). The Single-Case Reporting Guideline In BEhavioural Interventions (SCRIBE) 2016: Explanation and elaboration. *Archives of Scientific Psychology*, 4(1), 10-31.
- Taylor, B. M., Pearson, P. D., Peterson, D. S. & Rodriguez, M. C. (2005). The CIERA School Change Framework: An evidence-based approach to professional development and school reading improvement. *Reading Research Quarterly*, 40(1), 40-69.

- Tenorth, H.-E. (2006). Professionalität im Lehrerberuf. Ratlosigkeit der Theorie, gelingende Praxis. *Zeitschrift für Erziehungswissenschaft*, 9(4), 580-597.
- Terhart, E. (2011). Lehrerberuf und Professionalität. Gewandeltes Begriffsverständnis – neue Herausforderungen. In W. Helsper & R. Tippelt (Hrsg.), *Pädagogische Professionalität*, Beiheft, 57 (S. 202-224). Weinheim: Beltz.
- Tillmann, K.-J. (2011). Konzepte der Forschung zum Lehrerberuf. In E. Terhart, H. Bennewitz & M. Rothland (Hrsg.), *Handbuch der Forschung zum Lehrerberuf* (S. 232-240). Münster: Waxmann Verlag.
- Timperley, H. (2011). *Realizing the power of professional learning*. Berkshire: Open University Press.
- Timperley, H., Wilson, A., Barrar, H. & Fung, I. (2007). *Teacher professional learning and development: Best evidence synthesis iteration (BES)*. Wellington: Ministry of Education. Zugriff am 01.08.2024 auf https://thehub.sia.govt.nz/assets/documents/42432_TPLandDBESentireWeb_0.pdf.
- Toepell, S. (2010). *Lehrertrainings im deutschen Sprachraum. Einzeldarstellungen und Entwicklungsverläufe*. München: Herbert Utz Verlag.
- United Nations (2006). *Convention on the Rights of Persons with Disabilities*. Zugriff am 02.01.2023 auf <https://www.un.org/development/desa/disabilities/convention-on-the-rights-of-persons-with-disabilities/article-24-education.html>.
- Urban, M. (2022). Multiprofessionelle Kooperation, sonderpädagogische Beratungs- und Unterstützungssysteme und Professionelle Lerngemeinschaften als Beispiele der inklusionsinduzierten Ausdifferenzierung reflexionsermöglichender Strukturen im Schulsystem. In A. Schröter, M. Kortmann, S. Schulze, K. Kempfer, S. Anderson, G. Sevdiren, J. Bartz & C. Kreutchen (Hrsg.), *Inklusion in der Lehramtsausbildung – Lerngegenstände, Interaktionen und Prozesse* (S. 33-42). Münster: Waxmann.
- Urbutt, A. (2001). *Belastungen im Lehrerberuf und Aspekte der Bewältigung*. Lüneburg: GRIN.
- Urtou, K., Wilbert, J., & Hennemann, T. (2014). Der Zusammenhang zwischen der Einstellung zur Integration und der Selbstwirksamkeit von Schulleitungen und deren Kollegien. *Empirische Sonderpädagogik*, 6(1), 3-16.
- Van den Noortgate, W., & Onghena, P. (2003). Hierarchical linear models for the quantitative integration of effect sizes in single-case research. *Behavior Research Methods, Instruments, & Computers*, 35(1), 1-10. <https://doi.org/10.3758/BF03195492>.
- Van Mieghem, A., Verschueren, K., Petry, K. & Struyf, E. (2020). An analysis of research on inclusive education: a systematic search and meta review. *Journal of Inclusive Education*, 24(6), 675-689.
- Vannest, K. J. & Ninci, J. (2015). Evaluating Intervention Effects in Single-Case Research Designs. *Journal of Counseling & Development*, 93(4), 403-411.
- Veermann, J.W. & van Yperen, T.A. (2007). Degrees of freedom and degrees of certainty: A developmental model for the establishment of evidence-based youth care. *Evaluation and Program Planning*, 30, 212-221.
- Vehmeier, J., Kleickmann, T. & Möller, K. (2007). Zusammenhänge von Vorstellungen zum Lehren und Lernen mit unterrichtlichen Handlungen von Lehrkräften. In K. Möller, P. Hanke, C. Beinhreich, A. K. Hein, T. Kleickmann, R. Schages, *Qualität von Grundschulunterricht entwickeln, erfassen und bewerten* (S. 317-320). Bonn: VS Verlag für Sozialwissenschaften.

- Verband Bildung und Erziehung (2022, 25. Januar). *Politik verschleiert Realität: Bis 2030 fehlen mindestens 81.000 Lehrkräfte!* Zugriff am 03.12.2022 auf <https://www.vbe.de/presse/presseedienste/presseedienste-2022/standard-titel>.
- Visser, L., Büttner, G. & Hasselhorn, M. (2019). Kormobidität spezifischer Lernstörungen und psychischer Auffälligkeiten. Ein Literaturüberblick. *Lernen und Lernstörungen*, 8(1), S. 7-20.
- Volpe, R. J., & Briesch, A. M. (2012). Generalizability and dependability of single-item and multiple-item direct behavior rating scales for engagement and disruptive behavior. *School Psychology Review*, 41(3), 246-261.
- Volpe, R. J. & Briesch, A. M. (2015). Multi-item direct behavior ratings: Dependability of two levels of assessment specificity. *School Psychology Quarterly*, 30(3), 431-442.
- Vonderlin, E. & Pauen, S. (2013). Entwicklungsrisiken und Entwicklungsabweichungen. In F. Petermann (Hrsg.), *Lehrbuch der Klinischen Kinderpsychologie* (S. 77-100). Göttingen: Hogrefe.
- Voß, S., Blumenthal, Y., Ehrich, K. & Mahlau, K. (2020). Multimodale Diagnostik als Ausgangspunkt für spezifische Förderung – Eine Darlegung am Beispiel der Rechtschreibung. *Zeitschrift für Heilpädagogik*, 71(3), 88-100.
- Voß, S., Blumenthal, Y., Mahlau, K., Marten, K., Diehl, K., Sikora, S. & Hartke, B. (2016). *Der Response-to-Intervention-Ansatz in der Praxis. Evaluationsergebnisse zum Rügener Inklusionsmodell*. Münster: Waxmann.
- Wackermann, R. (2008). *Überprüfung der Wirksamkeit eines Basismodell-Trainings für Physiklehrer*. Berlin: Logos.
- Weidemann, G. (2019). Arbeitsplatzbedingungen für gute Inklusion. Was brauchen Schulen für Inklusion? *nds. Die Zeitschrift der Bildungsgewerkschaft*, (5), 28-29.
- Wilbert, J., & Lüke, T. (2021). *Scan: Single-case data analyses for single and multiple baseline designs (Version 0.53) [R]*. Zugriff am 01.08.2024 auf <https://cran.r-project.org/web/packages/scan/scan.pdf>.
- Wilbert, J., Bosch, J. & Lüke, T. (2021). Validity and Judgment Bias in Visual Analysis of Single-Case Data. *International Journal for Research in Learning Disabilities*, 5(1), 13-24. <https://doi.org/10.28987/ijrld.5.1.13>.
- Willmann, M. (2015). Zur geschichtlichen Entwicklung der schulischen Erziehungshilfe. In R. Stein & T. Müller (Hrsg.), *Inklusion im Förderschwerpunkt emotionale und soziale Entwicklung* (S. 44-76). Stuttgart: Kohlhammer.
- Wilson, D. B., Gottfredson, D. C. & Najaka, S. S. (2001). School-Based Prevention of Problem Behaviors: A Meta-Analysis. *Journal of Quantitative Criminology* volume, 17, 247-272.
- Wilson, S. J. & Lipsey, M. W. (2007). School-Based Interventions for Aggressive and Disruptive Behavior. *American Journal of Preventive Medicine*, 33, 130-143.
- Wirtz, M. A. (2021). *Dorsch. Lexikon der Psychologie* (20. Aufl.). Zugriff am 01.08.2024 auf <https://dorsch.hogrefe.com/stichwort/professionalisierung-von-lehrkraeften#search=59575dd6f2c904ff7e4171ff142410e&offset=0>.
- Wuntke, L. V., Blumenthal, Y., Föllmer, J. & Mahlau, K. (2022). Multisystemisches Arbeiten in der Schule anhand des Förderkonzepts Familienklassenzimmer. In D. Raufelder, O. Steinberg, C. Retzlaff-Fürst (Hrsg.), *Qualitätsoffensive Lehrer:innenbildung: Einflussfaktoren und Gelingensbedingungen. Ergebnisse aus Mecklenburg-Vorpommern* (S. 65-74). Opladen: Verlag Barbara Budrich.

Yoon, K. S., Duncan, T., Lee, S. W. Y., Scarloss, B. & Shapley, K. (2007). *Reviewing the evidence on how teacher professional development affects student achievement*. Issues and Answers Report, REL2007-No.033. Washington, DC: US Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Southwest.

Lena Varuna Wuntke Schulische Förderung der emotionalen und sozialen Entwicklung von Kindern durch Erweiterung spezifischer Lehrkraftkompetenzen

Zu den größten beruflichen Herausforderungen von Lehrkräften zählt die Förderung von Kindern mit Verhaltensschwierigkeiten. Der deutschsprachige Raum bietet dazu allerdings bislang kaum bewährte Professionalisierungsformate. Im vorliegenden Buch entwickelt und evaluiert die Autorin ein prozessbegleitendes Fortbildungskonzept. Die Ergebnisse geben einen konzeptionellen Rahmen für die Entwicklung zukünftiger Professionalisierungsmaßnahmen und stellen einen wichtigen Beitrag auf dem Weg zu einem inklusiven Bildungssystem dar.

Die Autorin:

Lena Varuna Wuntke, Wissenschaftliche Mitarbeiterin am Lehrstuhl für Sonderpädagogik und Inklusion (Akademische Rätin), Institut für Erziehungswissenschaften, Universität Greifswald

ISBN 978-3-96665-091-5



9 783966 650915

www.budrich-academic-press.de