

Mehrstufige Förderung im Kontext externalisierenden Verhaltens in der Grundschule

Inauguraldissertation zur Erlangung des Doktorgrades der Humanwissenschaftlichen Fakultät
der Universität zu Köln nach der Promotionsordnung vom 18.12.2018

Begutachtung

Diese Dissertation wurde von der Humanwissenschaftlichen Fakultät der Universität zu Köln im Juli 2024 angenommen.

Erstgutachten: Prof. Dr. Thomas Hennemann

Zweitgutachten: Prof.in Dr. Charlotte Hanisch

Vorwort

Diese Arbeit konnte entstehen, weil viele verschiedene Menschen, die sich mitunter gar nicht kennen, ohne es zu wissen ähnliche Werte, Vorstellungen und Ideen vertreten haben und somit an unterschiedlichen Schraubchen im System an der Entwicklung, Ausführung und Distribution dieser Arbeit beteiligt waren. Das Schwierige daran war, die vielen Gedanken und Ideen zusammenzubringen, sie zu strukturieren und sie in wissenschaftliche Studienformate und interpretierbare Ergebnisse zu übersetzen, sodass daraus etwas Sinnvolles entstehen konnte. Das, was nämlich alle Beteiligten eint, sind die ähnlichen gesellschaftspolitischen Werte, Ansichten und Überzeugungen. Das, was sie unterscheidet, ist das, was sie für mich sind: Freund*innen, Familienmitglieder, Kolleg*innen, Studienteilnehmer*innen, Chef*innen oder politische Entscheidungsträger*innen. Auch wenn hier nun mein Name vorne draufsteht, habe ich das dementsprechend nicht allein gemacht und möchte dafür einmal Danke sagen an:

- Tatjana: Ohne dich wäre ich wahrscheinlich nicht mal am Lehrstuhl gelandet. Danke dir für das Vertrauen und die Zusammenarbeit von Sekunde 1 an.
- Thomas & Charlotte: Danke für das Vertrauen, die Betreuung und die Möglichkeit, diese Arbeit überhaupt zu schreiben; die Stellen, die das möglich gemacht haben und die vielen Gedanken, die in dieser Arbeit stecken (plus Espresso).
- Das Multimo-Team und allen Co-Autor*innen: Ihr habt sehr großen Anteil an dieser Arbeit. Danke für die dauerhafte gegenseitige Unterstützung!
- Das gesamte ESE-Team: Danke fürs Anhören und Diskutieren meiner Beiträge und Ideen. Vieles davon findet ihr hier wieder, was mich sehr freut!
- Julian, Johanna, Jule, Leonie, Tobi, Meike, Silena, Hannah: Danke euch explizit für kollegiales Miteinander, Korrekturlesen, Kaffee-Trinken, Mittagessen, Mitdenken und all das, was viel mehr Freundschaft ist!
- All meinen Freund*innen, die als Ausgleich ungemein wichtig sind, insbesondere: Becca, Dora, Ezy, Hanna, Jannik, Johanna, Jonas, Joshi, Julius, Martin, Max, Paul, Tobi.
- Meiner Familie, insbesondere meiner Mutter fürs bedingungslose Da-Sein und unterstützen in jeglicher Hinsicht!
- Josi: Danke dir. Deine Perspektiven bereichern mein Leben und diese Arbeit enorm. Ich bin sehr froh, dich und deine Familie in meinem Leben zu haben. Ohne dich wären Teile dieser Arbeit nicht so, wie sie sind.

Inhaltsverzeichnis

VORWORT	II
INHALTSVERZEICHNIS	III
ABKÜRZUNGSVERZEICHNIS	VII
ABBILDUNGSVERZEICHNIS	IX
TABELLENVERZEICHNIS	X
BEITRÄGE DIESER RAHMENSCHRIFT	XI
1 EINLEITUNG	1
2 EXTERNALISIERENDES VERHALTEN	5
2.1 DIE PRÄVALENZ VON EXTERNALISIERENDEM VERHALTEN.....	6
2.2 DIE ÄTIOLOGIE VON EXTERNALISIERENDEM VERHALTEN	7
2.3 AUSWIRKUNGEN VON EXTERNALISIERENDEM VERHALTEN.....	8
2.4 EXTERNALISIERENDES VERHALTEN IM SCHULISCHEN KONTEXT.....	9
2.5 INTERVENTIONEN IM KONTEXT EXTERNALISIERENDEN VERHALTENS.....	10
3 MEHRSTUFIGE FÖRDERSYSTEME (MTSS)	13
3.1 POLITISCHE UND WISSENSCHAFTLICHE RAHMUNG GESTUFTER FÖRDERUNG	14
3.2 DIE PÄDAGOGISCHEN GRUNDPFILER GESTUFTER FÖRDERUNG.....	15
3.3 DIE FUNKTIONSWEISE GESTUFTER FÖRDERUNG.....	17
3.3.1 <i>Grundprinzipien und Säulen eines MTSS</i>	17
3.3.2 <i>Die Stufen eines MTSS</i>	19
3.3.3 <i>Das datengeleitete Vorgehen (decision making)</i>	20
3.3.4 <i>Multiprofessionelle Zusammenarbeit (problem-solving)</i>	22
3.4 MTSS IN DER GRUNDSCHULE IM INTERNATIONALEN VERGLEICH.....	23
3.4.1 <i>Systematische Überblicksarbeiten zum Thema MTSS</i>	24
3.4.2 <i>Response to Intervention (RTI)</i>	27
3.4.3 <i>Positive Behavior (Interventions and) Support (PB[I]S)</i>	28
3.4.4 <i>School-wide Positive Behavior (Interventions and) Support (SWPB[I]S)</i>	29
3.4.5 <i>Verschiedene MTSS-Formen und Forschungszugänge</i>	30
3.5 PÄDAGOGISCHER UNTERSTÜTZUNGSBEDARF IN DER GRUNDSCHULE IN DEUTSCHLAND.....	31
3.6 MTSS IN DER GRUNDSCHULE IN DEUTSCHLAND	33
3.6.1 <i>Das Rügener Inklusionsmodell (RIM)</i>	33
3.6.2 <i>Partizipation in einem Response-to-Intervention-Modell (PARTI)</i>	35
3.6.3 <i>Starke Schulen stärken Kinder</i>	35
3.6.4 <i>Offene Fragen & Ausblick</i>	36

3.7 EIN SYSTEMATISCHER LITERATURÜBERBLICK ÜBER MTSS MIT FOKUS AUF VERHALTENSBEZOGENE MAßNAHMEN (BEITRAG 1)	37
3.7.1 <i>Methodik</i>	38
3.7.2 <i>Ergebnisse</i>	39
3.7.3 <i>Implikationen und Ausblick</i>	42
4 DER MEHRSTUFIGE FÖRDERANSATZ MULTIMO	43
4.1 STUFE 1: DAS GOOD BEHAVIOR GAME ALS UNIVERSELLE FÖRDERMAßNAHME	44
4.2 STUFE 2: DIE DAILY BEHAVIOR REPORT CARDS ALS SELEKTIVE FÖRDERMAßNAHME	45
4.3 STUFE 3: DAS SCEP ALS INDIZIERTE FÖRDERMAßNAHME	46
4.4 EINORDNUNG & DISKUSSION DES FÖRDERANSATZES MULTIMO	47
5 WIRKSAMKEITSSTUDIEN IM KONTEXT MULTIMO.....	48
5.1 VOR- UND NACHTEILE VON EINZELFALLSTUDIEN UND IHREN DATENSÄTZEN	48
5.1.1 <i>Vorteile von Einzelfallstudien</i>	48
5.1.2 <i>Problembereiche von Einzelfallstudien</i>	49
5.1.3 <i>Auswertungsmethoden von Datensätzen von Einzelfallstudien</i>	50
5.2 EINZELFALLSTUDIEN IM KONTEXT VON SCHULISCHER INTERVENTIONSFORSCHUNG	52
5.3 WIRKSAMKEIT DES FÖRDERANSATZES MULTIMO AUF STUFE 1 & 2	55
5.3.1 <i>Methodik</i>	55
5.3.2 <i>Implementation & Durchführung</i>	56
5.3.3 <i>Auswertung & Analyse</i>	56
5.3.4 <i>Ergebnisse</i>	57
5.3.5 <i>Implikationen & Ausblick</i>	57
5.4 WIRKSAMKEIT DES FÖRDERANSATZES AUF STUFE 3 ANHAND VON EINZELFALLDATEN.....	58
5.4.1 <i>Methodisches Vorgehen der ergänzenden quantitativen Datenauswertung</i>	58
5.4.2 <i>Ergebnisse</i>	58
5.4.3 <i>Diskussion</i>	62
5.4.4 <i>Ausblick</i>	63
5.5 QUALITATIVE ANALYSEN DER STUFE 3: DAS SCEP-COACHING (BEITRAG 3)	63
5.5.1 <i>Methodik</i>	63
5.5.2 <i>Durchführung</i>	64
5.5.3 <i>Ergebnisse</i>	64
5.5.4 <i>Ausblick</i>	65
5.6 WIRKSAMKEIT DES GOOD BEHAVIOR GAMES AUF STUFE 1 (BEITRAG 4)	66
5.6.1 <i>Methodik</i>	66
5.6.2 <i>Durchführung</i>	67
5.6.3 <i>Auswertung & Analyse</i>	67
5.6.4 <i>Ergebnisse</i>	67
5.6.5 <i>Ausblick</i>	68
5.7 WIRKSAMKEIT DER DAILY BEHAVIOR REPORT CARDS (BEITRAG 5).....	69
5.7.1 <i>Methodik</i>	69

5.7.2 Durchführung.....	69
5.7.3 Auswertung & Analyse.....	70
5.7.4 Ergebnisse.....	70
5.7.5 Ausblick.....	71
6 IMPLEMENTATION GESTUFTER FÖRDERSYSTEME IN DEUTSCHLAND	72
6.1 BEFUNDE US-AMERIKANISCHER IMPLEMENTATIONSFORSCHUNG	74
6.2 HERAUSFORDERUNGEN UND GELINGENSBEDINGUNGEN DER IMPLEMENTATION VON MTSS IN DER GRUNDSCHULE (BEITRAG 6)	77
6.2.1 Methodik.....	77
6.2.2 Durchführung.....	77
6.2.3 Auswertung & Analyse.....	78
6.2.4 Ergebnisse.....	78
6.2.5 Ausblick.....	80
7 ZUSAMMENFASSENDER ERGEBNISDARSTELLUNG.....	81
7.1 EINE SYSTEMATISCHE LITERATURANALYSE ALS REFERENZ- UND AUSGANGSPUNKT	81
7.2 UNTERSUCHUNG DER STUFEN 1, 2 & 3	82
7.3 UNTERSUCHUNG IMPLEMENTATIONSBEZOGENER KONTEXTBEDINGUNGEN.....	83
7.4 EIN AUSBLICK AUF BASIS DER ERGEBNISSE	84
8 METHODENKRITIK.....	86
8.1 KRITISCHE ANALYSE DURCHGEFÜHRTER EINZELFALLSTUDIEN.....	86
8.1.1 Betrachtung der verschiedenen Studiendesigns.....	87
8.1.2 Analyse der Stichprobenqualität.....	88
8.1.3 Analyse der Datenqualität.....	88
8.1.4 Gängige Verfahren zur Auswertung von Einzelfalldaten	91
8.1.5 Analyse der verwendeten Auswertungsmethoden	93
8.1.6 Zusammenfassende Darstellung	95
8.2 QUALITATIVE FORSCHUNG IM KONTEXT MTSS	95
8.2.1 Analyse bestehender qualitativer Forschung im Kontext MTSS.....	96
8.2.2 Einordnung der durchgeführten Interviewstudien (Beitrag 3 & 6)	98
8.2.3 Zusammenfassende Darstellung	99
8.3 BEWERTUNG DER METHODISCHEN TRIANGULATION.....	100
8.4 DENKBARE STUDIENDESIGNS & METHODISCHE ZUGÄNGE IM KONTEXT MTSS	101
8.4.1 Gruppenbasierte empirische Forschungsdesigns	101
8.4.2 Skizzierung eines Regression Discontinuity Designs.....	103
8.4.3 Weitere vielversprechende methodische Zugänge.....	104
9 ABSCHLIEßENDE DISKUSSION.....	106
9.1 BEANTWORTUNG DER LEITFRAGEN 1-5.....	106
9.1.1 Beantwortung der Leitfragen 1 & 2: Der internationale Vergleich von MTSS	106

9.1.2 Beantwortung der Leitfragen 3 & 4: Effektivität von MTSS in der Grundschule.....	107
9.1.3 Beantwortung der Leitfrage 5: Implementationsbedingungen von MTSS in der Grundschule.....	108
9.2 DISKUSSION DES FÖRDERANSATZES MULTIMO.....	108
9.3 ADAPTIONSMÖGLICHKEITEN DES FÖRDERANSATZES MULTIMO	110
10 FAZIT & AUSBLICK	113
LITERATURVERZEICHNIS.....	115
ANTEILE DER EIGENLEISTUNG	154
ANHANG	156

Abkürzungsverzeichnis

ADHD:	Attention Deficit Hyperactivity Disorder (englisch für ADHS)
ADHS:	Aufmerksamkeitsdefizit-/Hyperaktivitätsstörung
APA:	American Psychological Association
ASA:	Academic and Behavior Response to Intervention School Assessment
BASC-3 BESS:	Behavioral and Emotional Screening System
CBM:	Curriculumbasiertes Messen
CIT:	Critical Incident Technique
DBR:	Daily Behavior Report (Direkte Verhaltensbeurteilung)
DBRC:	Daily Behavior Report Cards
DSM-V:	Diagnostisches und Statistisches Manual Psychischer Störungen, 5. Ausgabe
FBA:	Functional Behavior Assessment (Funktionale Verhaltensanalyse)
GBG:	Good Behavior Game
GDB:	Allgemeines Störverhalten (General Disruptive Behavior)
ICD-10:	Internationale Klassifikation der Krankheiten, 10. Revision
ITRF APD:	Integrated Teacher Rating Form Academic Productivity/Disorganization
ITRF OB:	Integrated Teacher Rating Form Oppositional/Disruptive Sub-Scale
ITRG Total:	Integrated Teacher Rating Form Total Score
KMK:	Kultusministerkonferenz
LSL:	Lehrereinschätzliste für Sozial- und Lernverhalten
LV:	Lernförderliches Verhalten
M:	Mittelwert/Mean
MBLR:	Mean Baseline Reduction
MIS:	Missings
MTSS:	Multi-tiered System(s) of Support
MTSS-B:	Multi-tiered System(s) of Support im Kontext Verhalten
N-of-1-Trial:	Einzelfallstudie mit einer Person
NA:	Fehlende Werte
NAP:	Nonoverlap of All Pairs

VIII

ODR(s)	Office Discipline Referral(s)
PARTI:	Partizipation in einem Response-to-Intervention-Modell
PB(I)S:	Positive Behavior (Interventions and) Support
PBLR:	Percentage of Baseline Level Responding
PEM:	Percentage of Exceeding the Median
PEM-T:	Percentage of Data Exceeding a Median Trend
PND:	Percentage of Non-overlapping Data
PND:	Percentage of Non-overlapping Data
PT:	Precision Teaching
PS:	Problem-Solving
QED:	Quasi-Experimentelle Designs
RCT:	Randomized Controlled Trial
RDD:	Regression Discontinuity Design
RIM:	Rügener Inklusionsmodell
RIRD:	Robust Improvement Rate Difference
RTI:	Response to Intervention
SCED:	Single-Case Experimental Design
SCEP:	Schulbasiertes Coaching bei Kindern mit expansivem Problemverhalten
SDQ:	Strength and Difficulties Questionnaire
SET:	School-Wide Evaluation Tool
SGB:	Spezifisches Zielverhalten (Specific Goal Behavior)
SSED:	Single-Subject Experimental Designs
SSRD:	Single-Subject-Research-Designs
SSV:	Störung des Sozialverhaltens
SRSS:	Student Risk Screening Scale
SSL:	Schülereinschätzliste für Sozial- und Lernverhalten
SV:	Störendes Verhalten
SWPB(I)S:	School-wide Positive Behavior (Interventions and) Support
SWPB(I)S TFI:	School-wide Positive Behavior (Interventions and) Support Tiered Fidelity Inventory
SSsK:	Starke Schulen stärken Kinder

Abbildungsverzeichnis

<u>Abbildung 1: Begrifflichkeiten im Kontext MTSS</u>	17
<u>Abbildung 2: Gestufte Förderung als Kontinuum verschiedener Interventionen</u>	19
<u>Abbildung 3: Der gestufte Förderansatz Multimo</u>	44
<u>Abbildung 4: Visuelle Darstellung für die Fälle C1 & C24</u>	60
<u>Abbildung 5: Visuelle Darstellung für die Fälle C7 & C9</u>	61
<u>Abbildung 6: Visuelle Darstellung für die Fälle C12 & C14</u>	61
<u>Abbildung 7: Visuelle Darstellung für die Fälle C17 & C20</u>	62

Tabellenverzeichnis

<u>Tabelle 1:</u> <i>Überblick über bestehende systematische Literaturarbeiten zum Thema MTSS</i>	26
<u>Tabelle 2:</u> <i>Deskriptive Daten des Subsamples (Stufe 3)</i>	59
<u>Tabelle 3:</u> <i>Effekte auf Fallebene des Subsamples (Stufe 3)</i>	59
<u>Tabelle 4:</u> <i>Variablen gelingender Implementation nach Fox et al. (2021)</i>	75
<u>Tabelle 5:</u> <i>Herausforderungen und Gelingensbedingungen zur Implementation von MTSS</i>	78
<u>Tabelle 6:</u> <i>Übersicht über vorhandene und fehlende Daten der Einzelfallstudien (Beitrag 2, 4 & 5)</i>	90
<u>Tabelle 7:</u> <i>Übersicht über die Verhaltensveränderung der Fälle der Einzelfallstudien (Beitrag 2, 4 & 5)</i>	91

Beiträge dieser Rahmenschrift

Beitrag 1

Nitz, J., Brack, F., Hertel, S., Krull, J., Stephan, H., Hennemann, T., & Hanisch, C. (2023). Multi-tiered systems of support with focus on behavioral modification in elementary schools: A systematic review. *Heliyon*, 9(6). <https://doi.org/10.1016/j.heliyon.2023.e17506>

Beitrag 2

Nitz, J., Hagen, T., Krull, J., Verbeck, L., Eiben, K., Hanisch, C., & Hennemann, T. (2023). Tiers 1 and 2 of a German MTSS: impact of a multiple baseline study on elementary school students with disruptive behavior. *Frontiers in Education*, 8. <https://doi.org/10.3389/educ.2023.1208854>

Beitrag 3

Nitz, J., Eiben, K., Hanisch, C., Hagen, T., Krull, J., Verbeck, L., Rauterkus, H., & Hennemann, T. (2024, im Druck). Mehrstufige Förderung in der inklusiven Grundschule – erste qualitative und quantitative Befunde zur Wirksamkeit und Umsetzung des Förderansatzes Multimo. In J. König, C. Hanisch, P. Hanke, T. Hennemann, K. Kaspar, M. Martens & S. Strauß (Hrsg.), *Teachers and their teaching matters – Auf die Lehrperson und ihren Unterricht kommt es an*. 10 Jahre empirische Forschung im IZEF der Universität zu Köln. Waxmann.

Beitrag 4

Hagen, T., Nitz, J., Brack, F., Hövel, D. C., & Hennemann, T. (2023). Effekte des Good Behavior Game bei Grundschüler_innen mit externalisierenden Verhaltensproblemen. *Lernen und Lernstörungen*, 1-14. <https://doi.org/10.1024/2235-0977/a000410>

Beitrag 5

Nitz, J., Volpe, R., Hagen, T., Krull, J., Hennemann, T., Hanisch, C. (voraussichtlich 2024). The Effects of the Daily Behavior Report Cards (DBRC) on the Disruptive Behavior and Specific Goal Behavior of Elementary School Children: A Multiple Baseline Design Study: *Dieser Beitrag ist eingereicht und befindet sich im Begutachtungsverfahren.*

Beitrag 6

Nitz, J., Niederelz, A., Hanisch, C., & Hennemann, T. (2024). Herausforderungen und Gelingensbedingungen erfolgreicher Implementation eines mehrstufigen Förderansatzes an Grundschulen. *Zeitschrift für Heilpädagogik*, 75(1), 4-17.

1 Einleitung

Bildungslandschaften weltweit befinden sich in einem stetigen Wandel, der neue und komplexe Herausforderungen an bestehende Schulsysteme, die Lehrkräfte und die jeweiligen nationalen Bildungspolitiken stellt (Lütje-Klose et al., 2014; Moldenhauer, 2023; Sterling & Huckle, 1996; Wilmers et al., 2021). Diese umfassende Diskussion bezüglich Veränderungsprozessen in der Schullandschaft verlangt auch in Deutschland nach einer Transformation bestehender Systeme und ihrer Komponenten (Friedrichs & Sanders, 2002; Kammerl et al., 2023; Mecheril & Füllekruss, 2023). Im Mittelpunkt dieses erforderlichen Wandels steht vor allem die Notwendigkeit einer erhöhten Flexibilität und Anpassungsfähigkeit in sämtlichen schulischen Bereichen, um den gesellschaftlichen Wandel und die zunehmende globale Vernetzung erfolgreich zu bewältigen. Dies führt dazu, dass Lehrkräfte immer häufiger mit einer Veränderung ihrer traditionellen Rollen konfrontiert sind: Ihre Aufgabenfelder weiten sich insbesondere im Rahmen inklusiver Bildungsangebote kontinuierlich aus (Götz et al., 2015; Lütje-Klose et al., 2014; Wohnhas et al., 2024). Dabei werden sie in komplexe, diagnostische Prozesse eingebunden und stehen vor der Aufgabe, Schüler*innen zu unterrichten, die durch psychische Probleme, Behinderungen, benachteiligte sozioökonomische Verhältnisse oder Flucht- und Migrationserfahrungen geprägt sind (Hurrelmann et al., 2010; Vierbuchen & Bartels, 2020; Wirth et al., 2023). Obwohl die Ausrichtung hinsichtlich inklusiver Prozesse politisch in Deutschland, unter anderem durch die UN-Behindertenrechtskonvention ¹, vorgegeben ist, stoßen die aktuell implementierten Systeme und Schulformen häufig an ihre Grenzen, wenn es darum geht, den spezifischen Herausforderungen psychisch und sozial belasteter Kinder gerecht zu werden (Baumgardt, 2016; Langer & Schnebel, 2020).

Aktuelle Studien für die Grundschule zeigen darüber hinaus, dass in der Praxis die Heterogenität der Schüler*innen nicht ausreichend berücksichtigt wird und dass eine individuelle und differenzierte Förderung weiterhin notwendig ist (Ernst, 2023; Liebers et al., 2023; Röhner et al., 2020). Wichtig erscheint vor diesem Hintergrund vor allem die Entwicklung und Durchführung qualitativ hochwertiger Fortbildungen und Konzepte zur Unterstützung der Lehrkräfte bezüglich des Umgangs mit psychisch belasteter Schüler*innen

¹ Deutschland unterzeichnete das Übereinkommen am 30. März 2007, ratifizierte es im Dezember 2008 und es trat am 1. Januar 2009 in Kraft. Nach Hinterlegung der Ratifikationsurkunde in New York am 24. Februar 2009 wurde das Übereinkommen am 26. März 2009 verbindlich.

(Baumgardt, 2016; Hanisch et al., 2023). Diese Befunde bestätigen, in Einklang mit anderen Untersuchungen, dass eben jene Kinder in psychischen und sozialen Problemlagen und die daraus resultierenden Schwierigkeiten die Hauptbelastungsfaktoren für Lehrkräfte darstellen (Agyapong et al., 2022; Lehr, 2014). Die Belastung äußert sich häufig in Form von externalisierenden Verhaltensweisen, wie z. B. disruptives, hyperaktives oder aggressives Verhalten (Agyapong et al., 2022; Jensen et al., 2023; Li et al., 2022). Ohne ausreichende Beachtung dieser Problemlagen sinkt das allgemeine Wohlbefinden der Lehrkräfte und der erlebte Stress samt dem Risiko an Burnout zu erkranken steigt (Li et al., 2022). Dieser Umstand hat auch Auswirkungen auf den Unterricht. So konnten Studien z.B. zeigen, dass stressbehaftete Lehrkräfte häufig qualitativ weniger guten Unterricht geben, der direkte negative Auswirkungen auf akademisches und sozial-emotionales Lernen der Schüler*innen haben kann (Agyapong et al., 2022; Evers et al., 2016). Dies betont umfassend die Notwendigkeit, Lehrkräfte in ihrem Umgang mit herausfordernden Verhaltensweisen zu stärken, sodass sie im Weiteren den Schüler*innen ein funktionierendes Lernumfeld bieten können. Erschwerend hinzukommt, dass insbesondere verhaltensauffällige Schüler*innen eine organisierte und strukturierte Lernumgebung samt individuellen Interventionen benötigen, um effektiv beschult werden zu können (Aldabbagh et al., 2022). Wird diese Umgebung nicht geschaffen, verstärken sich die Verhaltensprobleme, was die wechselseitige Beziehung zwischen Lehrkraftbelastung und dem Auftreten von z. B. störendem Verhalten unterstreicht (Breitenstein et al., 2009; Docherty et al., 2023; Müller, 2021).

Um auf diese Problemlagen der (inkluisiven) Grundschulen zu reagieren und aufbauend auf den vorhandenen Strukturen deutscher Grundschulsysteme eine systempädagogische Lösung anzustreben, bietet sich der Blick auf Strategien oder Systeme anderer europäischer und westlicher Länder mit vergleichbaren Bildungssystemen und Institutionen an. Hierbei fällt auf, dass sowohl z.B. in Norwegen und Finnland sowie in den USA gestufte Fördersysteme (engl. Multi-Tiered Systems of Support, MTSS) in Grundschulen implementiert wurden, welche die Lehrkraftbelastung und das externalisierende Verhalten der Schüler*innen reduzieren konnten (z. B. Eklund et al., 2020; Lee & Gage, 2020; Sørli & Ogden, 2007; Vainikainen et al., 2017). Diese Ergebnisse sind jedoch nicht direkt auf die deutschen Grundschulen übertragbar. Grund hierfür ist die individuelle Entwicklung eines gestuften Fördersystems entlang Voraussetzungen und Rahmenbedingungen der jeweiligen Schule samt dem umgebenden Bildungssystem (Fox et al., 2021; Schaffer, 2023). Auf erste inhaltliche Überlegungen zur Ausgestaltung solcher Förderansätze (z. B. Hanisch et al., 2019) samt empirischen Überprüfungen kann in Deutschland jedoch zurückgegriffen werden (z. B. Voß et

al., 2014). Es fehlt aber an breiten evidenten Forschungsergebnissen, die zusätzliche Aussagen über die potenzielle Effektivität der Implementation von MTSS in Deutschland möglich machen (Kap. 3). Aus diesem Desiderat und der damit verbundenen Möglichkeit inklusiv ausgerichteter Beschulung ergaben sich folgende Leitfragen dieser kumulativen Dissertation:

- 1) Welche MTSS mit verhaltensbezogener Ausrichtung im Grundschulbereich existieren weltweit und wie werden sie umgesetzt?
- 2) Welche Effekte lassen sich für MTSS im Grundschulbereich mit verhaltensbezogenen Interventionen im internationalen Vergleich ableiten?
- 3) Welche Wirksamkeit kann ein MTSS in der Grundschule in Deutschland zur Reduktion externalisierenden Verhaltens erzielen?
- 4) Wie wirksam sind die eingesetzten evidenzbasierten Interventionen eines MTSS zur Reduktion externalisierenden Verhaltens in der Grundschule in Deutschland?
- 5) Welche Herausforderungen und Gelingensbedingungen lassen sich bzgl. der erfolgreichen Implementation von MTSS in der Grundschule in Deutschland feststellen?
- 6) Welche Schritte erscheinen notwendig, um die nachhaltige Distribution der vorhandenen Forschungsergebnisse zu gewährleisten, um somit niederschwellig die Implementation von MTSS in der Grundschule in Deutschland zu unterstützen?

Um diese Leitfragen zu beantworten, bedarf es verschiedenen Studien samt unterschiedlichen Methodiken und Forschungszugängen und sind Grundlage dieser kumulativen Dissertation. Um sie zu beantworten, wurden sechs verschiedene Studien durchgeführt, die in dieser Rahmenschrift zusammengefasst werden. Der folgende Text stellt somit eine Rahmung der durchgeführten Untersuchungen und ihren Kernergebnissen dar. Auf diese Weise konnte ein umfassender Überblick über die Forschungslage zum Thema MTSS in der Grundschule in Deutschland entstehen: Beitrag 1 schafft in Form eines systematischen Literaturreviews einen Überblick über die bestehende internationale Forschungslandschaft zum Thema MTSS, welches den weiteren Studien als Fundament und Grundlage dient. Die Beiträge 2 und 6 befassen sich anschließend mit der Wirksamkeit und der Implementation des gestuften Förderansatz *Multimo* (Hanisch et al., 2019), welcher 2019 entwickelt und bis zur Umsetzung dieser Studien nicht empirisch überprüft wurde. Gestützt durch die Erkenntnisse aus Beitrag 1 eignet sich dieser Ansatz dazu, in der Grundschule durch verhaltensmodifizierende Interventionen externalisierendes Verhalten zu reduzieren. Diese Wirksamkeits- und

Implementationsforschung wird durch die Beiträge 3, 4 und 5 ergänzt, welche die Interventionen der Stufen 1, 2 und 3 des Ansatzes überprüfen. Auf Grundlage der Ergebnisse konnte die Effektivität der Interventionen in einer deutschen Grundschul-Sample bestätigt werden. Nahezu alle Ergebnisse der vorliegenden Arbeit sind bereits in wissenschaftlichen Fachzeitschriften veröffentlicht worden. Lediglich die Analysen der Einzelfalldaten der Stufe 3 (Kap. 5.4) sind für diese Arbeit ergänzend aufbereitet worden und bisher nicht in einer wissenschaftlichen Fachzeitschrift publiziert. Beitrag 5 befindet sich derzeit im Begutachtungsverfahren. Die vorliegende Arbeit schließt mit einer Methodik-Kritik und einer inhaltlichen Diskussion. Auf dieser Grundlage wird beurteilt, welche Schritte vollzogen werden können, um die nachhaltige Implementation gestufter Fördersysteme in Deutschland zu fördern und somit das Ziel, die inklusive Beschulung aller (psychisch und sozial belasteter) Kinder und die Entlastung der Lehrkräfte in der Grundschule, voranzutreiben. Zusätzlich wird die Bedeutung eines gelingenden Praxis-Wissenschaft-Transfers samt Umsetzungsmöglichkeiten diskutiert.

2 Externalisierendes Verhalten

Um im Folgenden einen der Hauptbelastungsfaktoren der Schule, das externalisierende Verhalten (Scherzinger & Wettstein, 2023), und die damit verbundenen Folgen samt möglichen schulischen Interventionen näher zu beleuchten, ist eine Gegenstandsbestimmung notwendig. Externalisierendes Verhalten stellt ein Sammelbegriff für unaufmerksames, hyperaktives, aggressives, oppositionelles und verweigerndes Verhalten dar (Döpfner & Banaschewski, 2022; Kauten & Barry, 2016). Insbesondere in der Schule kann externalisierendes Verhalten weitreichende Folgen für Lehrkräfte und die Schüler*innen haben (Aldabbagh et al., 2022; Hoglund et al., 2015). Es hängt z.B. mit erhöhtem Stress der Lehrkräfte, einem negativen Schul- und Klassenklima, sowie sozialer Ausgrenzung und ausbleibendem schulischen Erfolg der Schüler*innen zusammen (Agyapong et al., 2022; Evers et al., 2016; Li et al., 2022; Wartenberg et al., 2023). Auch im Feld der klinischen Störungsdiagnostik kommt dem externalisierenden Verhalten eine wichtige Bedeutung zu: So wird es z.B. im *Diagnostischen und Statistischen Manual Psychischer Störungen* (DSM-V, APA, 2018) als eine Gruppe an Verhaltensmerkmalen beschrieben, die durch „impulsive und substanzbezogene Symptome sowie disruptives Sozialverhalten“ (S. 17) gekennzeichnet ist. Dem externalisierenden Verhalten gegenüber steht das internalisierende Verhalten, welches zurückgezogene, ängstliche, gehemmte sowie depressive Verhaltensweisen umfasst (z.B. Laucht et al., 2000). Die Art und Weise sowie die Häufigkeit des Auftretens von externalisierendem (und internalisierendem) Verhalten kann entscheidend für die Vergabe einer psychischen Diagnose sein (APA, 2018). Ein Beispiel hierfür findet sich in den diagnostischen Kriterien für die *Störungen des Sozialverhaltens* (SSV, englisch Conduct Disorder² [CD]). Für die Diagnose dieser psychischen Störung ist z. B. das Auftreten von aggressivem Verhalten oder von regelverletzendem Verhalten erforderlich³. Darüber hinaus ist externalisierendes Verhalten in Form von Hyperaktivität oder Impulsivität ein wesentlicher Faktor für die Diagnose der *Aufmerksamkeitsdefizit-/Hyperaktivitätsstörung* (ADHS, englisch Attention Deficit Hyperactivity Disorder [ADHD] Döpfner & Banaschewski, 2022; von Wirth & Döpfner, 2023). Diese beschriebene Einteilung psychischer Störungsformen

² Klassifiziert nach: WHO. (2016). *International statistical classification of diseases and related health problems* (10 Aufl., Bd. 2).

³ Es liegt ein repetitives und anhaltendes Verhaltensmuster vor, durch das die grundlegenden Rechte anderer oder wichtige altersentsprechende gesellschaftliche Normen oder Regeln verletzt werden. Dies manifestiert sich im Auftreten von mindestens 3 der folgenden 15 Kriterien aus einer der nachfolgenden Kategorien während der letzten 12 Monate, wobei mindestens ein Kriterium in den letzten 6 Monaten erfüllt sein muss: Aggressives Verhalten gegenüber Menschen und Tieren, Zerstörung von Eigentum, Betrug oder Diebstahl, Schwere Regelverstöße (APA, 2018, S.645)

stellte einen kategorialen Ansatz psychischer Störungsdiagnostik dar (Döpfner, 2022). Entgegen dieser kategorialen Einteilung psychischer Störungen steht der dimensionale Ansatz psychischer Störungsdiagnostik: Hier werden psychische Störungen nicht durch die Einteilung in Kategorien festgestellt, sondern durch die Beschreibung entlang kontinuierlich verteilter Merkmalsdimensionen (Döpfner, 2022). Diese dimensionale Erfassung von Verhaltensauffälligkeiten ermöglicht somit genauere Beschreibungen des Merkmalszustandes, wie z.B. fließende Übergänge zwischen normalen und abnormen Verhaltensweisen. Der dimensionale Ansatz erweist sich insbesondere dann als zweckmäßig, wenn sich das zu beschreibende Verhalten durch das Fehlen klar definierbarer Grenzen auszeichnet (z. B. Lahey et al., 2017). Im Folgenden soll aufbauend auf diesen teilweise kategorialen und dimensional beschriebenen externalisierenden Verhaltensweisen ein Überblick über die Prävalenz, die Entstehung und die Folgen samt Interventionsmöglichkeiten geschaffen werden.

2.1 Die Prävalenz von externalisierendem Verhalten

Bezüglich der Auftretenshäufigkeit von externalisierendem Verhalten lässt sich nur schwer eine umfassende, allgemeingültige Prävalenzzahl angeben. Hierfür ist der bereits beschriebene breite Begriff samt unterschiedlichen Operationalisierungen verantwortlich (Keil & Price, 2006). Die Orientierung an Prävalenzraten psychischer Störungen und Auffälligkeiten eignet sich jedoch, um ein grobes Bild der Häufigkeit zu erlangen. Durch die KiGGS-Studie konnte festgestellt werden, dass 16,9 % der Kinder in Deutschland psychische Auffälligkeiten entlang des *Strengths and Difficulties Questionnaire* (SDQ, Goodman, 1997), aufweisen, nicht aber zwangsläufig eine klinische Diagnose erhalten haben (Klipker et al., 2018). Die Daten aus dem Elternbericht der selben Kohorte zeigen ergänzend, dass 4,4 % der 3- bis 17-jährigen Kinder und Jugendlichen mit ADHS diagnostiziert wurden (Göbel et al., 2018). In einem internationalen Vergleich kommt die Forscher*innengruppe um Husky et al. (2018) zu vergleichbaren Ergebnissen: Sie konnten zeigen, dass die Prävalenzraten für ADHS zwischen verschiedenen europäischen Ländern zwischen 1,6 % und 5,6 % variieren. Weitere umfassende Meta-Analysen bestätigen diese Prävalenzraten und beziffern sie auf mindestens 5-6 % weltweit (Faraone et al., 2024; Petermann, 2005; Polanczyk et al., 2007; Polanczyk et al., 2015; Polanczyk et al., 2014). Eine Studie aus Nordrhein-Westfalen ergab, dass Lehrkräfte 60,5 % der Schüler*innen mit festgestelltem Förderbedarf im Bereich Emotionaler und sozialer Entwicklung als auffällig oder sehr auffällig im Bereich ADHS einschätzten (Hanisch et al., 2023). Weiterhin wurden Auffälligkeiten bei 42 % im Bereich oppositioneller Störungen und

25,9 % im Bereich SSV festgestellt. Nur 9 % der Schüler*innen wurden als unauffällig eingestuft. Ergänzend beziffert eine Meta-Analyse die weltweite Prävalenz von SSV bzw. CD auf einen Wert von 8 % ([Mohammadi et al., 2021](#)). Weitere Studien bestätigen diesen Wert und berichten Prävalenzraten von bis zu 10 % ([Norozpour et al., 2023](#)). Somit zeigt sich durch die zahlreichen Studien, dass externalisierendes Verhalten und die häufig damit verbundenen psychische Auffälligkeiten und Störungen bei etwa 5-10 % der Kinder- und Jugendlichen weltweit auftreten, was die Notwendigkeit gezielter Präventions- und Interventionsmaßnahmen unterstreicht. Um diese erfolgreich einzusetzen, ist jedoch zunächst ein Blick auf die Entstehung externalisierenden Verhaltens notwendig.

2.2 Die Ätiologie von externalisierendem Verhalten

Die Entstehung externalisierenden Verhaltens ist oftmals das Ergebnis eines multifaktoriellen Zusammenspiels biologischer, psychologischer, sozialer sowie umweltbedingter Faktoren ([Bjorkqvist, 2018](#); [Cicchetti & Lynch, 1993](#); [Sturmey, 2022](#); [Zarakoviti et al., 2021](#)). Im Folgenden werden die verschiedenen Faktoren, die zur Entstehung externalisierenden Verhaltens beitragen können, näher beleuchtet. Bedeutend für die Entwicklung externalisierenden Verhaltens sind genetische Prädispositionen bzw. biologische Faktoren. Zwillings- und Adoptionsstudien konnten z. B. zeigen, dass bestimmte Verhaltensweisen, wie Aggressivität und Impulsivität, einen erblich bedingten Ursprung haben können ([Dobrosavljevic et al., 2023](#); [Nikstat & Riemann, 2023](#)). Diese Faktoren beeinflussen z.B. die Sensibilität eines Individuums für externe Umwelteinflüsse und können somit die Ausprägung externalisierenden Verhaltens beeinflussen ([Peel et al., 2023](#)). Ergänzend hierzu konnten neurobiologische Auffälligkeiten, die bei Personen mit externalisierendem Verhalten festgestellt wurden (z. B. [Baving et al., 2003](#); [Huffman, 2022](#)). Diese umfassen Veränderungen in der Struktur und Funktion von Gehirnregionen, die für die Emotionsregulation, Impulskontrolle und Informationsverarbeitung zuständig sind ([Baving et al., 2003](#); [Wilson et al., 2024](#); [Wong et al., 2021](#); [Zhang et al., 2022](#)).

Eng hiermit verknüpft sind psychologische Einflussfaktoren, welche eine entscheidende Rolle bei der Entwicklung und Aufrechterhaltung externalisierenden Verhaltens spielen ([Phillips et al., 2023](#)). Zu diesen Faktoren gehören Persönlichkeitsmerkmale wie hohe Impulsivität und eine niedrige Frustrationstoleranz ([Perry & Price, 2018](#)). Ergänzend kann eine verzerrte sozial-kognitive Informationsverarbeitung, die zu inadäquaten Interpretationen sozialer Interaktionen führt, zur Manifestation externalisierenden Verhaltens beitragen

(Brodbeck et al., 2022; Crick & Dodge, 1994). Oft geht dies mit Selbstregulationsdefiziten einher, die die Kontrolle über Gedanken, Emotionen und Verhalten umfassen (Calvete & Orue, 2012; Denham & Bassett, 2020). Diese Unfähigkeit, die eigene Emotionen und Impulse effektiv zu managen, kann somit ebenfalls zu externalisierendem Verhalten führen.

Als weitere wichtige Einflussfaktoren gelten die Familie und das soziale Umfeld (Buchanan-Pascall et al., 2018). Insbesondere die Familie, als primäre soziale Umgebung, spielt eine wesentliche Rolle bei der Entwicklung externalisierenden Verhaltens (Zarakoviti et al., 2021; Zhai et al., 2024). In Verbindung hiermit gelten Erziehungsstile, die durch Inkonsistenz, hohe Strenge oder mangelnde Wärme gekennzeichnet sind, als Risikofaktoren (Liu, 2004; Taillieu & Brownridge, 2015). Zudem kann das Erleben von Gewalt oder Vernachlässigung in der Familie die Entwicklung aggressiven Verhaltens fördern (Docherty et al., 2023; Zhai et al., 2024). Eng verknüpft ist hiermit oftmals der sozioökonomischer Status, da Stressfaktoren, wie Armut, beengte Wohnverhältnisse und Bildungsbenachteiligung ebenfalls die Entwicklung und Aufrechterhaltung von externalisierendem Verhalten beeinflussen können (Ayano et al., 2024; Chang & Qin, 2018; Gach et al., 2018; Reinelt et al., 2019). Weitere Umfeldfaktoren können z.B. negative Peer-Einflüsse und Medienkonsum, darstellen (Demirtaş-Madran, 2023; Norozpour et al., 2023). Insbesondere die Exposition gegenüber gewalttätigen Inhalten, kann Verhaltensweisen von Kindern und Jugendlichen beeinflussen und zur Exposition externalisierenden Verhaltens beitragen (Demirtaş-Madran, 2023; Farahani et al., 2023; Farmer & Xie, 2007; Janssens et al., 2015; Lee et al., 2021; Yoon et al., 2021).

Zusammenfassend verdeutlicht dies, dass die Entstehung externalisierenden Verhaltens nicht auf eine einzige Ursache zurückgeführt werden kann. Die Entstehung ist vielmehr das Ergebnis eines komplexen Zusammenspiels zwischen diversen biologischen, psychologischen, sozialen sowie umweltbedingten Faktoren.

2.3 Auswirkungen von externalisierendem Verhalten

Ähnlich divers, wie sich die Entstehung externalisierenden Verhaltens darstellt, sind die Auswirkungen jener Verhaltensweisen. Allgemein stellt externalisierendes Verhalten einen bedeutenden Risikofaktor und Prädiktor für spätere Jugenddelinquenz, Erwachsenenkriminalität und Gewalt dar (Lavigne et al., 2009; Liu, 2004; Matza et al., 2005; Weber, 2017; Willcutt, 2012). Genauer können Studien negative Konsequenzen für Gesundheit, Bildung, Beruf und die allgemeine soziale Stellung innerhalb der Gesellschaft zeigen (Cohn & Cicero, 2024; Deighton et al., 2018; Fry et al., 2018; Hornsveld & Kraaimaat,

2022; Keil & Price, 2006). Personen mit ADHS und SSV sind z. B. einem erhöhten Risiko für akademische Probleme, Arbeitslosigkeit, Kriminalität, Drogenmissbrauch und mentale Probleme ausgesetzt (Ayano et al., 2024; Erskine et al., 2016; Kulkarni et al., 2020) Eine Studie von Kuriyan et al. (2013) konnte zudem zeigen, dass junge Erwachsene, bei denen in der Kindheit ADHS diagnostiziert wurde, niedrigere Bildungs- und Berufserfolge erzielen als Menschen ohne ADHS. Grund hierfür sind schulische Leistungen und Disziplinprobleme sowie die damit verbundenen niedrigeren Chancen auf einen erfolgreichen Studienabschluss. Weitere Studien kommen ergänzend zu dem Ergebnis, dass jene persistierenden ADHS-Symptome im Erwachsenenalter negative Auswirkungen auf Bildungserfolg und Arbeitsplatzstabilität haben können (Hechtman et al., 2016; von Wirth et al., 2022). Diese weitreichenden Auswirkungen belasten nicht nur die Betroffenen und ihre Angehörigen, sondern verursachen auch erhebliche gesellschaftliche Kosten (Vergunst et al., 2023). Die frühzeitige Erkennung und Behandlung bzw. Prävention von externalisierendem Verhalten ist daher von entscheidender Bedeutung, um negative Entwicklungen und deren psychosozialen Folgen auf gesellschaftlicher und personeller Ebene zu verringern (Rivenbark et al., 2018). Aus diesem Grund wird externalisierendes Verhalten vor allem in den USA und weiteren europäischen Ländern, wie Deutschland, als ein gesellschaftliches Gesundheitsproblem angesehen (Beelmann et al., 2007; Campbell et al., 1995; Göbel et al., 2018; Polanczyk et al., 2015). Die erfolgreiche Prävention gilt daher als eine dringliche Aufgabe der Gesellschaft (Schubarth & Seidel, 2013; Vergunst et al., 2023). Zusammenfassend können diese Befunde die komplexen Wechselwirkungen von externalisierendem Verhalten, Gewalterfahrungen und Bildungserfolgen herausstellen, was zusätzlich die Bedeutsamkeit der Implementation von Interventionen, die jene Verhaltensweisen effektiv adressieren und abmildern können, untermauert. Ein wichtiger Ort hierfür ist neben dem familiären Umfeld die Schule der betroffenen Kinder.

2.4 Externalisierendes Verhalten im schulischen Kontext

Präventive Ansätze beinhalten im Sinne der multifaktoriellen Entstehung externalisierender Verhaltensweisen häufig ebenso multimodale pädagogische Interventionen (Fossum et al., 2008; Zarakoviti et al., 2021). Oftmals werden daher therapeutische Interventionen wie die kognitive Verhaltenstherapie, Problemlösetrainings, Selbstkontrolltechniken oder operante Ansätze angewandt und kombiniert (Buchanan-Pascall et al., 2018; Tully & Hunt, 2015). Ein wesentlicher Fokus liegt hierbei auch auf der Arbeit mit den Bezugspersonen des Kindes, vor allem Eltern und Lehrkräften, um ein unterstützendes Umfeld zu schaffen (Aldabbagh et al.,

2022). Dabei ist eine enge Zusammenarbeit zwischen verschiedenen Fachdisziplinen unerlässlich, um eine ganzheitliche Betrachtung des Kindes in seinem schulischen und familiären Umfeld zu gewährleisten (Bundschuh, 2023; Gebhardt et al., 2022; Voß et al., 2020). Der Schule kommt daher eine große Bedeutung in Bezug auf die Prävention und Intervention von externalisierendem Verhalten zu (Maynard et al., 2015; Scherzinger & Wettstein, 2023). Kinder mit externalisierendem Verhalten zeigen im schulischen Kontext oft eine Reihe von unangepassten Verhaltensmustern in Form von Hyperaktivität, Unaufmerksamkeit, Ungehorsam, Impulsivität oder Wutanfällen. Diese Verhaltensweisen führen nicht nur zu Störungen während des Unterrichts, sondern können auch die Peer-Beziehungen der Kinder beeinträchtigen (Krull et al., 2022; Müller, 2021). Kinder mit externalisierendem Verhalten erfahren somit mehr Ablehnung von Mitschüler*innen und haben größere Schwierigkeiten, Freundschaften aufzubauen (Guevremont & Dumas, 1994; Mikami et al., 2013). Sie zeigen außerdem Probleme in Gruppenarbeiten, was ihre soziale Entwicklung und Integration in die Klassengemeinschaft zusätzlich erschwert (Brundrett et al., 2022). Lehrkräfte, die Schüler*innen mit externalisierenden Verhaltensweisen unterrichten, sind demnach besonderen Belastungen ausgesetzt (Agyapong et al., 2022; Evers et al., 2016). Ihre Fähigkeiten der Klassenführung und ihr allgemeines Wohlbefinden können durch die ständige Auseinandersetzung mit externalisierenden Verhaltensweisen beeinträchtigt werden (Grayson & Alvarez, 2008; Lehr, 2014). Studien zeigen darüber hinaus, dass Lehrkräfte, die Schüler*innen mit externalisierendem Verhalten unterrichten, ein niedrigeres Selbstwirksamkeitserleben und ein höheres Stresslevel haben, was sich auf die angewendeten Unterrichtsmethoden und die Beziehungsgestaltung auswirkt (Hastings & Brown, 2002; Herman et al., 2017; Skaalvik & Skaalvik, 2010; Tschannen-Moran et al., 2016). Lehrkräfte, die sich unsicher fühlen, neigen demzufolge dazu, mehr Kritik und weniger Wärme zu zeigen, was wiederum das Verhalten der Schüler*innen negativ beeinflussen kann (Ahnert et al., 2012; Koenen et al., 2017). Schulen stehen daher vor der Herausforderung, effektive Strategien zur Bewältigung dieser externalisierenden Verhaltensweisen zu entwickeln und greifen hierbei auf unterschiedliche Interventionen im Bereich externalisierenden Verhaltens zurück, die im folgenden Kapitel näher betrachtet werden.

2.5 Interventionen im Kontext externalisierenden Verhaltens

Ebenso breit, wie sich die Forschungslage zur Ätiologie, Prävalenz und den Auswirkungen externalisierenden Verhaltens darstellt, ist sie bezogen auf die Intervention jener

Verhaltensweisen. Insbesondere im schulischen Kontext konnten diverse Meta-Analysen und systematische Überblicksarbeiten zeigen, dass externalisierendes Verhalten von Kindern durch den Einsatz von schulischen Interventionen reduziert werden kann (Durlak et al., 2011; Hövel et al., 2019). Oft sind diese multimodal ausgerichtet, sodass sie nicht nur in einem Setting eingesetzt werden, sondern sich neben der Schule auch auf die elterliche Erziehung oder andere Lebensbereiche des Kindes erstrecken (Tehrani et al., 2023). Aufgrund des starken Fokus auf das Verhalten des Kindes zielen Interventionen im Kontext externalisierenden Verhaltens auch auf verhaltensbestimmende und -aufrechterhaltende Faktoren, wie den elterlichen Erziehungsstil oder pädagogische Handlungen der Eltern ab. Demnach wurden neben vielen schulbezogenen Interventionen auch diverse elternbezogene Interventionsprogramme entwickelt (de Jong et al., 2023).

Zusätzlich zur multimodalen Ausrichtung ist eine frühzeitige Intervention von großer Bedeutung. Dies konnte z. B. durch eine Meta-Analyse von Luo et al. (2020) herausgestellt werden. Die Autor*innen überprüften den Effekt klassenweiter Interventionen für Vorschulkinder und konnten signifikante Verbesserungen der sozialen und emotionalen Kompetenzen sowie eine Reduktion des externalisierenden Verhaltens feststellen. Die Integration von familienbezogenen Interventionen konnte diese Effekte verstärken, was erneut die Relevanz einer umfassenden Betrachtung des kindlichen Umfelds betont. Parallel dazu zeigt eine weitere Meta-Analyse von Stoltz et al. (2012), dass individuelle, schulbasierte Interventionen externalisierendes Verhalten bei Grundschulkindern signifikant reduzieren können. Auch für diese Altersgruppe scheinen Effekte größer zu sein, wenn diese Interventionen zusätzliche Komponenten wie die Integration von Eltern, Lehrkrafttrainings oder Peer-Unterstützungsprogrammen beinhalten. Insbesondere Kinder mit schwereren Verhaltensproblemen profitieren besonders stark von Elterntrainings (Hautmann et al., 2010).

Die hier angesprochenen Interventionen können sich hinsichtlich ihrer konkreten Umsetzung und des vorrangigen Interventionsziels stark unterscheiden: So konnten Effekte nachgewiesen werden für gruppenbezogene Interventionen (Joslyn et al., 2019; Leidig et al., 2022; Pennington & McComas, 2017) oder individuumsbezogene Interventionen (Ackley, 2022; Pyle & Fabiano, 2017). Zusätzlich konnten Effekte gezeigt werden für die Reduktion von externalisierendem Verhalten in Form von Cyber-Mobbing (Polanin et al., 2022), physischer Gewaltanwendung (Bilz & Fischer, 2020; Bolton, 2010; Klomek et al., 2014), disruptiven Verhaltensweisen im Unterricht oder verweigernden Verhaltensweisen im Schulkontext (Buchanan-Pascall et al., 2018; de Jong et al., 2023; Hanisch et al., 2020; Maynard et al., 2015).

Eine weitere, insbesondere in den USA viel beforschte Möglichkeit externalisierendes Verhalten zu verringern, sind gestufte Fördersysteme bzw. MTSS (z.B. Schaffer, 2023). Diese Systeme variieren je nach Schule und Bildungssystem und wurden ursprünglich in den 1950er- Jahren für öffentliche Gesundheitsanliegen in den USA entwickelt. Die Besonderheit eines MTSS liegt in der gezielten Kombination unterschiedlicher Interventionen, um den vorhandenen Problembereichen eines Kindes möglichst passend zu begegnen. Sie bieten demnach keine eindimensionale Lösung, sondern adressieren verschiedene Bereiche der schulischen Unterstützung und beziehen dabei möglichst viele Personengruppen (Eltern, Psycholog*innen, Lehrkräfte usw.) in die Förderung mit ein (Burns et al., 2016; Walker et al., 2023). Diese gezielte Anpassung der Förderung kann sowohl akademische als auch verhaltensbezogene Variablen fokussieren (Harlacher et al., 2014). Viele Studien haben die Wirksamkeit gestufter Förderung bezüglich verschiedener Zielvariablen nachgewiesen: Es konnte gezeigt werden, dass die Implementation von MTSS unter anderem Auswirkungen auf die Selbstwirksamkeit und die Stresserfahrung der Lehrkräfte, das akademische Lernen, das externalisierende Verhalten sowie auf das Schulklima und internalisierende Verhaltensprobleme haben kann (Aasheim et al., 2018; Batsche, 2014; Berger, 2019; Clark & Dockweiler, 2020; Ruth A. Ervin et al., 2006; Lee & Gage, 2020; Simonsen, 2021; Stephan et al., 2015). Sie bieten somit eine Möglichkeit, externalisierendes Verhalten in Form eines multimodalen Systems zu reduzieren und sollen daher im Folgenden näher betrachtet werden.

3 Mehrstufige Fördersysteme (MTSS)

Im Kontext mehrstufiger Fördersysteme werden oft unterschiedliche Begriffe genannt, die zumeist das gleiche pädagogische System beschreiben (z. B. *gestufte Förderung*, *MTSS*, *tiered systems*). Zudem wird der Begriff *Response to Intervention* (RTI) häufig synonym verwendet zu MTSS (Avant, 2016; Burns et al., 2016). Es ist jedoch wichtig, hier eine Unterscheidung zu treffen und die Begriffe voneinander abzugrenzen: Historisch ist nachzuzeichnen, dass der Fokus auf verhaltensmodifizierenden Maßnahmen innerhalb der Fördersysteme erst in der jüngeren Vergangenheit zunahm und jene Systeme ursprünglich verstärkt auf die Förderung der akademischen Fähigkeiten der Schüler*innen ausgerichtet waren (Bradley et al., 2005; Lee & Gage, 2020; Shepley et al., 2020). Vor diesem Hintergrund wurden Systeme wie *School-wide Positive Behavior Support* (SWPBS), *Response to Intervention* (RTI) oder *Positive Behavior Support* (PBS) unter dem Begriff *Multi-tiered Systems of Support* subsumiert. McIntosh und Goodman (2016) beschreiben MTSS dementsprechend wie folgt:

An integrated MTSS model provides all students with the best opportunities to succeed both academically and behaviorally in school. MTSS focuses on providing high-quality instruction and interventions matched to student need across domains and monitoring progress frequently to make decisions about changes in instruction or goals. It is not simply the implementation of both academic RTI and PBIS systems. There is a systematic and careful integration of these systems to enhance the efficiency and effectiveness of all school systems. (S. 6)

Hier wird deutlich, dass innerhalb eines MTSS sowohl akademische als auch verhaltensbezogene Maßnahmen zum Einsatz kommen (können). Beide Interventionsdimensionen (akademisch und verhaltensbezogen) sind in einen Prozess schulischer Diagnostik und Entscheidungsfindung eingebunden. Um diese Grundprinzipien und ihre Entwicklung näher zu verstehen, ist ein historischer Abriss, orientiert an wichtigen politischen, wissenschaftlichen und gesellschaftlichen Entwicklungen im Ursprungsland der gestuften Förderung, den USA, hilfreich.

3.1 Politische und wissenschaftliche Rahmung gestufter Förderung

Maßgeblich verantwortlich für die Entwicklung und Umsetzung gestufter Fördersysteme waren gesellschaftliche Entwicklungen in den USA rund um die Rechte und Teilhabe von Menschen mit Behinderungen (Dobbins & Bieber, 2014; Goodman-Scott et al., 2023). In den 1960er- und 1970er Jahren wurden auf dieser Grundlage Kindern und Jugendlichen mit Behinderung im Sinne einer bildungspolitischen Neuorientierung umfängliche Rechte und Zugänge zu individualisierter Bildung zugestanden. Verantwortlich hierfür waren vor allem das *Public Law 94-142* (National Education Association of United States, 1978) und der *Individuals with Disabilities Education Act* (IDEA) und seine Neuauflage (United States Congress, 2004, 2015; Dobbins & Bieber, 2014). Sie definieren Bildungsgrundsätze, wie z. B. ein Recht auf freie und öffentliche Bildung, individualisierte Bildungszugänge samt möglichst barrierefreiem Lernen, Frühförderung und eine datengestützte Entscheidungsfindung im Kontext von Förderung. Wissenschaftlich gerahmt wird jene politische Entwicklung durch Autoren wie Reynolds (1962) und Deno (2018), die politische Entscheidungen und Entwicklungen mit Bezug zur pädagogischen Praxis aufbereiten und ausformulieren. Reynolds (1962) betont vor allem die Notwendigkeit von differenzierenden Bildungsprogrammen und schulischen Interventionen, welche notwendig sind für eine gelingende gestufte sonderpädagogische Unterstützung. Seine Grundgedanken und Systembeschreibungen ähneln einem Leitfaden für pädagogische Fachkräfte und politische Entscheidungsträger*innen, um effektive Strategien zur Unterstützung von Kindern mit besonderen Bedürfnissen erarbeiten zu können. Kurze Zeit später beschreibt Deno (2018) weitere Potenziale der Sonderpädagogik und unterstreicht die Dringlichkeit von Veränderungen in den gängigen sonderpädagogischen Praktiken. Es wird argumentiert, dass der vorherrschende Fokus auf Pathologie in der Sonderpädagogik unzweckmäßig ist und er plädiert für eine stärkere Ausrichtung auf die Förderung der Selbstverwirklichung („adequate problem-solvers“ S. 230). Zudem betont er nachdrücklich die Wichtigkeit der Anerkennung der individuellen Menschlichkeit jedes Kindes und formuliert einen aus heutiger Sicht inklusiven Gedanken von systematisierter Bildung. Alles in allem lassen sich in diesen beiden Werken die Grundpfeiler einer inklusiv ausgerichteten, systemischen, multiprofessionellen und multimethodalen (sonderpädagogischen) Schulpädagogik finden, die im Folgenden entlang ihrer rechtlichen Rahmung politisch umgesetzt werden sollte.

3.2 Die pädagogischen Grundpfeiler gestufter Förderung

In den 1960er-Jahren wurde ausgehend von Deno und Mirkin die Entwicklung und Anwendung des *Curriculumbasierten Messens* publik (CBM, Deno & Mirkin, 1977; Jenkins & Fuchs, 2012). Im Zentrum stand die Überwachung des Fortschritts und die Entscheidung über potenzielle Fördermaßnahmen samt einer Anpassung des Unterrichtsmaterials für die Schüler*innen. Ein Augenmerk lag zusätzlich auch auf der Binnendifferenzierung des Lernstoffs, sodass auch Kinder mit Behinderung verstärkt in den Unterricht integriert werden sollten. Zunächst lag der Fokus auf der Alphabetisierung und den basalen Rechenfähigkeiten. Durch diese kurzen, präzisen Messungen (basierend auf dem lokalen akademischen Lehrplan) konnten Aussagen über eine notwendige Adaption des Unterrichts oder der Lerngruppe getroffen werden. Parallel entstand ein Ansatz namens *Precision Teaching* (PT, Evans et al., 2021; Kunzelmann, 1970; White, 1986). PT ist eine frühe Form des *decision making* (Solesbee & Davies, 2021) gestufter Förderansätze und bezieht im Vergleich zum CBM ein vielfältigeres Set von akademischen und verhaltensbezogenen Dimensionen in den Entscheidungsprozess über die Förderung belasteter Kinder und Jugendlichen ein. In Kombination bieten CBM und PT-Fachkräften einen organisatorischen Rahmen, der hilft, die Erfassung relevanter Daten zu systematisieren und die daraus resultierende Förderentscheidung zu verfeinern.

Eine weitere konkrete schulpädagogische und multiprofessionell orientierte Umsetzung jener Grundsätze ist das *Individualized Education Program* (IEP), welches in den 1970er-Jahren implementiert wurde (United States Congress, 2004)⁴. Das IEP ist ein spezielles Bildungsprogramm für Schüler*innen, die zusätzliche Unterstützung in der Schule benötigen und findet bis heute Anwendung. Es zielt grundsätzlich auf die erfolgreiche, inklusive Beschulung benachteiligter Kinder ab und beinhaltet außerdem einige Prinzipien gestufter Förderung. Dazu gehören Aspekte wie teambasierte Planung und zielorientierte Entscheidungen, die auf dem aktuell gezeigten Verhalten Kindes basieren. Angewandte Interventionen und unterrichtliche Methoden sollten evidenzbasiert sein und die Fortschritte der Schüler*innen müssen kontinuierlich monitorisiert werden. Zudem ist das Programm durch fortlaufende Screenings auf universeller Förderebene, dem sog. *Child Find* präventiv ausgerichtet (Goodman-Scott et al., 2023). Potenziell zu fördernde Schüler*innen können auf diese Weise zuverlässig identifiziert werden.

⁴ Zitiert wird die aktuelle Fassung des Gesetzestextes, der fortgeschrieben wurde.

Eng damit verbunden war zusätzlich die Implementation weitreichender Beratungsangebote für Lehrkräfte (Chalfant et al., 1979). Diese gingen einher mit der Einrichtung der sog. *Resource Rooms*, welche allen Schüler*innen, die individuellen oder Kleingruppenunterricht benötigen, zur Verfügung gestellt wurden (O'Connor et al., 2016). Diese Interventionen wurden in das Schulwesen implementiert, um den individuellen Bedürfnissen einiger Schüler*innen umfassender gerecht zu werden. In der Folge kooperierten Lehrkräfte aus dem allgemeinen und dem sonderpädagogischen Bereich eng mit Schulberater*innen und Psycholog*innen. Ihr Ziel war es, eine integrative Lernerfahrung für alle Schüler*innen zu schaffen und spezielle pädagogische Unterstützungen für Kinder mit Lern- und Verhaltensschwierigkeiten zu implementieren (Stoiber & Gettinger, 2016). Besondere Aufmerksamkeit galt dabei der positiven Wechselbeziehung zwischen akademischer Leistung und Verhaltensmodifikation. Auch hier spiegelt sich demnach ein Grundsatz gestufter Förderung wider.

Diese Entwicklungen der 60er und 70er-Jahre mündeten schlussendlich in den ersten gestuften Fördersystemen rund um das Konzept *Response to Intervention* (RTI, Bradley et al., 2005; Fuchs & Fuchs, 1986; Fuchs & Fuchs, 2007), welches ursprünglich in sonderpädagogischen Kontexten Anwendung fand (Berkeley et al., 2009). In den 90 er-Jahren wurde RTI auch für die allgemeine Schule adaptiert und implementiert. Diese historische Einordnung jener Einflussfaktoren auf die Entwicklung gestufter Fördersysteme zeigt, dass aus einer politischen und gesellschaftlichen Grundhaltung, ein inklusives Schulsystem aufzubauen, langfristig ein durch verschiedene Hilfesysteme ausgestattetes Fördersystem entstanden ist. Der präventive Grundgedanke hat sich dementsprechend fortlaufend verändert. Diese Entwicklung kann zudem nicht als abgeschlossen betrachtet werden, da weiterhin neue Fördersysteme auf gleicher Basis entwickelt werden, um neuen schulischen und pädagogischen Anforderungen gerecht zu werden (Bellmer et al., 2020; Hubert & Grosche, 2012; Voß et al., 2016). Schaffer (2023) z. B. gliedert MTSS in einem aktuellen Überblickswerk MTSS auf in: *RTI, School-wide Positive Behavior Support, Social-Emotional RTI & Trauma Informed Practicies* und *Suicide Prevention & Intervention*. Dieses erweiterte Verständnis von MTSS ist jedoch erst seit kurzer Zeit Bestandteil des Literaturkanons rund um gestufte Fördersysteme und zeigt die fortlaufende Weiterentwicklung und Adaption bestehender Systeme und ihres Grundverständnisses.

3.3 Die Funktionsweise gestufter Förderung

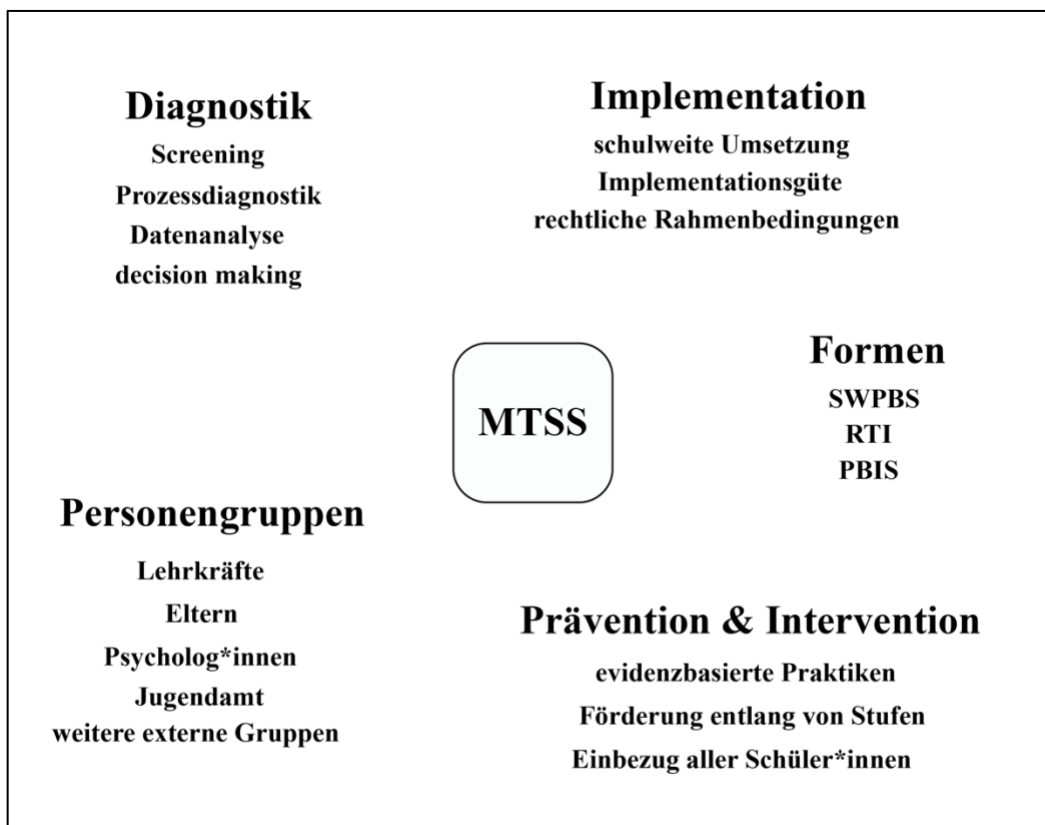
Gestufte Förderung speist sich, wie im vorherigen Kapitel dargelegt, aus verschiedenen pädagogischen Grundsätzen. Diese Grundsätze können in verschiedenen Gesellschaften oder Schulsystemen unterschiedlich inhaltlich gefüllt werden. Aus diesem Grund gibt es nicht *die eine* gestufte Förderung oder *das* MTSS. Gestufte Förderung orientiert sich immer an den jeweiligen Ressourcen, Schutzfaktoren, Risikofaktoren und dem Umfeld eines Kindes – oder globaler einer Schule. Die Berücksichtigung dieser Grundbedingung ist bei der Beschreibung der basalen Funktionsweise von gestufter Förderung wichtig, um nicht den Eindruck zu erwecken, jedes System ist in sich geschlossen gleichermaßen ausgestaltet.

3.3.1 Grundprinzipien und Säulen eines MTSS

Auch wenn es viele unterschiedliche Formen gestufter Fördersysteme gibt, lassen sich gewisse Regelmäßigkeiten und Grundlagen bezüglich der Funktionsweise von MTSS erkennen. Zusätzlich spielen immer wieder ähnliche Personengruppen, diagnostische Prozesse oder die Grundhaltung des Einsatzes präventiver Maßnahmen eine Rolle (vgl. Abb. 1). Harlacher et al. (2014) beschreiben hierauf aufbauend sechs Grundprinzipien von MTSS: Zunächst sollen alle

Abbildung 1

Begrifflichkeiten im Kontext MTSS



Schüler*innen die Möglichkeit bekommen, von dem Fördersystem zu profitieren. Zweitens sollte die Grundausrichtung des MTSS immer eine präventive und proaktive sein, um frühzeitig bestimmen zu können, welche Schüler*innen aufgrund ihrer akademischen und/oder verhaltensbezogenen Bedürfnisse mehr Unterstützung benötigen. Das dritte Grundprinzip ist die Verzahnung von Forschung und Praxis. Es sollten empirisch validierte Interventionen implementiert werden, um die bestmögliche Förderung der Schüler*innen zu garantieren. Als viertes Grundprinzip wird die Verwendung von objektiven Daten (Diagnostik), um Entscheidungen über die Förderung zu treffen, genannt. Fünftens soll das System gänzlich kindzentriert sein und zuletzt einen Ansatz bieten, schulweite Reformen der Schulkultur zu initiieren. Der Einbezug aller relevanten Personengruppen ist hierfür notwendig.

Schaffer (2023) beschreibt darüber hinaus vier Grundsäulen von MTSS, die im Folgenden den Ausgangspunkt für das hier definierte Verständnis von gestufter Förderung darstellen sollen: Er beschreibt die Funktionsweise und die verschiedenen Säulen, die miteinander das Gesamtkonzept eines MTSS ausmachen, als *MTSS-House*. Die verschiedenen Säulen bieten die Basis eines jeden MTSS und ohne die Berücksichtigung dieser würde das Gesamtkonstrukt dysfunktional bleiben bzw. zusammenbrechen. Die erste Säule (*Varying levels of preventative evidence-based supports and interventions*), die er benennt, umfasst die verschiedenen Präventionsstufen eines MTSS, die in Form evidenzbasierter Interventionen allen Schüler*innen akademische und sozial-emotionale Unterstützung bieten sollen. Hierzu gehört im Weiteren eine Variation an Interventionen und eine stete Schüler*innenzentriertheit. Die zweite Säule wird abgebildet durch das *Universal Screening*, welches typischerweise drei Mal im Jahr mit allen Schüler*innen durchgeführt wird, um potenzielle risikobehaftete Schüler*innen zu identifizieren. Dies schließt emotionale, soziale und akademische Dimensionen ein. Die dritte Säule hängt mit der zweiten Säule unmittelbar zusammen und beschreibt das *Progress Monitoring*. Hiermit ist eine fortlaufende Überprüfung des Erfolgs der Maßnahmen auf den Stufen 2 und 3 gemeint. Die Überprüfung der Stufe 2 sollte mindestens alle zwei Monate erfolgen und die der dritten Stufe jede Woche oder zumindest alle zwei Wochen. Um hierauf aufbauend eine Förderentscheidung zu treffen, ist die vierte Säule notwendig. Hier wird das *Data-Based Decision Making* als letzte tragende Säule angeführt, was den gesamten Prozess der Datenerhebung, den Screening-Prozess und weitere Datengewinnungsmethoden wie die Diskussion von Schulnoten o. ä. miteinbezieht. Sollte sich durch die Analyse der Daten herausstellen, dass z. B. ein*e Schüler*in nicht von dem gestuften System und seinen Interventionen profitiert, können auf Grundlage der Daten weitere intensivpädagogische Maßnahmen (*special education referral*) in Betracht gezogen werden.

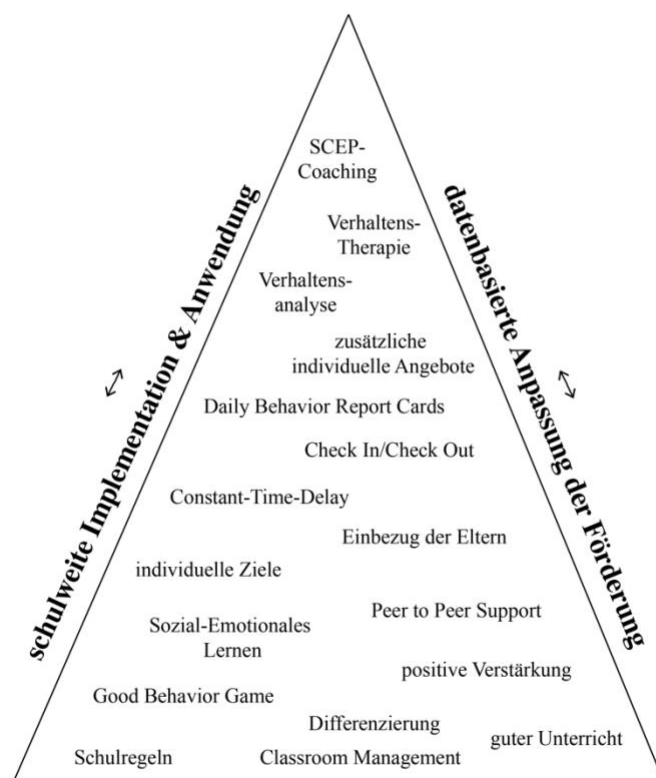
Entlang dieses Verständnisses soll im Folgenden auf die einzelnen Unterpunkte, ergänzt durch weitere Forschungsergebnisse und Konzeptionen, eingegangen werden.

3.3.2 Die Stufen eines MTSS

Einer der wichtigsten und prominentesten Aspekte gestufter Fördersysteme ist die Förderung in Form von Stufen (*tiered system*) (Wexler, 2018). Auf jeder Stufe werden unterschiedliche Hilfen und Interventionen implementiert, die von Stufe 1 zu Stufe 3 bezogen auf die Intensität zunehmen (Stoiber & Gettinger, 2016). Die Anzahl der Schüler*innen, die jene Förderung entlang der unterschiedlichen Stufen erhalten, ist hierbei unterschiedlich: Stufe 1 sollte bis zu 100 % erreichen, also alle Kinder; auf Stufe 2 werden ca. 10-15 % der Schüler*innen gefördert und auf Stufe 3 in etwa 1-5 % (Batsche, 2014; Goodman-Scott et al., 2023; Simonsen, 2021). Diese Werte sind jedoch bloß als grobe, zumeist deskriptive Beschreibungen entlang von

Abbildung 2

Gestufte Förderung als Kontinuum verschiedener Interventionen



Prävalenzraten zu verstehen (Goodman-Scott et al., 2023). Jene Interventionen selbst sind zudem nicht als unveränderbar und abgeschlossen zu betrachten, sondern je nach Kind flexibel adaptierbar⁵. Die hieraus resultierende Darstellung eines Förderansatzes, der auf den verschiedenen Stufen unterschiedliche Interventionen bereithält, mag daher oftmals Missverständnisse produzieren. Vielmehr stellt jedes Kind mit seinen Ressourcen, Benachteiligungen und eventuellen Defiziten einen Anspruch an gewisse Interventionen. Diese verteilen sich

⁵ Im Sinne der Güte jener evidenzbasierten Maßnahmen ist die Umsetzungstreue hierbei nicht außer Acht zu lassen. Siehe hier z. B. Rojas-Andrade, R., & Bahamondes, L. L. (2018). Is Implementation Fidelity Important? A Systematic Review on School-Based Mental Health Programs. *Contemporary School Psychology*, 23(4), 339-350. <https://doi.org/10.1007/s40688-018-0175-0>

auf einem Kontinuum verschiedener Unterstützungsstufen und werden von einem gestuften Fördersystem adressiert und aufgefangen (Abb. 2). Somit wird dem Eindruck einer Etikettierung in Form der Zuweisung einer Stufe zu einem Kind vorgebeugt. Diese Förderung kann auch individuumsbezogen sehr heterogen sein: Ein*e Schüler*in kann sich z. B. im Fach Mathematik auf der universalen Förderstufe (Stufe 1) befinden und im Bereich des Verhaltens entlang der indizierten Förderstufe (Stufe 3) gefördert werden.

Zusammenfassend kann festgehalten werden, dass sich der systemische Grundgedanke gestufter Fördersysteme in einer hierarchisierten Abfolge an evidenzbasierten Interventionen zu unterschiedlichen Entwicklungsbereichen niederschlägt, die den Anspruch haben, möglichst individualisiert und adaptiert im Rahmen des Systems den Bedürfnissen der Schüler*innen zu entsprechen.

3.3.3 Das datengeleitete Vorgehen (decision making)

Ein weiterer tragender Punkt hinsichtlich der Funktionsweise gestufter Fördersysteme ist das datengeleitete Vorgehen in Verbindung mit den Förderentscheidungen (*decision making*) (Oslund et al., 2021; Solesbee & Davies, 2021). Um sicherzustellen, dass Förderdimensionen nicht fälschlich festgelegt oder Defizit- und Kompetenzbereiche übersehen werden, sollten alle Anpassungen in der Interventionsplanung auf fundierten Daten beruhen (z. B. Bundschuh, 2023; Gebhardt et al., 2022). Hier wird der Bezug zum CBM oder PT deutlich (Kap. 3.2). Beide Verfahren oder Ansätze wurden bereits vor der ersten Entwicklung von MTSS in den USA umgesetzt, da sie im Sinne der inklusiven Beschulung Bedarfe und Kompetenzen der Schüler*innen offenbaren. Durch diese Verfahren konnte akademisches und sozial-emotionales Lernen bzw. die Förderung jener Defizit- oder Kompetenzbereiche effektiver gestalten werden.

Auf der universalen Ebene (1. Stufe) werden im Kontext des *decision making* verhaltensbezogene Screenings eingesetzt, um Schüler*innen mit Unterstützungsbedarf zu identifizieren. Die Bandbreite solcher Verfahren ist recht umfassend: Bekannte Screenings sind z. B. der SDQ (Goodman, 1997), die Risk Screening Scale (SRSS, Drummond, 1994) oder das Behavioral and Emotional Screening System (BASC-3 BESS, Kamphaus & Reynolds, 2015). Für den deutschsprachigen Bereich sind z. B. ergänzend die Lehrereinschätzliste für Sozial- und Lernverhalten (LSL, Petermann & Petermann, 2013) und die Schülereinschätzliste für Sozial- und Lernverhalten (SSL, Petermann & Petermann, 2014) zu nennen⁶. Zusätzlich finden

⁶ Diese Auflistung ist nur beispielhaft und keinesfalls vollständig

in Ergänzung dazu auch akademische Universal-Screenings (siehe CBM) Anwendung. Dies sind oft kurze, standardisierte Tests im Bereich Lesegeschwindigkeit oder Textverstehen (Schaffer, 2023). Beliebte Verfahren für solche Screenings sind z. B. *DIBELS* (Good & Kaminski, 2002) und *AIMSweb* (Shinn, 2006). Die Anwendung eigener nicht standardisierter Abfragen ist ebenfalls möglich und kann den Prozess der Entscheidungsfindung unterstützen.

Häufig liegt die Zuständigkeit für die Durchführung der Screenings bei den Klassenlehrkräften oder bei Lehrkräften, die intensiven Kontakt zu einer Klasse haben. Wichtig ist, dass die jeweiligen Personen das Verhalten der Schüler*innen differenziert einschätzen können. In weiten Teilen der USA werden ergänzend lediglich Disziplinarverweise (*Office Discipline Referral*, ODR) der Schüler*innen (digital) festgehalten und dienen als Indikatoren für problematisches Verhalten (Clonan et al., 2006; J. Nitz et al., 2023). Dieses Vorgehen wird z. T. kritisiert und Validitätsprüfungen hinsichtlich der Aussagekraft unterzogen (Martella et al., 2010; K. McIntosh et al., 2009).

Der nächste Schritt des *decision making* sieht vor, dass diese Daten von (unterschiedlich operierenden) Teams ausgewertet und diskutiert werden. Die Zusammensetzung dieser Teams kann sehr unterschiedlich sein; das Ziel ist aber immer, dass die Lehrkräfte einen möglichst neutralen Blick auf die Klasse und die einzelnen Schüler*innen erhalten. Auf dieser Grundlage soll eine Förderentscheidung (z. B. eine mögliche Interventionsplanung auf Stufe 2 oder 3) ermöglicht werden (Pirani-McGurl et al., 2022; Powers & Mandal, 2011). Ein wichtiger Aspekt dieser Auswertung sind häufig verschiedene Benchmarks⁷, die zeigen, wie ein Kind im Vergleich zu anderen lokalen oder nationalen Vergleichsgruppen derselben Klassenstufe abschneidet. Oft werden jene Daten auch visualisiert aufbereitet (Goodman-Scott et al., 2023; Schaffer, 2023). Die visuelle und meist farbliche Darstellung erleichtert im praktischen Kontext den Entscheidungsprozess und kann zudem in Gesprächen mit Eltern oder externen Therapeut*innen hinzugezogen werden. Sollten die Ergebnisse zeigen, dass weniger als 80 % ihrer Schüler*innen den Benchmark der Stufe 1 erreichen, können die Schulen diese Daten nutzen, um den Kernunterricht und Lehrplan der Stufe 1, sowohl akademisch als auch in sozio-emotionaler und verhaltensbezogener Hinsicht, anzupassen (Goodman-Scott et al., 2023).

Für die Schüler*innen, die sich bereits auf der zweiten und dritten Stufe befinden, gibt es weitere datengestützte Verfahrensweisen, um Förderentscheidungen anzupassen. Das Hauptziel der Datenerfassung ist es hier, den Fortschritt über die Zeit zu überprüfen, indem die

⁷ Ähnlich zu Grenzwerten und Normwerten: Festgelegte Werte, ab denen ein Verhalten oder eine Leistung als auffällig eingestuft wird.

aktuell erhobenen Daten mit den Ausgangsdaten verglichen werden und somit festgestellt werden kann, ob die Förderung effektiv ist oder nicht. Hierzu werden die zuvor erwähnten Tests und Screenings erneut angewendet oder durch weitere Verfahren ergänzt. Im US-amerikanischen Schulsystem wird häufig ein Zeitraum von sechs Wochen (oder sechs Datenpunkten) angegeben, für den die Verhaltensänderung analysiert wird (Schaffer, 2023). Hier greifen dann je nach MTSS unterschiedliche Regeln, die eine Anpassung der Intervention vorgeben. In jedem Fall sollte durch fortlaufendes *progress monitoring* der Fördererfolg der Maßnahmen überprüft und angepasst werden.

3.3.4 Multiprofessionelle Zusammenarbeit (problem-solving)

Ein weiterer Grundbestandteil von MTSS ist die multiprofessionelle Zusammenarbeit mehrerer Fachkräfte, die je nach MTSS-Form unterschiedlich gestaltet werden kann. In der US-amerikanischen Implementationsforschung wird diese Zusammenarbeit durch das bestehende Problem und die verschiedenen Lösungswege und Ansätze der Professionen begründet und auch oft mit *problem-solving* (PS) umschrieben. Diese PS-Modelle wurden ursprünglich in Anlehnung an *behavioral consultation models* entwickelt (Erchul & Martens, 2010) und basieren auf dem Grundsatz, dass multiprofessionelles Arbeiten und das Verbinden unterschiedlicher Professionsdimensionen einen Mehrwert für die Analyse und das interventionsbezogene pädagogische Handeln haben kann (Besa et al., 2023; Demmer & Hopmann, 2020; Halbheer & Kunz, 2011; Pool Maag, 2022; te Poel et al., 2023).

Der klassische Ablauf der PS-Teams sieht vor, dass sie zunächst die Stärken und Bedürfnisse der*s betreffenden Schüler*in basierend auf Daten analysieren und im Folgenden auf dieser Grundlage geeignete Interventionen ableiten. Das Team ist hieran anknüpfend auch für die Umsetzung und Implementation jener Interventionen im Rahmen des MTSS zuständig und verfolgt die Entwicklung dieser. Erweiterte PS-Modelle fügen neben der Problemidentifikation und der Problemanalyse, Problementwicklung und Implementation noch die (datenbasierte) Evaluation dem Prozess hinzu (Schaffer, 2023). Hier wird dementsprechend der multiprofessionelle Charakter gestufter Fördersysteme deutlich, welcher sich auch auf alle im schulischen Kontext eingesetzten Mitarbeiter*innen erstreckt. Es können neben Verwaltungsangestellten, Ergotherapeut*innen oder Logopäd*innen auch die Schulbusfahrer*innen oder Ganztagsbetreuungsangestellten miteinbezogen werden, um im Sinne eines PS-Prozesses vulnerable Bereiche zu identifizieren, Interventionen zu überlegen und gemeinsame pädagogische Strategien zu erarbeiten (Brown-Chidsey, 2016; Harrison & Thomas, 2014; Hierck & Weber, 2022). Insbesondere im Kontext der Beschreibung von

psychischen Störungen und Auffälligkeiten ist der Einbezug verschiedener Settings sinnvoll (Bradshaw et al., 2010; Opp, 1998; Sturmey, 2022).

Auf Seiten der Lehrkräfte kann diese multiprofessionelle Zusammenarbeit zudem für Entlastung rund um die Förderplanung sorgen und die alleinige Verantwortung des Fördererfolgs verringern (Demmer & Hopmann, 2020). Zusätzlich kann die hiermit verbundene kollegiale Zusammenarbeit innerhalb der Schule die allgemeine Schulqualität, das Schulklima, die Innovationsbereitschaft samt der allgemeinen Entwicklung der Lehrkräfte positiv beeinflussen (Halbheer & Kunz, 2011; te Poel et al., 2023). Im Kontext inklusiver Beschulung spielen diese Kooperationsprozesse demnach eine bedeutende Rolle (Pool Maag, 2022). Diese Vernetzung von interner schulischer Kooperation und gleichzeitiger externer schulischer (multiprofessioneller) Kooperation ist somit ein wesentlicher Bestandteil gestufter Förderung und findet sich in verschiedenen anderen Bereichen (z. B. *decision making, evidence based interventions, tiered support*) wieder. Die jeweilige Zielgruppe und das konkrete MTSS bestimmen hierbei die Qualität und die Ziele der multiprofessionellen Zusammenarbeit, die hauptsächlich dem Grundsatz eines effektiven PS unterliegen und im besten Fall positive Auswirkungen auf alle Beteiligten eines MTSS haben.

3.4 MTSS in der Grundschule im internationalen Vergleich

Nach diesen Erläuterungen zur Entstehung und der Funktionsweise von MTSS und ausgehend von den anfangs aufgestellten Leitfragen bietet sich ein genauerer Blick auf die derzeitige Forschungslage und den damit verbundenen internationalen Vergleich verschiedener MTSS an. Neben der Grundcharakteristik, den eingesetzten Interventionen, den adressierten Personenkreisen, der Wirkung und den Einsatzorten gestufter Fördersysteme weltweit, sollte hier zusätzlich auch die Frage nach funktionierenden Forschungszugängen im Vordergrund einer solchen Analyse stehen. Grund hierfür sind die multimethodalen Zugänge, welche die empirische Bildungsforschung in den letzten vier bis fünf Dekaden entwickelt hat, um Wirkprozesse auf systemischer und interventionsbezogener Ebene zu untersuchen (Renner et al., 2012; Suri & Clarke, 2009). In Verbindung damit steht, dass sich bestehende MTSS in der Vergangenheit dauerhaft transformiert und erweitert haben. Dies ging einher mit einer weltweiten Verbreitung und kulturellen Adaption wirksamer Fördersysteme. Im Zuge einer solchen systematischen Aufarbeitung sollten bestehende systematische Reviews und Metaanalysen ebenfalls in Betracht gezogen werden, um als Referenzpunkt für weitere

Analysen genutzt werden zu können. Im Kontext gestufter Förderung lassen sich somit diverse Beiträge mit unterschiedlichen Schwerpunkten identifizieren (Tabelle 1).

3.4.1 Systematische Überblicksarbeiten zum Thema MTSS

Das zuletzt veröffentlichte Scoping Review von Castillo et al. (2022) widmet sich dem Bereich der Professionalisierung (bzw. Fortbildung) im Kontext von MTSS und fasst 79 Studien mit unterschiedlichen Forschungszugängen zusammen. Das Ergebnis ist eine systematische Auflistung von Forschungsanliegen und Forschungszugängen sowie Aussagen über hinderliche und förderliche Faktoren rund um die Professionalisierung von Lehrkräften im Kontext von MTSS. Darüber hinaus untersucht die Forscher*innengruppe rund um Briesch et al. (2019), inwieweit US-amerikanische Bundesstaaten einheitliche Richtlinien für gestufte Fördersysteme (im Kontext Verhalten, *MTSS-B*) in das Bildungswesen integriert haben. Trotz der Wichtigkeit dieser Verfahrensweisen zeigt sich, dass nur etwa die Hälfte der Bundesstaaten solche grundlegenden Orientierungshilfen bereitstellt. Oft beinhalten diese bildungspolitischen Vorgaben Angaben zu den Arten der Interventionen, vernachlässigen jedoch Leitlinien zu den Verfahrensweisen der Datenerhebung und -auswertung (*decision making*). Die Ergebnisse weisen auf die Notwendigkeit hin, lokale Implementationsstrategien zu verbessern und zukünftige staatliche Richtlinien für MTSS-B zu stärken– auch im Bereich der Grundschulen.

Ein weiterer großer Teil der systematischen Überblicksarbeiten fokussiert die verschiedenen Interventionen auf den Ebenen der MTSS. So untersuchen Estrapala et al. (2020) die eingesetzten Interventionen auf der universalen Förderebene. Grasley-Boy et al. (2020) fassen darüber hinaus Interventionen zusammen, die in *Alternative Education Settings* eingesetzt werden. Diese Einrichtungen bieten einen speziell angepassten pädagogischen Rahmen, um auf die Bedürfnisse dieser Schüler*innengruppe einzugehen, die oft intensivere Verhaltensunterstützung benötigen. Zu den alternativen Bildungseinrichtungen können spezielle Schulen für Kinder mit Behinderungen, Wohnheimrichtungen oder Justizvollzugsanstalten zählen. Viele Schüler*innen mit externalisierendem Verhalten werden in den USA in solchen Bildungseinrichtungen beschult. Die Überblicksarbeit konnte anhand von 19 Studien zeigen, dass die Implementation von verhaltensverstärkenden Interventionen verschiedene positive Einflüsse haben kann. Insgesamt verringerte sich laut Aussage der Studie die Häufigkeit von Zwangs- und Isolationsmaßnahmen nach Umsetzung der Unterstützungsangebote auf universalen Präventionsebene. Je nach Art der Verhaltensauffälligkeit hatten die eingesetzten Interventionen jedoch unterschiedlich hohe Effekte. In Bezug auf High-Schools deuten die Ergebnisse von Estrapala et al. (2020) darauf

hin, dass gestufte Förderung das Schulklima und die Abschlussquoten verbessern kann. Ungeachtet dessen sind in beiden Schulformen wenige gestufte Ansätze implementiert, da Faktoren wie die Schulgröße und die Organisationskultur die Einführung und Umsetzung erschweren.

In ähnlicher Art und Weise fassen weitere Überblicksarbeiten die eingesetzten Interventionen auf der 2. Stufe zusammen. Die Beiträge legen nahe, dass Interventionen auf der selektiven Präventionsebene MTSS übergreifend wirksam sind, insbesondere mit Fokus auf den Bereichen Verhalten und Lesefähigkeiten (Mitchell et al., 2017; Wanzek et al., 2016). Die Studien weisen zudem auf die Bedeutung von geeigneten Screening-Methoden hin, um die richtigen Schüler*innen für diese Interventionen zu identifizieren, und betonen die Notwendigkeit einer treuen Umsetzung der Interventionen (Blair et al., 2021; Bruhn et al., 2013). Die Analysen von Blair et al. (2021) zeigen Effektstärken in Einzelfallstudien zwischen $TauU = 0.26$ und 0.98 . Im Bereich „Training sozialer Fertigkeiten“ ließen sich die größten Effekte nachweisen, während Check-in/Check-out, Gruppenkontingenzverfahren und Interventionspakete mittlere Effektgrößen aufwiesen. Moderator-Analysen zeigten darüber hinaus unterschiedliche Effektstärken für verschiedene Interventionstypen und durchführende Personen. Das Feld der selektiven Interventionsmaßnahmen lässt sich somit als breites Forschungsfeld mit diversen Analyseverfahren und Evaluationszugängen beschreiben. Die individuellen Entwicklungsverläufe der Schüler*innen stehen hier stark im Vordergrund und lassen bisher weniger stark beforschte Bereiche wie die Implementationsqualität und Wirksamkeit kombinierter Methoden erkennen.

Ein weiterer Fokus systematischer Überblicksarbeiten ist die Untersuchung der unterschiedlichen Personengruppen. Verschiedene Forscher*innengruppen haben dementsprechend Vorschulkinder oder Kinder mit Behinderungen und/oder sozial-emotionalen Auffälligkeiten in den Blick genommen, um zu analysieren, inwiefern gestufte Systeme auf diese besonderen Bedürfnisse oder Bedingungen abgestimmt sind: Shepley et al. (2020) fassten aufbauend auf einer vorherigen Analyse (Shepley & Grisham-Brown, 2019) Einzelfallstudien und ihre Wirksamkeit zusammen und kommen für Vorschulkinder zu dem Schluss, dass MTSS, insbesondere für die sozial-emotionale Entwicklung im frühen Kindesalter, vorteilhaft sein können, aber auch, dass aufgrund der Variabilität der Wirksamkeit in verschiedenen Entwicklungsbereichen eine sorgfältige Gestaltung und Implementation jener Systeme erforderlich ist. Einen anderen Schwerpunkt im Bereich von verhaltensunterstützenden Maßnahmen im Kontext gestufter Förderung legte die Forscher*innengruppe rund um McLaughlin et al. (2011). Der Fokus lag hier auf der Umsetzung der Maßnahmen in den

Familien, wobei ein positiver Einfluss (Reduktion) der Interventionen auf herausforderndes Verhalten junger Kinder festgestellt werden konnte. Eine wichtige Ergänzung in diesem Zusammenhang ist jedoch, dass 98 % der Kinder in den 18 zugrundeliegenden Studien eine diagnostizierte Form von Autismus oder pervasiver Entwicklungsstörung⁸ aufwiesen.

Tabelle 1

Überblick über bestehende systematische Literaturarbeiten zum Thema MTSS

Bereich	Studien
Stufe 1 Interventionen	<u>Allen-Meares et al. (2013); Estrapala et al. (2020); Grasley-Boy et al. (2020)</u>
Stufe 2 Interventionen	<u>Blair et al. (2021); Bruhn et al. (2013); Mitchell et al. (2017); Wanzek et al. (2016)</u>
MTSS in der Vorschule	<u>McLaughlin et al. (2011); Shepley und Grisham-Brown (2019); Shepley et al. (2020)</u>
High-School & weiterführende Schule	<u>Estrapala et al. (2020); Freeman et al. (2019); Guest (2011)</u>
Allgemein	<u>Charlton et al. (2020); Öğülmüş und Vuran (2016); Solomon et al. (2012); Lee und Gage (2020)</u>
Professionalisierung & Implementation	<u>Briesch et al. (2019); Castillo et al. (2022)</u>
Herausforderndes Schüler*in-Verhalten	<u>Walker et al. (2023); Simonsen et al. (2013)</u>

Anmerkung: Diese Zusammenstellung wurde durch systematische Suchen via Academia Search Ultimate im März 2024 erstellt.

Diverse weitere Überblicksarbeiten können zusätzlich zu diesen Ergebnissen Effekte von MTSS auf Bereiche wie die Verringerung von Problemverhalten und der Steigerung des Sozialverhaltens (Öğülmüş & Vuran, 2016), Auswirkungen auf das Schulklima (Charlton et al., 2020) und eine Verbesserung der schulischen Leistungen (Lee & Gage, 2020) nachweisen. Der

⁸ Nicht anderweitig spezifiziert

zentrale Fokus liegt in allen Studien auf grundlegenden Konzepten von MTSS, die sich über die Zeit fortlaufend weiterentwickelt haben. Allgemein kann MTSS daher im Grundschulbereich, aber auch darüber hinaus eine breite Wirksamkeit auf viele schulische Facetten nachgewiesen werden (Tabelle 1)

Wie bereits dargelegt, sind gestufte Fördersysteme aus verschiedenen bildungspolitischen Umständen heraus entstanden und treten heute in diversen Unterformen entsprechend ihrem vorrangigen Ziel auf. Wissenschaftlich und historisch lässt sich diese Entwicklung nachzeichnen und vorrangig auf drei Grundkonzepte gestufter Förderung zurückführen (Goodman-Scott et al., 2023; Schaffer, 2023; Shepley & Grisham-Brown, 2019; Sugai & Horner, 2019):

1. Response to Intervention (RTI)
2. Positive Behavior (Interventions and) Support (PB[I]S)
3. School-wide Positive Behavior (Interventions and) Support (SWPB[I]S)

Da diese Konzepte eine tragende Rolle in der Ausgestaltung und Entwicklung neuer Konzepte für MTSS in der Grundschule spielen und oftmals Referenzpunkt für Studienergebnisse, systematische Analysen und Evaluationskonzepte sind, werden sie im Folgenden kurz vorgestellt.

3.4.2 Response to Intervention (RTI)

Response to Intervention (RTI) ist eine Form gestufter Förderung, die zuerst in den USA in das Bildungswesen integriert wurde (Bradley et al., 2005; Fuchs & Fuchs, 2007; Grosche & Volpe, 2013). Zunächst lag der Fokus verstärkt auf der Prävention akademischer Problemlagen. Das heutige Verständnis umfasst von RTI umfasst jedoch auch die Prävention verhaltensbezogener Risiken (Grosche & Volpe, 2013; Hawken et al., 2008; Hubert & Grosche, 2012). Die zuvor dargelegten Kriterien gestufter Förderung (Kap. 3.3) lassen sich allesamt in RTI-Ansätzen wiederfinden (Gresham, 2016; Jimerson et al., 2016). RTI ist dementsprechend auch dreistufig aufgebaut und basiert auf der Früherkennung auffälliger Verhaltensweisen und akademischen Leistungen entlang eines strukturierten diagnostischen Prozesses. Trotz fehlender allgemeingültiger Definition umfasst die Förderung systemübergreifend zusammengefasst folgende Ebenen (Pullen et al., 2018):

- **Stufe 1** umfasst Förderung in Form von evidenzbasiertem Kernunterricht für alle Schüler*innen (Murawski & Hughes, 2009; Pullen et al., 2018).
- **Stufe 2** zeichnet sich durch zusätzliche Förderung in kleineren Gruppen aus für Schüler*innen, die auf Stufe 1 nicht profitiert haben. Dies kann z. B. mehrmalige Förderung in der Woche von ca. 20-40 Minuten umfassen (Grosche & Volpe, 2013; Lembke et al., 2010; Pullen et al., 2018).
- **Stufe 3** umfasst intensivere, oft mit Einzelförderung verbundene Interventionen, für Schüler*innen, die auch auf Stufe 2 nicht respondiert haben. Dies kann z. B. mehrmalige Förderung in nochmals kleinerer Gruppe mehrmals bis täglich in der Woche von ca. 40-60 Minuten umfassen (Grosche & Volpe, 2013).

In bestimmten schulischen Kontexten kann eine unzureichende Leistung auf der dritten Stufe zur Überweisung des Schülers bzw. der Schülerin auf eine Förderschule (*special education*) führen. Das primäre Ziel von RTI ist jedoch, allen Schüler*innen den Zugang zu evidenzbasiertem Unterricht zu gewährleisten, bevor eine sonderpädagogische Förderung in Erwägung gezogen wird (Bradley et al., 2005; Huber, 2013). Auch wenn RTI neben anderen gestuften Fördersystemen als Ansatz beschrieben wird, der auch Verhaltensförderung miteinschließt, ist die Entwicklung von RTI stark auf die Verbesserung der Lese- und Rechtschreibfähigkeiten bzw. grundlegenden akademischen Fähigkeiten zurückzuführen (Hawken et al., 2008). Aus diesem Grund werden RTI-Ansätze weiterhin mit der bloßen Förderung akademischer Kompetenzen in Verbindung gebracht, was der Komplexität heutiger Ansätze zumeist aber nicht entspricht.

3.4.3 Positive Behavior (Interventions and) Support (PB[I]S)

Einen stärkeren Fokus auf das in der Schule gezeigte Verhalten legt der *Positive Behavior (Interventions) and Support* (PBS & PBIS). Im Vordergrund steht hier die Förderung von positiven Verhaltensweisen und die Reduzierung von Problemverhalten wie Unterrichtsstörungen und aggressivem Verhalten (Cheney & Jewell, 2012; James et al., 2019). Der Einsatzort ist jedoch nicht immer auf die Schule festgelegt (Don Kincaid, Heather Peshak George, & Karen Childs, 2016). So gibt es auch PBIS-Systeme in Einrichtungen wie Strafvollzugsanstalten o. ä. (Scheuermann & Nelson, 2019). In der Schule basiert PBIS zumeist auf Daten, um die Wirksamkeit von Interventionen auf Schulebene, Klassenebene und Schüler*innenebene zu beurteilen und die Förderung gegebenenfalls anzupassen (Arco, 2003). Die Stufen des PBIS bzw. PBS gliedern sich häufig wie folgt, auch wenn hier Abweichungen

der expliziten Ausgestaltung der Gemeinden, Schulen oder weiteren (pädagogischen) Institutionen möglich sind:

- **Stufe 1** umfasst universale Maßnahmen für alle Schüler*innen in einer Schule. Sie beinhaltet das Festlegen von sozialen Erwartungen (z. B. Respekt, Verantwortung, Sicherheit), das Unterrichten dieser Erwartungen und das positive Verstärken von erwartetem Verhalten (Estrapala et al., 2020).
- **Stufe 2** richtet sich an Schüler*innen, die Schwierigkeiten haben, die sozialen Erwartungen der Stufe 1 zu erfüllen. Sie umfasst zielgerichtete Interventionen für diese Schüler*innen, wie z. B. Mentoring, erhöhte Überwachung, tägliche Fortschrittsberichte und soziale Kompetenztrainings (Hoyle et al., 2011; Rodriguez et al., 2016).
- **Stufe 3** ist für die Schüler*innen (5 % oder weniger), die individualisierte Interventionen benötigen. Sie sieht z. B. die Durchführung einer funktionalen Verhaltensanalyse sowie die Umsetzung eines individuellen Förderplans vor und kann auch andere Ansätze wie spezielle therapeutische Bildungsprogramme beinhalten (Debnam et al., 2011; Voorhees et al., 2013).

PBIS hat sich als effektiv erwiesen, um problematische Verhaltensweisen zu reduzieren sowie soziale und akademische Fähigkeiten zu verbessern (Gage et al., 2020; Gage, Whitford, & Katsiyannis, 2018; Donald Kincaid et al., 2016; Pas et al., 2019). Die Umsetzung von PBIS erfordert von Schulen jedoch häufig, ihre disziplinarischen Ansätze neu auszurichten und sie stattdessen proaktiver, instruktiver und verstärkender zu gestalten. Auf diese Weise können die Schüler*innen die Schule als einen positiven Ort erleben.

3.4.4 School-wide Positive Behavior (Interventions and) Support (SWPB[I]S)

Den dritten MTSS-Ansatz stellt der *School-wide Positive Behavior Support* (SWPBS, oder auch *School-wide Positive Behavior Interventions and Support*, SWPBIS) dar. Dabei handelt es sich um eine auf die Schule bezogene Form von PBIS (Cheney & Jewell, 2012; Walker et al., 2016) und bildet ebenfalls ein MTSS ab, das auf die Prävention von externalisierendem Verhalten von Kindern und Jugendlichen in der Schule ausgerichtet ist (Lee & Gage, 2020). Ein wesentlicher Bestandteil ist auch hier das schulbasierte, datengesteuerte Vorgehen bezüglich der Identifikation von Schüler*innen mit Bedarf an intensiveren pädagogischen Maßnahmen (Noltmeyer et al., 2019). Zusätzlich kann auch die akademische Förderung im

Fokus stehen (Gage et al., 2013; Nelen et al., 2021). Auch hier finden sich zumeist drei Stufen, auf denen die Förderung der Schüler*innen stattfindet (Gage, Lee, et al., 2018):

- **Stufe 1** bietet universelle Unterstützung für alle Kinder und beinhaltet z. B. pädagogische Maßnahmen wie das Aufstellen schulweiter Regeln, Etablieren von Schulroutinen, Classroom Management und Anerkennungssysteme für angemessenes Verhalten (Timothy et al., 2014).
- **Stufe 2** fokussiert vermehrt kleinere Schüler*innengruppen mit Strategien wie *Check and Connect* (Karhu et al., 2020), Sozialkompetenztrainings (Clifford et al., 2020) oder Maßnahmen wie die *Daily Behavior Report Cards* (Volpe et al., 2013).
- **Stufe 3** beinhaltet verhaltensbezogene Interventionspläne, die durch funktionale Verhaltensanalysen erstellt werden, um den intensiven pädagogischen Bedarf individuell zu adressieren (Lower et al., 2016).

SWPBIS konzentriert sich darüber hinaus global auf die Schaffung eines sicheren, gerechten und integrativen schulischen Umfelds und hat somit positive Auswirkungen auf die schulischen Leistungen der Schüler*innen sowie die Arbeitszufriedenheit des Schulpersonals und der Lehrkräfte (Walker et al., 2023). Derzeit haben ca. 25.000 Schulen weltweit SWPBS implementiert, sodass diesem eine hohe Verbreitung und Adaption nachgewiesen werden kann (Lee & Gage, 2020).

3.4.5 Verschiedene MTSS-Formen und Forschungszugänge

Zusammenfassend sollte deutlich werden, dass die zuvor vorgestellten MTSS-Formen große Überschneidungen, aber auch einige Unterschiede aufweisen. PB(I)S wurde z. B. in den 1980er- Jahren als Reaktion auf den Bedarf an Verhaltensstrategien im Kontext des Umgangs mit Menschen mit schweren Behinderungen entwickelt. Der Ansatz basiert somit auf den Prinzipien der angewandten Verhaltensanalyse (Solomon et al., 2012). SWPBS hingegen ist ein präventiver und interventionsorientierter gestufter Förderansatz, der kontinuierlich durch Forschung und evidenzbasierte Praxis weiterentwickelt wird. SWPBS ist somit nicht als *ein* bestimmtes Programm oder *ein* Interventionskanon zu verstehen, sondern vielmehr als ein allgemeiner Ansatz zur Verbesserung der Integration und Implementation sowie nachhaltiger Anwendung verschiedener evidenzbasierter Praktiken im Zusammenhang mit Verhaltensmanagement, Classroom Management und schulischen Disziplinsystemen. SWPBS unterscheidet sich von individuell angewandtem PBS durch seinen stark präventiven Fokus,

einen lehrorientierten Ansatz, die Anwendung evidenzbasierter Verhaltenspraktiken und einen systemischen Ansatz, der bestehende Schulressourcen und -strukturen nutzt.

Diese Transformation bzw. Weiterentwicklung von PBS zu SWPBS kann sinnbildlich für die Flexibilität und Adaptionfähigkeit gestufter Fördersysteme gesehen werden (Kittelman et al., 2019; Voorhees et al., 2013). Für die erfolgreiche Implementation gestufter Förderung in Deutschland erscheint es somit sinnvoll, sich zunächst verschiedene Formen von MTSS im internationalen Vergleich anzuschauen und diese hinsichtlich des zugrundeliegenden Bildungssystems miteinander zu vergleichen. Auch sollte hierbei ein Vergleich der Wirksamkeit und der Studiendesigns samt Fragestellungen betrachtet werden. Die Fülle an Studien, die weltweit entstanden sind, können zusammengenommen eine Möglichkeit bieten, erste Ideen für die Entwicklung eines deutschen MTSS abzuleiten. Auffällig erscheint ergänzend hierzu, dass für die Beforschung von gestufter Förderung häufig drei methodische Zugänge gewählt werden (Gage, Whitford, & Katsiyannis, 2018; Lee & Gage, 2020; J. Nitz et al., 2023):

- 1) Quasi-Experimentelle Designs & Randomisierte Kontrollgruppen-Designs
- 2) Einzelfallstudien
- 3) Qualitative Interviewstudien

Dies erscheint vor dem Hintergrund, dass unterschiedliche MTSS-Formen verschiedene Hauptziele verfolgen, zunächst sinnvoll. Je nach inhaltlichem Schwerpunkt unterscheidet sich somit auch die zugrundeliegende Forschungsmethodik. Es besteht also häufig ein Zusammenhang zwischen eingesetzter Forschungsmethodik und inhaltlichem Schwerpunkt des MTSS: Aufgrund der Fülle der verschiedenen Systeme und der weitreichenden globalen Verbreitung ist es nahezu unmöglich, alle implementierten Systeme in der Gänze zu erfassen. Bestehende Überblicksarbeiten beschränken sich daher zumeist auf einen Teilbereich (z. B. den Fokus auf angewandte Interventionen; Tabelle 1). Um vor diesem Hintergrund ein geeignetes MTSS samt geeignetem Forschungszugang zu entwickeln oder zu bewerten, müssen also der Einsatzort und die damit verbundenen Rahmenbedingungen näher betrachtet werden.

3.5 Pädagogischer Unterstützungsbedarf in der Grundschule in Deutschland

Grundschulen in Deutschland sind für die meisten Kinder der erste gemeinsame Lernort. Im Jahr 2022 besuchten laut Kultusministerkonferenz 3.022.187 Schüler*innen in Deutschland

eine Grundschule⁹ (KMK, 2024). Förderschulen wurden von 337.663 Schüler*innen besucht. Im Gegensatz zu weiterführenden Schulen, die Schüler*innen basierend auf ihren Leistungen besuchen können (z. B. Hauptschule, Realschule, Gymnasium), lernen nahezu alle Kinder in Deutschland demnach zunächst gemeinsam. Ein potenzielles Fördersystem würde einen Großteil der schulpflichtigen Bevölkerung dieser Altersspanne (ca. 6-10 Jahre) erreichen. Tritt in dieser Altersstufe eine psychische Auffälligkeit auf, so ist die frühe Intervention und pädagogische Begegnung dieser in Form von einer systematischen und zielgeleiteten Förderung oftmals effektiver, als wenn sich psychische Auffälligkeiten und Störungen manifestieren und ihnen erst zu einem späteren Zeitpunkt therapeutisch und pädagogisch begegnet werden (Bachmann et al., 2010; Stoltz et al., 2012). Aus diesem Grund kommt der Grundschule eine wichtige Bedeutung im Kontext der Prävention externalisierenden Verhaltens zu (Reiss et al., 2023).

Diese präventive Arbeit wird jedoch erschwert, da das inklusive Schulsystem insbesondere in der Grundschule in vielerlei Hinsicht stark belastet ist (Lütje-Klose et al., 2014; Skorsetz et al., 2020). Lehrkräfte beklagen auf struktureller Ebene z. B. personelle und zeitliche Mängel sowie eine unzureichende Ausstattung (Baumgardt, 2016; Prengel, 2013). Zusätzlich zeigen Studien, dass sich auch in der Grundschule Kinder mit psychischer Auffälligkeiten (vor allem im Bereich externalisierenden Verhaltens) befinden und Lehrkräften Unterstützung benötigen, mit diesen Verhaltensweisen adäquat umzugehen (Klipker et al., 2018; Polanczyk et al., 2015; Scherzinger & Wettstein, 2023). Ergänzend sind pädagogische Systeme und Rahmungen, ähnlich der der Förderschulen, häufig in der Grundschule nicht implementiert (Prengel, 2013). Hierunter fällt z. B. der systematische Einbezug weiterer (externer) Professionen wie Psycholog*innen, Therapeut*innen oder beratenden Personen aus Jugendamt und Jugendhilfe. Aktuelle Forschungen deuten zudem darauf hin, dass im Grundschulbereich die Heterogenität der Lerngruppen oft unzureichend berücksichtigt wird (Diehm, 2020; Langer & Schnebel, 2020; Skorsetz et al., 2020). Dies unterstreicht die die Notwendigkeit individueller Förderkonzepte oder Systeme (Ernst, 2023; Liebers et al., 2023; Röhner et al., 2020). Auch zur Unterstützung der Lehrkräfte im Umgang mit psychisch belasteten Schüler*innen werden qualitativ hochwertige Fortbildungen als essenziell erachtet (Baumgardt, 2016; Hanisch et al., 2023), da Studien bestätigen, dass insbesondere Kinder mit psychischen und sozialen Herausforderungen auf Lehrkraftseite zu Stress und einem erhöhten Burnout-Risiko führen

⁹ 3.131.362 insgesamt: Die Übrigen besuchten integrierte Gesamtschulen (78.798) oder freie Walddorfschulen (30.385).

kann ([Agyapong et al., 2022](#); [Lehr, 2014](#); [Li et al., 2022](#)). Diese Belastung beeinträchtigt die Unterrichtsqualität negativ, was sich wiederum negativ auf das erfolgreiche Lernen der Schüler*innen auswirken kann ([Agyapong et al., 2022](#); [Evers et al., 2016](#)). Zusammenfassend heben die verschiedenen Befunde die Wichtigkeit hervor, Lehrkräfte darin zu unterstützen, mit herausforderndem Verhalten umzugehen und gleichzeitig ein Lernumfeld zu schaffen, das (verhaltensauffälligen) Schüler*innen die notwendige Förderung entgegenbringt ([Aldabbagh et al., 2022](#); [Breitenstein et al., 2009](#); [Docherty et al., 2023](#); [Müller, 2021](#)).

3.6 MTSS in der Grundschule in Deutschland

Nachdem die Bedarfe, die bildungspolitische Motivation und die Relevanz der Entwicklung eines MTSS für den deutschen Grundschulbereich abgeleitet und dargelegt wurden, ergibt sich zusätzlich die Frage nach bereits bestehenden Systemen in Deutschland. Auch wenn diese Teil eines solchen Literaturüberblicks ([Beitrag 1](#)) sein werden, da sie von der Forschungsfrage zum großen Teil eingebunden und durch eine systematische Analyse identifiziert werden, kann eine zusätzliche, vorgeschaltete Recherche bezüglich deutschen Forschungsprojekten und Fördersystemen gewinnbringend für die Auseinandersetzung mit dem Thema sein. Aus diesem Grund werden im Folgenden bestehende deutschen MTSS-Ansätze bzw. Projekte kurz vorgestellt¹⁰.

3.6.1 Das Rügener Inklusionsmodell (RIM)

Das *Rügener Inklusionsmodell* (RIM) ist ein gestuftes Förderkonzept, das auf der Insel Rügen implementiert wurde und auf dem RTI-Ansatz basiert ([Mahlau et al., 2011](#)). Es zielt darauf ab, eine inklusive Schulstruktur zu schaffen, die es ermöglicht, Kinder mit Förderbedarf, insbesondere in den Bereichen Lernen, sprachliche Entwicklung und emotional-soziale Entwicklung, effektiv zu unterstützen. Eine der Herausforderungen bei der Entwicklung des MTSS war die Anpassung an die traditionellen, stark differenzierten Förderstrukturen in Deutschland, die wenig Raum für eine systematische Zusammenarbeit zwischen allgemeiner Pädagogik und Sonderpädagogik bieten. Zudem mangelt es im deutschsprachigen Raum an einer ausreichenden Auswahl an evidenzbasierten Lern- und Förderkonzepten sowie an Verfahren zur Lernfortschrittsdiagnostik. Ein wesentlicher Schwerpunkt des RIM liegt daher in

¹⁰ Basierend auf einer Recherche über EBSCO-Host und FIS-Bildung im Februar 2024 mit eingängigen Schlagwörtern: vgl. [Beitrag 1](#)

der Etablierung einer Infrastruktur für die Kommunikation und Kooperation zwischen der allgemeinen Schule und den Förderschulen. Ebenso wichtig ist die Entwicklung von systematischen Verfahren für das Monitoring des Lernfortschritts der Schüler*innen. Zu diesem Zweck wurden unterschiedliche Materialien für verschiedene schulische Bereiche entwickelt, die über die Website www.lernfortschrittsdokumentation-mv.de zugänglich sind. Diese Plattform dient auch der internetgestützten Dokumentation des Lernfortschritts – einem zentralen Element des RIM. Ein weiterer wichtiger Aspekt der Implementation des RIM in den Rügener Grundschulen ist ein umfassendes Fortbildungsprogramm für Lehrkräfte und Schulleitungen. In über 300 Fortbildungsstunden wurden die Teilnehmenden nicht nur in allgemeinen Themen wie den Bedingungsfaktoren der Schulleistung geschult, sondern auch in der Anwendung der im RIM verankerten Unterrichts-, Förder- und Trainingsmaterialien sowie diagnostischer Verfahren in den Bereichen Mathematik, Deutsch, Sprache sowie emotionale und soziale Entwicklung.

Die Ergebnisse erster Studien zeigen, dass Schüler*innen in der Kontrollgruppe in Mathematik leicht besser abschnitten als die RIM-Gruppe, während im Lesen keine signifikanten Unterschiede zwischen den Gruppen festgestellt wurden (Voß et al., 2014). Die praktische Umsetzung des RIM stieß auf Herausforderungen, insbesondere bei der Einhaltung der theoretischen Konzeption, und erforderte Anpassungen im Lehrplan sowie zusätzliche Lehrkraftfortbildungen. Trotz dieser Schwierigkeiten führte das RIM zu wichtigen Veränderungen in der pädagogischen Praxis und stellt einen Fortschritt in der inklusiven Bildung dar. Positive Wirkungen des Ansatzes zeigen sich nur hinsichtlich bestimmter Faktoren, darunter Problemverhalten, prosoziales Verhalten, das Gefühl des Angenommenseins, des Klassenklimas, des Selbstkonzepts und der sozialen Integration (Blumenthal & Voß, 2016). Diese positiven Effekte treten vor allem bei Kindern mit leichten schulischen Minderleistungen auf. Der stärkste Effekt wurde für das selbstwahrgenommene Klassenklima bei Kindern mit deutlichen schulischen Minderleistungen beobachtet ($d = 0.69$).

Die geringen Effekte des RIM könnten teilweise auf die langfristigen und komplexen schulischen Umstrukturierungen zurückzuführen sein. Während bei Kindern der RIM-Gruppe vereinzelt positive Effekte, insbesondere in emotional-sozialen Bereichen, beobachtet wurden, ist eine eindeutige Kausalinterpretation aufgrund der vorhandenen Querschnittsdaten und weiterer einschränkender Faktoren wie soziokulturelle und leistungsbezogene regionale Unterschiede nicht möglich (Blumenthal & Voß, 2016). Auch die unklare Umsetzung der Maßnahmen im Klassenkontext und die potenzielle Verzerrung durch sozial erwünschtes Antwortverhalten in den Fragebögen schränken die Aussagekraft der Ergebnisse ein. Kinder

ohne oder mit leichten schulischen Minderleistungen zeigen durchschnittliche bis überdurchschnittliche emotionale und soziale Werte, während Kinder mit deutlichen Minderleistungen schlechtere Ergebnisse, insbesondere in schulmotivationalen Aspekten und sozialer Integration, aufweisen. Die Ergebnisse deuten darauf hin, dass das RIM keine negativen Auswirkungen auf Schüler*innen ohne Minderleistungen hat und dass weiterführende Forschung erforderlich ist, um die festgestellten Herausforderungen zu ergründen und die Wirksamkeit des Programms umfassend zu bewerten.

3.6.2 Partizipation in einem Response-to-Intervention-Modell (PARTI)

Das Projekt *Partizipation in einem Response-to-Intervention-Modell für den Förderschwerpunkt Emotionale und soziale Entwicklung* (PARTI) konzentriert sich auf die Unterstützung von Grundschullehrkräften bei der Inklusion von Schüler*innen mit externalisierenden Verhaltensproblemen ([Bartling et al., 2021](#)). Das Projekt ist zudem eine Reaktion auf Kritik an separierenden Fördermaßnahmen und der Trennung der Zuständigkeiten zwischen verschiedenen Berufsgruppen. Im PARTI-Projekt wurden daher Förder- und Diagnostikmethoden entwickelt, die im regulären Unterricht durchgeführt werden, um so die Verantwortungsaufteilung zwischen verschiedenen Professionen vermeiden. Das Fortbildungskonzept von PARTI wurde in elf Grundschulen implementiert und evaluiert, wobei die Lehrkräfte die methodische Umsetzung, die Relevanz der Inhalte und die Umsetzung im Schulalltag positiv bewerteten. Es wurde allerdings nur die unmittelbare Reaktion der Teilnehmenden erfasst; weitere Evaluationsebenen wie der tatsächliche Einfluss auf die Praxis und die Effekte auf Schüler*innen sind noch zu untersuchen.

3.6.3 Starke Schulen stärken Kinder

Ein weiterer Ansatz aus dem Märkischen Kreis ist *Starke Schulen stärken Kinder* (SSsK), der ebenfalls auf dem Prinzip des RTI basiert ([Bellmer et al., 2020](#)). Dieser Ansatz zielt darauf ab, Schulen systemische Unterstützungsmöglichkeiten entlang gestufter Fördersystemprinzipien zu geben, indem Schüler*innen bei Lernschwierigkeiten frühzeitig in den Blick genommen werden. Die Implementation des Ansatzes variiert von Schule zu Schule stark. Laut Aussage der Autor*innen zeigt die Umsetzung aber eine positive Entwicklung: Lehrkräfte und Schulen zeigen eine hohe Bereitschaft zur Mitarbeit und persönlichen Weiterentwicklung. Fortbildungsinhalte werden aktiv in den Schulalltag integriert, was zu einer transparenteren Aufgabenverteilung und zu einem fachlichen Erkenntnisgewinn führt. Es gibt eine zunehmende Tendenz zur Vernetzung und Zusammenarbeit zwischen den Schulen. Befragungen ergaben

außerdem, dass Lehrkräfte in diesem Projekt generell eine positive Haltung gegenüber standardisierten Verhaltensbeobachtungen für alle Kinder auf der ersten Förderebene zeigten. Viele Lehrkräfte sind auch bereit, Screenings für alle Kinder ihrer Klasse durchzuführen, trotz des damit verbundenen hohen Arbeitsaufwands. Dies deutet darauf hin, dass sie den Nutzen und Gewinn einer fundierten präventiven Förderung erkennen. Zudem zeigt sich eine hohe Bereitschaft, die eigene diagnostische Praxis zu erweitern. Insgesamt bestätigen die Ergebnisse, dass das zentrale Anliegen des Projekts bei den Lehrkräften auf Zustimmung stößt und viele bereit sind, sich im Bereich der Diagnostik weiterzuentwickeln.

3.6.4 Offene Fragen & Ausblick

Zusammenfassend lässt sich erkennen, dass die Forschungslandschaft zum Thema MTSS in Deutschland bereits in Teilen erste Ergebnisse und Ansätze vorweisen kann (z. B. [Bartling et al., 2021](#); [Mahlau et al., 2014](#); [Mahlau et al., 2016](#); [Voß et al., 2016](#)). Verschiedene Konzepte sind in den vergangenen ca. 15 Jahren entstanden, die jedoch nicht umfassend implementiert und evaluiert sind. Insbesondere fällt auf, dass alle drei Ansätze (RIM, PARTI, SSsK) auf dem RTI-Ansatz aufbauen. Jedes Projekt legt einen leicht veränderten Schwerpunkt auf die Ausbildung und Unterstützung von Lehrkräften durch unterschiedliche Fortbildungsprogramme, um die jeweiligen Ansätze erfolgreich zu implementieren. In allen drei Projekten zeigt sich eine allgemeine Bereitschaft der Lehrkräfte, sich mit neuen diagnostischen Verfahren und Fördermethoden auseinanderzusetzen und diese in den Schulalltag zu integrieren. Anschließend hieran ergeben sich jedoch offene Fragen, die für zukünftige Forschung mitbedacht und die auch in eine weitere internationale Sondierung der bestehenden Studien miteinbezogen werden sollten.

Zunächst erscheint relevant, wie die bestehenden Herausforderungen, die sich im RIM gezeigt haben (die Anpassung an differenzierte Förderstrukturen und die Entwicklung von Lernfortschrittsdiagnostik), effektiver adressiert werden können. Mit anderen Worten: Wie müssen Fortbildungskonzepte geplant und umgesetzt werden, sodass sie nicht zu einer Mehrbelastung der Lehrkräfte führen? Zusätzlich erscheint in diesem Zusammenhang die Frage, inwieweit die positiven Effekte des RIM, insbesondere in emotional-sozialen Bereichen, weiter gestärkt und auf andere Bereiche ausgedehnt werden können. Welche Veränderungen könnten entlang der vorliegenden Förderansätzen vorgenommen werden, um die Effektivität dieser zu erhöhen? Um die Effektivität und langfristigen Auswirkungen der vorgestellten pädagogischen Programme in der Grundschule in Deutschland vollständig zu verstehen, ist somit weiterführende Forschung unerlässlich. Diese Forschung sollte sich nicht nur auf die

kurzfristigen Effekte konzentrieren, sondern auch darauf, wie sich diese Programme langfristig auf die Schüler*innen auswirken.

Nicht zuletzt ist es von großer Bedeutung, alle beteiligten Stakeholder (Lehrkräfte, Schulleitungen, Eltern und Schüler*innen) in den Evaluations- und Entwicklungsprozess dieser Ansätze einzubeziehen. Durch ihre aktive Teilnahme und ihr Feedback können die MTSS besser auf die Bedürfnisse und Anforderungen aller Beteiligten abgestimmt werden, was wiederum deren Effektivität und Akzeptanz in der Schulgemeinschaft erhöhen würde (Fox et al., 2021). Eine solche ganzheitliche Herangehensweise ist entscheidend, um die Bildungslandschaft nachhaltig multiprofessionell auszurichten und alle Schüler*innen bestmöglich zu unterstützen.

Zusammengefasst erscheint es somit notwendig, einzelne Teilbereiche gestufter Fördersysteme in Deutschland kleinschrittig, samt ihren Wirkweisen und Effekte, stärker in den Fokus zu rücken und diese in Form qualitativer und quantitativer Studien genauer zu erheben. Ergänzend gilt es, Herausforderungen und Gelingensbedingungen von MTSS in Deutschland zu erforschen, sodass Desiderate, wie die hier aufgeführten, zuverlässig im Kontext der Neukonzeption von MTSS beforscht und geschlossen werden können. Darüber hinaus ergibt sich der besondere Bezug zu verhaltensbezogenen Maßnahmen nicht nur aus den in diesem Kapitel präsentierten Ergebnissen, sondern auch aus den zuvor vorgestellten bildungspolitischen Herausforderungen im Zusammenhang mit der Prävalenz und Persistenz externalisierenden Verhaltens im Kindes- und Jugendalter.

3.7 Ein systematischer Literaturüberblick über MTSS mit Fokus auf verhaltensbezogene Maßnahmen (Beitrag 1)¹¹

Aufbauend auf den geschilderten Grundlagen, Bedarfen und den erprobten deutschen Fördersystemen samt den damit verbundenen Forschungsdesideraten ergab sich die Leitfrage nach international implementierten und MTSS-Ansätzen mit Fokus auf verhaltensmodifizierenden Maßnahmen. Zusätzlich zu einer Neuentwicklung und späteren Adaption eines deutschen MTSS in der Grundschule vermag ein systematischer Vergleich bestehender MTSS-Systeme eine reflexive und kritische Funktion in der Evaluation jener Ansätze einzunehmen.

¹¹ Beitrag 1: Nitz, J., Brack, F., Hertel, S., Krull, J., Stephan, H., Hennemann, T., & Hanisch, C. (2023). Multi-tiered systems of support with focus on behavioral modification in elementary schools: A systematic review. *Heliyon*, 9(6). <https://doi.org/10.1016/j.heliyon.2023.e17506>

3.7.1 Methodik

Methodisch orientiert sich Beitrag 1 demnach an den gängigen PRISMA-Vorgaben¹² zur Erstellung systematischer Überblicksarbeiten für systematische Reviews und Meta-Analysen (Page et al., 2021). Hierzu gehört die Entwicklung einer Suchmatrix mit den vordergründig relevanten Begriffen, das systematische Sichten der Beiträge in Abstract- und Volltextanalyse sowie die Auswertung und Zusammenstellung der Ergebnisse in Form verschiedener Tabellen und Übersichten. Die übergeordneten Fragestellungen des Beitrags lauten in der deutschen Übersetzung:

- 1) Was sind die Charakteristika von MTSS im Grundschulbereich im internationalen Vergleich?
 - a. Welche verschiedenen Formen von MTSS wurden bisher international evaluiert?
 - b. Welche verhaltensbezogenen Interventionen werden auf den unterschiedlichen Stufen eingesetzt?
 - c. Welche Personengruppen sind in die Ausgestaltung des MTSS involviert?

- 2) Was sind die Charakteristika von empirisch-quantitativer Forschung bezüglich MTSS in der Grundschule im internationalen Vergleich?
 - a. Was ist der derzeitige Stand internationaler Forschung im Kontext von MTSS im Bereich der Grundschule?
 - b. Welche Variablen werden erhoben, um die Effektivität oder Implementation von MTSS zu erforschen?
 - c. Welche Effekte werden für verschiedene MTSS berichtet?

Die eingesetzte Suchmatrix, die für die Suchmaschine EBSCO Host¹³ verwendet wurde, lautete:

SU tier OR AB (“MTSS” OR “RTI” OR “PBIS” “PBS” OR “SWPBS” “SWPBIS” OR “response to intervention*” OR “positive behavio*” OR school-wide OR “integrated approach”) AND AB (*school* OR *grader* OR “primary” OR pupil* OR child* OR learner OR kid* OR student* OR educat* OR elementary) NOT AB (college OR “high school”)*

¹² Preferred Reporting Items for Systematic reviews and Meta-Analyses

¹³ <https://search.ebscohost.com>

Durchsucht wurden folgende Datenbanken: *American Psychological Association (APA) PsycArticles*, *APA PsycInfo*, *ERIC* und *MEDLINE* sowie die *FIS-Bildung* mit einer entsprechenden deutschen Suchmatrix. Ergänzend wurden bestehende Reviews und Metaanalysen in die Suche miteinbezogen. Schlussendlich wurden 8041 Artikel mit einem Projektteam von vier Mitarbeiter*innen mit einer Interrater-Reliabilität von $k = 0,64$ (Übereinstimmung von 91,3 %) codiert. Letztlich waren 40 Studien das Ergebnis für die Volltextanalyse, die im Folgenden die Grundlage der Ergebnisse darstellen. Die Analysen wurden mit den Programmen *Covidence* und *Rayyan* durchgeführt (Babineau, 2014; Couban, 2016; Kellermeyer et al., 2018).

3.7.2 Ergebnisse

Die Ergebnisse des systematischen Reviews mit Bezug zur Grundschule samt verhaltensmodifizierenden Maßnahmen zeigen, dass die Implementation von MTSS wirksam sein kann hinsichtlich einer Verbesserung positiver Schulerfahrungen und einer Verringerung negativer Bildungseinflüsse. Insbesondere im Kontext von störendem Verhalten und emotionalen sowie verhaltensbezogenen Störungen erweisen sie sich als erfolgreich.

3.7.2.1. Weltweite Verbreitung

Ein Großteil der untersuchten Studien (72,5 %), haben ihren Ursprung in den Vereinigten Staaten, was darauf hindeutet, dass in den USA eine besonders aktive und umfangreiche Forschungsgemeinschaft zu diesem Thema existiert (siehe z.B. auch Lee & Gage, 2020). Andere Länder, die identifiziert wurden und in denen ebenfalls Forschung zu MTSS betrieben wird, sind Australien, Kanada, Finnland, Deutschland, Neuseeland und Taiwan (z.B. Vainikainen et al., 2017; Wu et al., 2019; Yeung et al., 2009). Aus jedem dieser Länder stammt jeweils eine Wirksamkeitsstudie im Kontext MTSS mit Bezug zur Grundschule. Eine Ausnahme stellt Norwegen dar, wo im Vergleich eine höhere Forschungsaktivität zu verzeichnen ist: Insgesamt entstanden hier fünf Studien, was 12,5 % der gesamten untersuchten Studien entspricht (z.B. Sørli & Ogden, 2015; Sørli & Ogden, 2007). Diese globale Verteilung der Forschungsarbeiten spiegelt die internationale Relevanz und das Interesse an MTSS in verschiedenen Bildungssystemen wider.

3.7.2.2 Evidenzbasierte Interventionen und involvierte Personen

Nicht alle Studien machen präzise Angaben darüber, welche Interventionen im Rahmen der Implementation Anwendung fanden. Neun Studien gaben lediglich an, dass sie sich grundlegend an dem Konzept der gestuften Förderung orientierten (Bradshaw & Pas, 2011; Bradshaw et al., 2012; Debnam et al., 2011; R. A. Ervin et al., 2006; Feuerborn & Tyre, 2012; Lane et al., 2008; Savage et al., 2011; Weiland et al., 2014; Yeung et al., 2009). Aus den übrigen 31 Studien ließen sich für die verschiedenen Stufen der untersuchten MTSS-Formen bestimmte Interventionskategorien ableiten: Die Interventionen der ersten Stufe umfassen schulweite Regeln und Trainings sowie direkte Maßnahmen für Schüler*innen und Lehrer*innen, um eine grundlegende Struktur zu schaffen. Die Interventionen der zweiten Stufe beinhalten spezifischere Trainings, Kleingruppen- und Einzelinterventionen, Beratungsangebote sowie Unterstützung für Personal und Eltern. Die Interventionen auf der dritten Förderstufe waren intensivere, individuelle Maßnahmen, einschließlich konfliktbezogener und familiärer Interventionen. Die adressierten Personengruppen in den verschiedenen Interventionsebenen waren Schüler*innen, Eltern, Lehrer*innen, angehende Lehrer*innen, (weiteres) Schulpersonal, Psycholog*innen und Sozialarbeiter*innen an Schulen. In 28 % der Studien wurde angegeben, ganzheitliche und multimodale Förderung umzusetzen. Allumfassend scheinen Personen außerhalb des Schulsystems jedoch seltener in die Ansätze und Erhebungen integriert zu sein als Personen innerhalb der Schule.

3.7.2.3 Forschungsüberblick und Effekte gestufter Förderung

Die meisten Studien (65 %) waren in einem quasi-experimentellen Design (QED) ohne beschriebene Randomisierung angelegt (z. B. Albrecht et al., 2015; Debnam et al., 2011; Martella et al., 2010). Sechs Studien (15 %) waren Einzelfallstudien mit Verhaltensbeobachtungen (Franzen & Kamps, 2008; MacLeod et al., 2016; Nelson et al., 2018; Pearce, 2009; Sanetti & Collier-Meek, 2015; Wu et al., 2019), drei Studien (7,5 %) beschrieben Randomized Controlled Trials (RCT) (Algozzine & Algozzine, 2007; Bradshaw et al., 2012; Vainikainen et al., 2017) und fünf Studien (12,5 %) wurden in einem Mixed-Methods-Ansatz durchgeführt (Goodman-Scott, 2013; Greulich et al., 2014; Weiland et al., 2014; Wu et al., 2019). Zudem zeichnen sich viele Erhebungen durch große Stichproben aus und nutzen meist Daten aus Schulregistern (ODRs). Detaillierte deskriptive Daten und Inferenzstatistiken fehlen häufig. Stattdessen werden Veränderungen in Prozent oder Unterschiede zwischen Gruppen berichtet. Dennoch zeigen viele Studien positive Auswirkungen von MTSS auf problematisches Schüler*innenverhalten. Änderungen in der Klassenstruktur oder andere Einflussfaktoren

werden oft nicht in die Analysen einbezogen. Die spezifischen Interventionen in den verschiedenen MTSS-Stufen werden selten detailliert beschrieben.

Ein Großteil der Studien (60 %) konzentriert sich zudem auf die Untersuchung von emotionale und sozialen Verhaltensdimensionen (z. B. Blumenthal & Voß, 2016; Pearce, 2009; Tobin & Sugai, 2005; Vincent et al., 2011). Oftmals werden in diesem Zusammenhang die ODRs als Outcome Measure (45 %) definiert (z. B. Vincent et al., 2011; Walker et al., 2016). In 32,5 % der Studien wird außerdem die Implementationsqualität einzelner Maßnahmen oder des gesamten MTSS untersucht (z.B. Sanetti & Collier-Meek, 2015; Sørli & Ogden, 2007; Utley, 2012). Weitere 27,5 % der Studien erforschen die akademische Leistung (Albrecht et al., 2015; Algozzine & Algozzine, 2007; Borgen et al., 2020; Bradshaw et al., 2012; Greulich et al., 2014; Kelm et al., 2014; Sherrod et al., 2009; Walker et al., 2016; Weiland et al., 2014), 12,5 % den Schul- und Klassenkontext sowie Wahrnehmungen von Lehrkräften, Eltern und anderen Stakeholdern (Borgen et al., 2020; Martella et al., 2010; Pearce, 2009; Yeung et al., 2009).

Auffällig ist, dass die sechs Einzelfallstudien detailliertere Ergebnisse präsentieren als viele der größer angelegten Studien. Bei diesen ist der Stichprobenumfang herabgesetzt, jedoch fällt auf, dass studienübergreifend positive Effekte auf disruptives Verhalten (z. B. Sanetti & Collier-Meek, 2015) und den Aufbau arbeitsbezogener, akademischer Verhaltensweisen nachweisbar sind (z. B. Nelson et al., 2018).

Kurz vor Veröffentlichung dieser Rahmenschrift wurde die Suchmatrix erneut durchlaufen. Es konnten für den Zeitraum 2021 bis 2024¹⁴ weitere 351 Artikel identifiziert werden, von denen aber nur fünf Beiträge nach Entfernung der Duplikate für eine Volltextanalyse in Betracht gezogen wurden (Bohanon et al., 2023; Morse, 2023; Robinson et al., 2023; Sutherland et al., 2023; Weist et al., 2022). Alle Beiträge ergänzen den bisherigen Literaturkorpus, erweitern diesen jedoch nicht um widersprüchliche oder abweichende Informationen, sondern reihen sich in die bestehenden Ergebnisse ein. Zudem untersuchen sie (lediglich) spezifische Stufen und Zusammenhänge innerhalb der gestuften Fördersysteme und stellen keine Studienergebnisse dar, die sich auf die Implementation neuer MTSS beziehen. Beachtenswert ist allerdings, dass sich verstärkt Beiträge mit antirassistischer Kritik finden lassen und die Forschungslandschaft um diese wichtige Sichtweise erweitert (Bettors-Bubon et al., 2022; Mayes & Byrd, 2022). Der Ausdruck „*MTSS for all*“ wird hier hinterfragt und auf den Prüfstand gestellt und es werden praktische Implikationen vorgeschlagen, wie MTSS

¹⁴ Bis einschließlich März 2024

antirassistisch bzw. antidiskriminierend aufgestellt werden können (Edirmanasinghe et al., 2022; Fallon et al., 2021). Diese Ergebnisse werden im Kontext der Weiterentwicklung bestehender deutscher MTSS in Kap. 9 näher erläutert.

3.7.3 Implikationen und Ausblick

Die Effektivität von gestuften Fördersystemen hinsichtlich verschiedener Entwicklungsbereiche von Kindern in der Grundschule wurde durch dieses systematische Review bestätigt. Darüber hinaus konnten positive Auswirkungen von MTSS auf die Lehrkraftgesundheit belegt werden. Zudem wurden Forschungslücken sichtbar, die für den weiteren Verlauf der Studienplanung und die weitere Auseinandersetzung mit dem Thema relevant erscheinen: Das Studiendesign vieler großer Erhebungen erlaubt wenige Aussagen über intraindividuelle Verläufe. Eine Lösung und damit einen vertieften Einblick bieten Einzelfallstudien und qualitative Forschungszugänge. Im Kontext mehrstufiger Förderung ist dieser spezifische Blick auf die individuellen Lern- und Entwicklungsverläufe relevant und aussagekräftig. Darüber hinaus könnte eine Erweiterung um Interventionen zur Förderung lehrkraftspezifischer Kompetenzen eine Möglichkeit darstellen, die Kohärenz und Effektivität von MTSS zu stärken. Eine weitere Leerstelle ist der (fehlende) Einbezug aller beteiligten Gruppen bereits in der Entwicklungsphase gestufter Fördersysteme. Laut internationaler Implementationsforschung ist eben jener Einbezug ausschlaggebend hinsichtlich der gelingenden Umsetzung von gestuften Fördersystemen (Fox et al., 2021). Zuletzt wird deutlich, dass Konzepte, die auf dem SWPBS basieren, oftmals einen stärkeren Fokus auf die Reduktion von Verhaltensproblematiken legen. Sie erscheinen somit für die Entwicklung eines MTSS, welches die Reduktion externalisierenden Verhaltens adressiert, am geeignetsten.

Es lässt sich zusammenfassen, dass der systematische Literaturüberblick eine fundierte Grundlage für die Entwicklung und Bewertung relevanter gestufter Fördersysteme in der Grundschule für den deutschpolitischen Bildungsbereich bietet und gleichermaßen Forschungsdesiderate offenbart, die in zukünftigen Forschungsprozessen adressiert werden sollten. Auf dieser Grundlage ist eine dezidierte Entwicklung bzw. Bewertung verschiedener Fördersysteme denkbar und möglich.

4 Der mehrstufige Förderansatz Multimo

Wie bereits dargelegt, erscheint die Entwicklung und Implementation eines gestuften Förderansatzes im deutschen (inkluisiven) Grundschulsystem möglich und gleichzeitig vielversprechend. Ein vordergründiger Schwerpunkt in der Ausgestaltung jener Systeme vermag im Sinne der Studienlage (Kap. 3) und der Prävalenzzahlen (Kap. 2.4) im Bereich der Unterstützung emotionaler und sozialer Kompetenzen sowie der Verringerung von externalisierendem Problemverhalten zu liegen. Aus diesem Grund entwickelte die Forscher*innengruppe um Hanisch et al. (2019) den Förderansatz *Mehrstufige, multimodale und multiprofessionelle Förderung bei externalisierendem Problemverhalten in der Grundschule* (Multimo, Abb. 3), der bis zum Jahr 2020 nicht implementiert und evaluiert wurde. *Multimo* erscheint nach den Ergebnissen des Literaturüberblicks (Beitrag 1), der Analyse der praktischen Umsetzung gestufter Förderung in Deutschland samt bekannter Herausforderungen und erster Ergebnisse als sinnvolle theoretische Grundkonzeption eines MTSS für deutsche Grundschulen. Um dies näher zu erläutern, wird das Konzept im weiteren Verlauf vorgestellt:

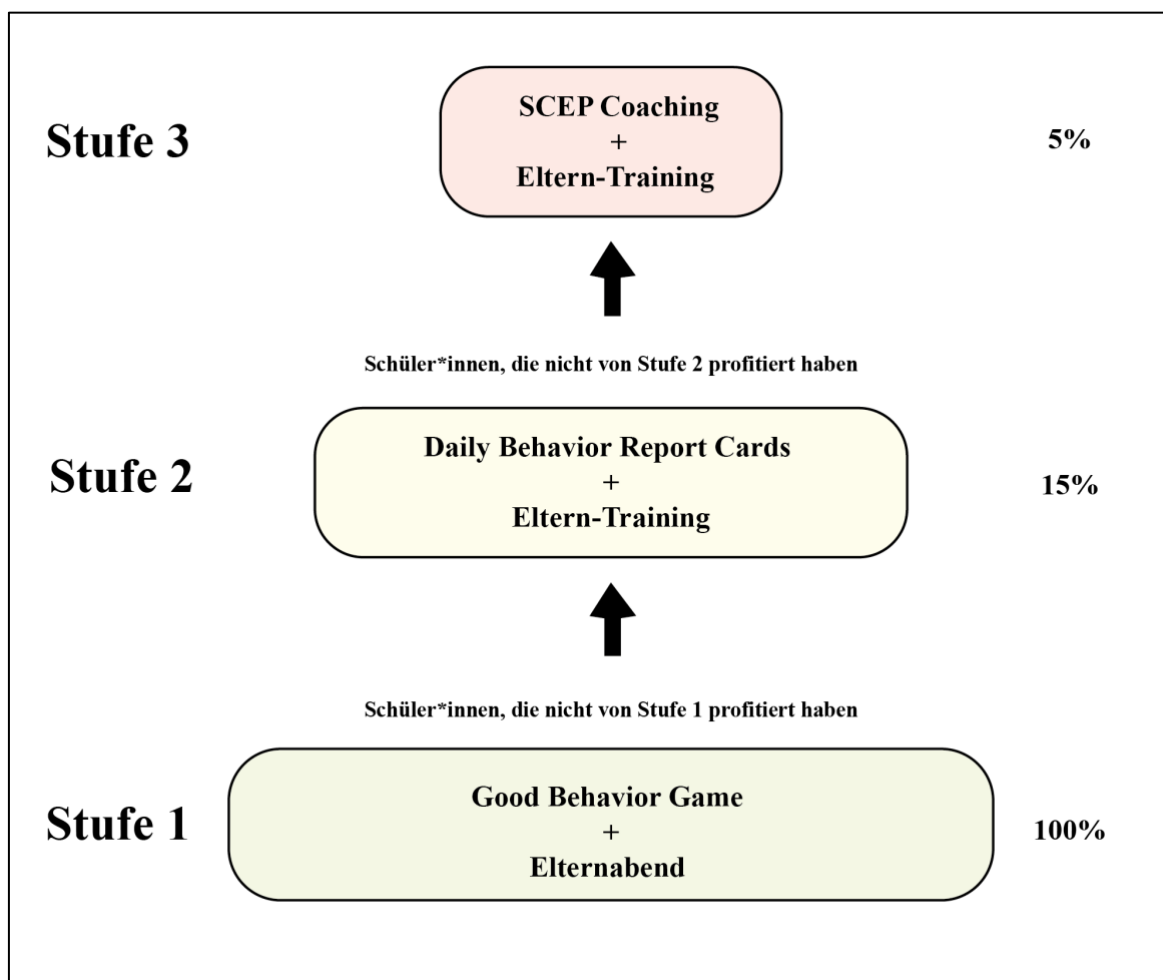
Multimo ist ein mehrstufiges, multimodales Förderkonzept für Grundschulen, das sich auf Schüler*innen mit expansivem bzw. externalisierendem Problemverhalten konzentriert (Hanisch et al., 2019). Externalisierendes Verhalten umfasst, wie bereits dargelegt, hyperaktives, aggressives, oppositionelles und unkonzentriertes Verhalten (Kap. 2). Das Konzept basiert auf dem Ansatz des SWPBS und beinhaltet evidenzbasierte Interventionen, die stufenweise an die individuellen Bedürfnisse der Kinder angepasst sind. *Multimo* umfasst dementsprechend drei Stufen: Auf der ersten Stufe (universal) sollen alle Schüler*innen der Klasse mit der Intervention des *Good Behavior Games* (GBG, Hillenbrand & Pütz, 2008) erreicht werden (Kap. 4.1). Diese Intervention hat ihren Ursprung in den USA und wird dort seit über 50 Jahren beforscht (Barrish et al., 1969). Die zweite Stufe (selektiv) sieht den Einsatz der *Daily Behavior Report Cards* (DBRC, Volpe et al., 2013) vor. Diese Intervention basiert auf einem verhaltensbezogenen Ziel, welches im Kontext der Maßnahme erreicht werden soll (siehe Kap. 4.2). Wenn ein Kind auf der zweiten Stufe ohne gewünschten Fördererfolg bleibt, folgt eine dritte (indizierte) Stufe. Auf dieser Stufe wird ein Lehrkraft-Coaching, das *Schulbasierte Coaching bei Kindern mit expansivem Problemverhalten* (SCEP, Hanisch et al., 2018), durchgeführt (Kap. 4.3). Hier wird die Beziehung zwischen der Lehrkraft und der Schüler*in sowie das Verhalten beider Interaktionspartner*innen in den Mittelpunkt der

Betrachtung gerückt, um durch eine Analyse der Bedingungsfaktoren und etwaige Neuausrichtung auf die herausfordernden Situationen zu reagieren (Hanisch et al., 2020).

Das Konzept sieht zusammenfassend ein umfassendes Fortbildungsprogramm für Lehrkräfte und Schulleitungen vor und zielt allgemein darauf ab, sowohl die schulischen Herausforderungen bei externalisierendem Verhalten zu bewältigen als auch die Lehrkraftbelastung zu reduzieren. Im Folgenden sollen die einzelnen Stufen samt ihren Interventionen näher vorgestellt werden.

Abbildung 3

Der gestufte Förderansatz Multimo



4.1 Stufe 1: Das Good Behavior Game als universelle Fördermaßnahme

Die erste Förderstufe des *Multimo*-Ansatzes sieht die universelle Förderung aller Schüler*innen durch das GBG vor (im Deutschen: KlasseKinderSpiel, Hillenbrand & Pütz, 2008). Es handelt

sich dabei um eine Intervention, die in der Schule zur Förderung von angemessenem Verhalten und zur Unterbindung unerwünschter Verhaltensweisen eingesetzt wird (Barrish et al., 1969; Leidig et al., 2022). Die Klasse wird hierzu in Teams aufgeteilt und muss sich während einer vorher festgelegten Spielzeit an vereinbarte Verhaltensregeln halten. Regelverstöße werden mit Fouls für das jeweilige Team geahndet. Das Team mit den wenigsten Fouls gewinnt am Ende der Spielzeit eine Belohnung. Der Aufbau und die Umsetzung des GBG ist vergleichsweise einfach und es ist gut in den regulären Unterricht integrierbar (Sewell, 2019, 2020). Zudem ist es in hohem Maße adaptierbar und leicht an den jeweiligen Klassenkontext anzupassen (Hillenbrand & Pütz, 2008).

Meta-Analysen können belegen, dass das GBG störendes Verhalten im Klassenzimmer reduziert und positives Verhalten fördert (Flower et al., 2014). Zudem ist bestätigt, dass das GBG bei der Reduzierung externalisierender Verhaltensprobleme eine hohe Wirkung erzielen kann (Breeman et al., 2015; Donaldson et al., 2017; Rubow et al., 2018; Ruiz - Olivares et al., 2010). Verschiedene Studien, darunter viele Einzelfallstudien, zeigen demnach positive Effekte bei Schüler*innen mit psychischen Störungen oder Aufmerksamkeitsproblemen (Jornevald et al., 2023). Die meisten Studien betrachten jedoch ganze Klassen oder Gruppen als Einzelfälle, wodurch weniger Informationen über die Wirkung auf einzelne Schüler*innen verfügbar sind. Bis 2022 fehlte es an Studien im deutschsprachigen Raum mit Kindern mit Verhaltensauffälligkeiten, jedoch konnten für die Grundschule Effekte bezüglich der Reduktion von externalisierendem Verhalten nachgewiesen werden (Leidig et al., 2022).

Im Sinne gestufter Fördersysteme erreicht die erste Förderstufe bestenfalls alle Schüler*innen des Systems (Wexler, 2018). Schüler*innen, die erhöhten Förderbedarf zeigen, können durch Screenings wie die LSL (Petermann & Petermann, 2013), *Integrated Teacher Rating Form* (ITRF, z.B. Daniels et al., 2014) oder den SDQ (Goodman, 1997) identifiziert und bei anhaltendem, ausbleibendem Fördererfolg auf die 2. Stufe, die selektive Förderstufe, transferiert werden.

4.2 Stufe 2: Die Daily Behavior Report Cards als selektive Fördermaßnahme

Auf der zweiten Förderstufe werden die DBRC (Volpe et al., 2013) implementiert und ergänzend wird eine Fortbildung für Eltern bezüglich beziehungsförderlicher und positiver Erziehungsweisen angeboten (Hanisch et al., 2019). Jene Inhalte aus der Fortbildung können mit der Umsetzung der DBRC in den Schulen verbunden werden, da die Maßnahme auch außerhalb der Klasse einsetzbar ist. Dies ist möglich, da bei der DBRC ein spezifisches

verhaltensbezogenes Ziel samt Überprüfungsskala individuell für eine*n Schüler*in festgelegt wird und in verschiedenen Kontexten als Feedbacksystem genutzt werden kann (Iznardo et al., 2020; Riden et al., 2018). Jene Unterstützung erhalten erwartungsgemäß ca. 15 % der Schüler*innen einer Klasse (Wexler, 2018).

Zu den Hauptelementen der DBRC gehören demnach die Festlegung spezifischer Verhaltensziele, die tägliche Bewertung des gezeigten Schüler*innenverhaltens und der regelmäßige kommunikative Austausch über die Ergebnisse zwischen den relevanten beteiligten Personen wie Eltern, Lehrkräften oder den Schüler*innen selbst (Volpe et al., 2013). Um die Methode einzuführen, wird zunächst ein Ziel (oder mehrere Ziele) mit dem*r Schüler*in festgelegt. Das gezeigte Verhalten wird dann auf einer individuellen Skala (z. B. 1-5, in oder Form von Smileys) möglichst zeitnah bewertet. Ebenfalls denkbar ist die Erfassung in Form absoluter Zahlen wie z. B. sechs Meldungen im Unterricht. Die Schüler*innen erhalten somit ein meist unmittelbares Feedback bezüglich ihres gezeigten Verhaltens. Des Weiteren erhalten sie Hilfen, um ihre Ziele zu erreichen und werden beim Scheitern an zukünftige Möglichkeiten erinnert. Wird eine bestimmte Anzahl an Verhaltensweisen gezeigt, kann eine positive Konsequenz (Belohnung) folgen, die settingübergreifend festgelegt werden kann. Die DBRC kann darüber hinaus als Grundlage für Feedback-Gespräche dienen (Ackley, 2022). Die Elternintervention findet in Form eines Elterntrainings statt, das auf dem *Präventionsprogramm für Expansives Problemverhalten* basiert (PEP, Plück et al., 2006)

Überprüft wird der Fördererfolg durch eine engmaschige Begleitung in Form von Verlaufsdagnostik (z. B. in Form täglicher Verhaltensbeobachtung). Hierzu bewerten die Lehrkräfte das gezeigte (problematische) Verhalten täglich und schaffen auf diese Weise eine Datengrundlage, die der Anpassung der Fördermaßnahme dient. Bei ausbleibendem Fördererfolg werden die Schüler*innen entlang der dritten Förderstufe gefördert. Bei Fördererfolg wird je nach Datengrundlage entschieden, ob das Kind auf der 2. Förderstufe verweilt oder ob die Intervention keinen weiteren verhaltensbezogenen Nutzen mehr hat (Hanisch et al., 2019).

4.3 Stufe 3: Das SCEP als indizierte Fördermaßnahme

Wenn die DBRC nicht ihren erwarteten Effektzeitigt und das externalisierende Verhalten fortbesteht, wird im Rahmen der die dritten Stufe das SCEP (Hanisch et al., 2018) durchgeführt. Das SCEP ist ein wissenschaftlich fundiertes Coachingprogramm, das sechs bis acht Sitzungen umfasst und speziell für Lehrkräfte konzipiert ist (Hanisch et al., 2020). Hier wird basierend

auf einer funktionalen Verhaltensanalyse ein maßgeschneidertes Coaching aus zwölf verschiedenen Modulen für die Lehrkraft zusammengestellt. Ziel ist es, die Lernumgebung und die Interaktionen zwischen Lehrkraft und Schüler*in zu verbessern. Im Fokus steht dabei insbesondere die Wahrnehmung von sowohl problematischem als auch erwünschtem Verhalten und die Reaktion auf solches sowie die Förderung der kindlichen Selbstregulation. Die Inhalte des SCEP umfassen unter anderem die Definition individueller Probleme und Ziele, die Gestaltung der Lernumgebung sowie Interventionen zur Stärkung der Beziehung zwischen Lehrkraft und Schüler*innen. Ein wesentlicher Bestandteil ist zudem das Augenmerk auf die Förderung emotional-sozialer Kompetenzen der Lehrkräfte ([Hanisch et al., 2020](#)). Parallel dazu erhalten die Eltern der ausgewählten Kinder auch auf dieser Stufe eine individuell angepasste Fortbildung ([Hanisch et al., 2019](#)).

Das datengeleitete Vorgehen wird hier in derselben Art und Weise wie auf der 2. Stufe fortgesetzt. Die Entscheidung über eine Fortführung der Förderung auf der 3. Stufe wird dementsprechend ebenfalls von dem Fördererfolg des Kindes abhängig gemacht, auch wenn das Coaching stark auf die Bedarfe der Lehrkräfte fokussiert ist.

4.4 Einordnung & Diskussion des Förderansatzes *Multimo*

Der Förderansatz *Multimo* entspricht im Sinne seiner Grundkonzeption den Prinzipien des SWPBS und bildet somit einen gestuften Förderansatz ab, der in der Lage dazu scheint, adäquat auf die Bedarfe von Kindern (und Jugendlichen) mit externalisierendem Problemverhalten zu reagieren. Insbesondere die Kombination von multiprofessionellen Fortbildungskonzepten und dem damit verbundenen Einbezug verschiedener Personengruppen, macht den Ansatz relevant für den Einsatz in der (inkluisiven) Grundschule. Herausfordernd erscheint der Einsatz verschiedener diagnostischer Instrumente durch die Lehrkräfte, da diese auch gemäß [Blumenthal und Voß \(2016\)](#) nicht immer Teil der pädagogischen Praxis von Grundschullehrkräften sind. Zudem ist zu beachten, dass durch den starken Fokus auf das externalisierende Verhalten die Förderung lernbezogenen Verhaltens in den Hintergrund rückt. Dies stellt zunächst eine geringe Limitation dar. Aufgrund der bestehenden Herausforderungen im Grundschulbereich ([Kap. 2](#)) und der Forschungslandschaft in Deutschland zum Thema MTSS ([Kap. 3](#)) erscheint dieser Fokus auf den Rückgang externalisierenden Verhaltens und Aufbau von emotionalen und sozialen Kompetenzen dennoch gewinnbringend und Desiderats schließend.

5 Wirksamkeitsstudien im Kontext Multimo

Die Ergebnisse des systematischen Reviews haben eine breite Studienlage bezüglich der Beforschung gestufter Fördersysteme offengelegt (Beitrag 1). Um die Wirksamkeit von unterschiedlichen MTSS zu untersuchen, wurden international vor allem kontrollierte Einzelfallstudien, qualitative Interviewstudien, Quasi-Experimentelle-Designs und Randomized Controlled Trials durchgeführt. Insbesondere die kontrollierten Einzelfallstudien haben sich in diesem Zusammenhang jedoch als aufschlussreich bezüglich intraindividuellen Verhaltens- und Bildungsverläufen gezeigt und bieten das Potential, auch größere Datenmengen auf Einzelfallebene und Gesamtpopulationsebene belastbar und stichhaltig auszuwerten (Wilbert et al., 2022). Im Kontext einer Wirksamkeitsüberprüfung eines deutschen MTSS im Grundschulbereich erscheinen sie daher als mögliche Studienform, um die Effektivität und den Nutzen gestufter Förderung zu erheben. Um diese Ergebnisse weiter einordnen zu können, sollen Einzelfallstudien im Folgenden theoretisch im Kontext schulischer Interventionsforschung erläutert werden.

5.1 Vor- und Nachteile von Einzelfallstudien und ihren Datensätzen

Einzelfallstudien-Designs erleben seit geraumer Zeit eine weit verbreitete Anwendung neben traditionellen experimentellen Gruppendesigns (Kratochwill et al., 2012; Smith, 2012). Sie bieten die Möglichkeit, tiefgreifende Einblicke in individuelle Veränderungsprozesse zu gewinnen und können hoch valide Ergebnisse produzieren. Dies gilt besonders für die Interventionsforschung (Epstein & Dallery, 2022; Wilbert et al., 2022). Einzelfallstudien werden hier z. B. häufig eingesetzt, um die Wirksamkeit von Interventionen zur Bewältigung von Verhaltensproblemen in Bildungseinrichtungen zu untersuchen (Lipien et al., 2023; Moeyaert et al., 2023). Um sie in der Forschung zum Bereich MTSS gewinnbringend einsetzen zu können, erscheint ein kritischer Blick auf Vor- und Nachteile sowie auf mögliche Auswertungsmethoden und Anwendungsszenarien sinnvoll.

5.1.1 Vorteile von Einzelfallstudien

Einzelfallstudien bieten in der Bildungs- und Psychologieforschung verschiedene Vorteile (Epstein & Dallery, 2022; Kratochwill et al., 2012; Ledford et al., 2022; Lipien et al., 2023; Smith, 2012; Wilbert et al., 2022). Sie ermöglichen z. B. eine intensive und detaillierte

Untersuchung einzelner Fälle (Kinder), was besonders in Bereichen wichtig ist, in denen individuelle Unterschiede und spezifische Kontexte eine Rolle spielen (Lipien et al., 2023). Diese individuumsbezogene Ausrichtung ist z. B. in der Sonderpädagogik wichtig, da hier oftmals Interventionen entlang individueller Bedürfnisse geplant werden, um effektiv auf Problemverhalten oder geringes lernbezogenes Verhalten zu reagieren. Dies macht sie dementsprechend attraktiv für die Forschung im Kontext von MTSS (Beitrag 1).

Darüber hinaus eignen sie sich als erster Ausgangspunkt für den Erkenntnisgewinn von bisher wenig erforschten Themenbereichen (Smith, 2012). Selbst bei geringen Stichprobengrößen (z. B. $N = 1$) können sie valide Daten mit geringem Aufwand erbringen und somit größere Studien vorbereiten (Kratochwill & Levin, 2010; Manolov & Solanas, 2008). Ebenso können Forscher*innen Bedingungen oder Umstände beleuchten und so zu einem umfassenderen Verständnis des Untersuchungsgegenstandes beitragen. Dies ist in größeren Studienformaten (z. B. experimentellen Kontrollgruppendesigns) aufgrund von umfangreichen Stichproben kaum detailliert möglich (Shadish & Sullivan, 2011). Zusätzlich bieten sie ein hohes Maß an Flexibilität in Bezug auf Forschungsdesign und -methodik, was es ermöglicht, spezifische Fragestellungen zu beantworten.

5.1.2 Problembereiche von Einzelfallstudien

Eine frühe jedoch weiterhin aktuelle und umfangreiche systematische Literaturübersicht mit der Bewertung verschiedener Einzelfallstudien-Designs liefert Smith (2012). Der Beitrag konzentriert sich vorrangig auf verwendete Messmethoden und analytische Techniken, die angewandt werden. Besondere Aufmerksamkeit gilt der statistischen Testkraft und methodologischen Herausforderungen, die vor allem im Bereich einer gütegeleiteten Umsetzung und Auswertung liegen. Der Beitrag offenbart verschiedene Problembereiche, die durch weitere Studien im Folgenden kontextualisiert werden.

Ein Hauptproblem der Einzelfallforschung ist die Erhebung einer repräsentativen Ausgangslage (Baseline) und das damit verbundene Management der Nicht-Unabhängigkeit sequenzieller Beobachtungen, die auf wiederholten Messungen innerhalb eines einzelnen Subjekts oder einer Untersuchungseinheit basieren. Der hierfür verwendete Begriff „Autokorrelation“ beschreibt somit die Tendenz von Datenpunkten, die in einer Phase erhoben werden, sich ähnlich zu verhalten oder voneinander abhängig zu sein (Borckardt et al., 2008). Wenn beispielsweise ein bestimmtes Verhalten oder ein Messwert zu einem Zeitpunkt A erhoben wird, kann das Ergebnis des nächsten Messzeitpunkts B teilweise durch das Ergebnis des vorherigen Zeitpunkts (A) beeinflusst werden. Diese Abhängigkeit der Messwerte

voneinander stellt eine Herausforderung für die statistische Analyse dar, da viele gängige Auswertungsmethoden von einer kohärenten Datenunabhängigkeit ausgehen (Snijders & Bosker, 2011). Die Bewältigung dieses Problems erfordert somit spezielle Auswertungsmethoden, die in der Lage dazu sind, die interne Struktur wiederholter Messungen zu berücksichtigen. Beispiele hierfür sind Multilevel-Modelle oder Zeitreihenanalysen. Diese Methoden können helfen, die Effekte der Autokorrelation zu kontrollieren und genauere Schlussfolgerungen über die Daten zu ermöglichen (Smith, 2012).

Ein weiteres Problem sind falsche oder fehlende Daten (Missings), die häufig im Kontext von Wiederholungsmessungen auftreten. Methoden wie die tägliche Verhaltensbeurteilung (engl. Direct Behavior Rating [DBR], Huber & Rietz, 2015) sollen das Risiko von falschen Daten reduzieren, indem sie möglichst kleinschrittig genaue Echtzeitinformationen erfassen. Allerdings sind auch diese Methoden anfällig für fehlende Daten aufgrund von Vergesslichkeit, Nichtbesitz des Materials, welches zur Erfassung notwendig ist oder absichtlicher bzw. systematischer Nichtbefolgung (Smith et al., 2012). Unabhängig davon ergeben sich aus Einzelfallerhebungen häufig Datensätze, die nicht normalverteilt sind (Wilbert et al., 2022). Hierzu gehören z. B. auch von Beobachter*innen bewertetes Verhalten (bspw. Lehrkraft-Beurteilung von Schüler*innen). Gängige statistische Methoden, insbesondere solche, die sich auf die Modellfindung konzentrieren, gehen jedoch von normalverteilten Daten aus. Die nähere Betrachtung valider Auswertungsmethoden, die jene Problembereiche berücksichtigen, erscheint demnach notwendig.

5.1.3 Auswertungsmethoden von Datensätzen von Einzelfallstudien

Eine Möglichkeit, nicht normalverteilte Daten erfolgreich und valide zu analysieren, stellen z. B. die Zero-Inflated Poisson Regression, negative Binomialregression oder hierarchische lineare Modelle sowie Multilevel-Analysen und Strukturgleichungsmodellierungen dar (Smith, 2012). Auch wenn eine verzerrte Datenverteilung durch die Erhebungssituation nicht verhindert werden kann, können sie nicht normalverteilte Datensätze valide analysieren. Darüber hinaus sind jene Techniken besonders hilfreich für die Auswertung von Trajektorien (Entwicklungspfade) und für Steigungen in Längsschnittdaten sowie für die Verknüpfung von Veränderungen mit potenziellen Kovariaten. Multilevel-Analysen und verwandte hierarchische lineare Modelle können demnach z. B. auch die Beziehung zwischen verschiedenen Variablen beleuchten und klären, ob sich die Ausprägungen dieser zwischen den Proband*innen unterscheiden (Manolov et al., 2014).

Neben regressionsbasierten Auswertungsmethoden können auch Ansätze wie die Berechnung standardisierter Mittelwertdifferenzen (z. B. Cohens d , oder Hedge's g) für die Auswertung von Einzelfalldaten verwendet werden (Smith, 2012). Ein geringer Nachteil dieser Ansätze ist, dass sie nicht für die Verarbeitung von autokorrelierten Daten entwickelt wurden. Dennoch gelten sie als robust gegenüber nicht normalverteilten Daten, sodass die Autokorrelation die Effektgrößen, die mit diesen Ansätzen berechnet werden, wenig stark beeinflussen (Manolov & Solanas, 2008). Wichtig ist jedoch zu vermerken, dass diese Methoden nur geeignet sind, wenn signifikante Unterschiede zwischen den Phasen der Studie untersucht werden sollen. Sie können keine Trajektorien oder Beziehungen zwischen Variablen beleuchten.

Weitere Ansätze zur Analyse vor allem von Niveau- und Steigungsveränderungen in kurzen (weniger als 30 Beobachtungen pro Phase), aber ebenfalls autokorrelierten Datensätzen sind nichtparametrische Methoden (Chen et al., 2016; Lenz, 2017; Smith, 2012). Hierzu gehören z. B. die Analysen der Datenpunkte, die einander überlappen. Ein Beispiel ist der *Non-Overlap of All Pairs* (NAP, Parker & Vannest, 2009). Der NAP wird berechnet, indem man jedes Datenpaar aus unterschiedlichen Phasen miteinander vergleicht: Ein nicht überlappendes Paar liegt vor, wenn ein Datenpunkt aus Phase B größer als sein entsprechender Datenpunkt aus Phase A ist und der ermittelte Wert ergibt sich aus der Anzahl der nicht überlappenden Paare geteilt durch die Gesamtzahl der Vergleiche. Der NAP kann z. B. manuell aus einem Datenplot, durch Zwischenstatistiken aus dem Wilcoxon-Rangsummentest oder als Fläche berechnet werden (Alresheed et al., 2013; Parker & Vannest, 2009).

Ein weiterer Index ist der *Percentage of All Non-Overlapping Data* (PAND, Parker et al., 2016), für dessen Berechnung der Vergleich aller Datenpunkte zwischen zwei Phasen (z. B. Baseline und Intervention) erfolgt. Für jedes Paar von Datenpunkten aus den verschiedenen Phasen wird geprüft, ob (k)eine Überlappung vorliegt. Der ermittelte Wert ist dann der prozentuale Anteil der nicht überlappenden Paare an der Gesamtzahl der Paarvergleiche. Dieses Maß berücksichtigt dementsprechend alle Datenpunkte des Datensatzes und ist damit umfassender als andere Maße, die sich nur auf bestimmte Extremwerte konzentrieren, z. B. der *Percentage of Non-Overlapping Data* (PND, Scruggs et al., 2016). Zudem kann dieser Index manuell berechnet werden, indem die Datenpunkte grafisch analysiert und ausgezählt werden.

Jamshidi et al. (2020) untersuchen in ihrer Studie die Wirksamkeit verschiedener Mehrebenenmodelle im Umgang mit Abhängigkeiten zwischen mehreren Regressionskoeffizienten. Durch den Einsatz von Monte-Carlo-Simulationen wird die Leistung dieser Modelle bei der genauen Schätzung von Gesamtbehandlungseffekten und

Varianzkomponenten bewertet. Die Forschenden empfehlen die Verwendung eines multivariaten Mehrebenen-Ansatzes zur Modellierung der bestehenden Abhängigkeit zwischen verschiedenen Regressionskoeffizienten (z. B. der Level- oder der Slope-Effekt)¹⁵, um die Aussagekraft der Studien weiter zu erhöhen. In komplexeren Modellen, wie Mehrebenen- oder hierarchischen linearen Modellen, können diese Effekte auf verschiedenen Ebenen der Datenstruktur berechnet werden. Insbesondere in der Schulforschung kommt solchen Verfahren eine hohe Aussagekraft zu, die Daten oft eine genestete Datenstruktur aufweisen, die nicht immer berücksichtigt werden kann. Im Folgenden sollen somit Einzelfallstudien im Kontext schulischer Interventionsforschung näher beleuchtet werden.

5.2 Einzelfallstudien im Kontext von schulischer Interventionsforschung

Einzelfallstudien stellen einen häufig genutzten Forschungsansatz in der schulischen Interventionsforschung dar, um Verhaltensänderungen zu verstehen und sie zu quantifizieren (Smith, 2012). Typischerweise werden sie mit Hilfe von DBR durch Lehrkräfte durchgeführt und ermöglichen Einblicke in die Entwicklung spezifischer Verhaltensweisen von Schüler*innen über einen festgelegten Zeitraum (Harrison et al., 2014; Smith et al., 2018). Im Unterschied, z. B. zu Studien im Wartekontrollgruppen-Design, generieren die Teilnehmenden in Einzelfallstudien ihre eigenen intraindividuellen Vergleichsdaten. Dies geschieht, indem für ein spezifisches Verhalten desselben Individuums wiederholt Bewertungen vorgenommen werden. Somit entsteht ein Datensatz, der eine präzise Nachverfolgung des personenbezogenen Verhaltens und zeitgleich den Vergleich zwischen den Teilnehmenden ermöglicht (Maggin et al., 2018). Veränderungen können dann entlang definierter Phasen errechnet werden. Eine gängige Methode zur Auswertung der gewonnenen Daten sind Regressionsanalysen (Kratochwill et al., 2022). Die Grundlage für diese Auswertung ist ein Studiendesign, welches verschiedene Phasen der Datenerhebung vorsieht, die zeitversetzt starten (Multiple-Baseline-Design, Levin et al., 2018). In der Forschungspraxis wird zwischen verschiedenen Forschungsdesigns von Einzelfallstudien unterschieden, die im Folgenden kurz erläutert werden.

Ein prominentes Forschungsdesign ist das *Single-Case Experimental Design* (SCED), auch *Reversal/Withdrawal-* oder *ABAB-Versuchsdesign*, genannt. In so einem ABAB-Design

¹⁵ Level-Effekte und Slope-Effekte werden im Kapitel 5.2 ausführlicher erklärt.

wird zunächst eine Baseline (A-Phase) festgelegt, in der von dem zu beobachtenden Verhalten ohne Intervention Daten gesammelt werden. Darauf folgt die B-Phase, in der die Intervention durchgeführt wird. Anschließend kehrt die Studie in eine erneute A-Phase (*Reversal* oder *Withdrawal*) zurück, in der die Intervention entfernt wird, um zu beobachten, ob das Verhalten zum Ausgangszustand zurückkehrt. Schließlich wird in der zweiten B-Phase die Intervention erneut eingeführt, um die Konsistenz der Effekte zu bestätigen (Krasny-Pacini & Evans, 2018).

Single-Subject Experimental Designs (SSED) sind ähnlich aufgebaut mit dem Unterschied, dass sie häufig in medizinischen oder therapeutischen Kontexten angewandt werden (Epstein & Dallery, 2022; Krasny-Pacini & Evans, 2018). Ein typisches Beispiel hierfür ist das *N-of-1-Trial*, bei dem eine einzelne Versuchsperson über mehrere Phasen hinweg beobachtet wird. Diese Art von Design wird oft verwendet, um die Wirksamkeit einer Behandlung für einen individuellen Fall zu bestimmen.

Weitere im pädagogischen Bereich verbreitete Ansätze sind *Single-Subject-Research-Designs* (SSRD). Sie umfassen in der Regel Multiple-Baseline-Designs, die sich in drei Kategorien einteilen lassen: über Teilnehmer*innen hinweg, über Settings hinweg und über Verhaltensweisen hinweg (Krasny-Pacini & Evans, 2018). Bei einem Design über Teilnehmer*innen hinweg wird die Intervention zu unterschiedlichen Zeiten bei verschiedenen Teilnehmer*innen eingeführt, wohingegen bei einem Design über Settings hinweg die Intervention in verschiedenen Umgebungen oder Kontexten angewandt wird. Bei einem Design über Verhaltensweisen hinweg werden unterschiedliche Verhaltensweisen derselben Person untersucht.

Diese Designs sind besonders nützlich, um die Generalisierbarkeit der Ergebnisse zu testen und die Wirksamkeit der Intervention in verschiedenen Kontexten oder auf verschiedene Verhaltensweisen hin zu überprüfen. Die Auswertung jener Daten, die aus Einzelfallstudien gewonnen werden können, erfolgt in den meisten Fällen ungeachtet des Designs mit denselben Indizes (Lenz, 2017). Zugrunde liegt zumeist eine visuelle Inspektion der Daten, die einen groben, eher unsystematischen Überblick über die Datenlage ermöglicht. Kenngrößen, die auf mathematischen Berechnungen und dem Verhältnis der ermittelten Daten basieren, ermöglichen zusätzlich Aussagen über die Signifikanz und Größe der Effekte (Wilbert et al., 2022). Hierunter fallen z. B. Effekte wie der Level-, der Slope- oder der Trend-Effekt (Wilbert & Lueke, 2022):

- **Level-Effekt:** Der Level-Effekt beschreibt die Änderung einer Variable zwischen zwei Phasen. Die Größe gibt eine unmittelbare Verschiebung im Basisniveau der Daten zum Zeitpunkt des Phasenübergangs an.
- **Slope-Effekt:** Der Slope-Effekt betrifft eine Änderung in der Steigung (oder Änderungsrate) im Regressionsmodell. Die Größe gibt an, wie schnell die Daten über die Zeit, bezogen auf einen Messzeitpunkt, zunehmen oder abnehmen.
- **Trend-Effekt:** Der Trend-Effekt gibt systematische Änderungen in den Daten über die Zeit hinweg (über alle Phasen) pro Messzeitpunkt an.

Level-, Slope- und Trend-Effekte geben demnach wichtige Einblicke in die Art und Weise, wie sich Daten über die Zeit entwickeln, und sind entscheidend für das Verständnis der Wirkung von Interventionen in Einzelfallstudien. Sie ermöglichen, subtilere Veränderungen und Muster in den Daten zu identifizieren, die über einfache Vorher-Nachher-Vergleiche hinausgehen. Darüber hinaus bieten folgende weitere Kenngrößen Informationen über die Daten und die Entwicklung des beobachteten Verhaltens (Chen et al., 2016; Lenz, 2017; Parker, Vannest, & Davis, 2011; Wilbert et al., 2022):

- **PND (Percentage of Non-overlapping Data):** Diese Kenngröße gibt an, wie viel Prozent der Daten in der Interventionsphase sich nicht mit den Daten der Baseline-Phase überlappen. Ein höherer PND-Wert deutet auf eine effektivere Intervention hin (Ma, 2006).
- **PET (Percentage of Exceeding Targets):** Diese Kennzahl misst, wie oft das Verhalten während der Interventionsphase ein vordefiniertes Zielniveau überschreitet (Krasny-Pacini & Evans, 2018).
- **NAP (Nonoverlap of All Pairs):** Ähnlich wie PND misst NAP die Unterschiede zwischen den Daten vor und nach einer Intervention, berücksichtigt aber alle möglichen Paarvergleiche zwischen den Datenpunkten (Parker & Vannest, 2009).
- **PAND (Percentage of All Non-overlapping Data):** Dieser Index ist eine Erweiterung des PND und betrachtet alle nicht-überlappenden Datenpunkte in einem Datensatz und ist robust gegenüber Ausreißern (Yucesoy - Ozkan et al., 2019).
- **PEM (Percentage of Exceeding the Median):** Der PEM misst, wie oft das beobachtete Verhalten während der Interventionsphase den Median der Baseline-Phase übersteigt (Ma, 2006).

- **PEM-T (Percentage of Data Exceeding a Median Trend):** Der PEM-T misst, wie viele Datenpunkte der Interventionsphase über dem Baseline-Trend liegen (Spilles & Hagen, 2019).
- **Tau-U:** Tau-U ist eine statistische Kennzahl, die die Nicht-Überlappung von Daten unter Berücksichtigung von Trends in der Baseline-Phase misst. Tau-U ist hilfreich, um sowohl die Richtung als auch das Ausmaß der Veränderung zwischen Phasen zu bestimmen (Chen et al., 2016; Yucesoy - Ozkan et al., 2019).

Zusammenfassend kann demnach gesagt werden, dass Einzelfallstudien durch eine Vielzahl an Designs und Auswertungsmethoden eine hohe Aussagekraft und Güte vorweisen können. Insbesondere für die Auswertung von Daten gestufter Fördersysteme zeigen sie Potenzial, da sie die verschiedenen Stufen eines Systems in Form von verschiedenen Phasen bzw. Förderung abbilden können. Sie ermöglichen somit ein hohes Maß an Flexibilität und Adaptionenmöglichkeit, was der praktischen Umsetzung gestufter Förderung entgegenkommt.

5.3 Wirksamkeit des Förderansatzes *Multimo* auf Stufe 1 & 2¹⁶

Anschließend an die vorherigen Überlegungen und theoretischen Ausarbeitungen erfolgte die erste Wirksamkeitsstudie des Förderansatzes *Multimo* in Form einer Einzelfallstudie. Die Stufen 1 und 2 wurden in Form von gruppenbezogenen Analysen über alle Fälle der Stichprobe hinweg evaluiert. Die dritte Stufe ist in dem publizierten Beitrag aufgrund einer geringeren Stichprobe und dem damit verbundenen Wechsel des Auswertungsdesigns nicht enthalten. Im folgenden Kapitel werden jene Ergebnisse der Stufe 3 jedoch ergänzend bereitgestellt (Kap. 5.4). Zusätzlich wird die Stufe 3 in Beitrag 3 und in einer sich im Druck befindlichen Publikation der Forschungsgruppe¹⁷ gesondert betrachtet.

5.3.1 Methodik

Die Stichprobe der Studie umfasste 29 Schüler*innen der ersten und zweiten Klasse aus 18 inklusiven Grundschulen in Nordrhein-Westfalen. Diese Schüler*innen zeigten laut Lehrkrafturteil externalisierende Verhaltensweisen. Die Studie wurde über etwa sieben Monate

¹⁶ Nitz, J., Hagen, T., Krull, J., Verbeck, L., Eiben, K., Hanisch, C., & Hennemann, T. (2023). Tiers 1 and 2 of a German MTSS: impact of a multiple baseline study on elementary school students with disruptive behavior. *Frontiers in Education*, 8. <https://doi.org/10.3389/educ.2023.1208854>

¹⁷ Zum aktuellen Zeitpunkt ist keine vorläufige Literaturangabe möglich (März, 2024)

durchgeführt und beinhaltete die Implementation von zwei verhaltensbezogenen Interventionen: das GBG in Stufe 1 und die DBRC in Stufe 2 entlang des Förderansatzes *Multimo* (Hanisch et al., 2019). Das Design der Studie entsprach einer Einzelfallstudie mit Multiplen-Baseline-Design, bei dem die Klassen der Schulen zu unterschiedlichen Zeiten mit den Interventionen begannen. Daten zum störenden (disruptiven) Verhalten wurden mit Hilfe der DBR auf einer Skala von 0 bis 10 erhoben.

5.3.2 Implementation & Durchführung

Um *Multimo* in den Schulen umzusetzen, wurde ein Multiplikator*innenkonzept genutzt (z. B. Behr et al., 2020; Pearce et al., 2012). Hierzu wurden zunächst 18 erfahrene Lehrkräfte und Schulleitungen von Universitätspersonal in Bezug auf die Verfahrensweisen von *Multimo* geschult. Diese Moderator*innen schulten dann das Schulpersonal hinsichtlich der Implementation und begleiteten diesen Prozess. In ganztägigen Schulungen wurden das Konzept der gestuften Unterstützung, individuelle Interventionen und diagnostische Methoden vermittelt. Der Prozess beinhaltete wöchentliche Austauschtreffen mit den Moderator*innen und dem Universitätsprojektteam. Die Implementation und Durchführung der Studie dauerte etwa sieben Monate: Die erste Stufe wurde Anfang Dezember 2021, die zweite Ende Januar 2022 und die dritte im März 2022 eingeführt. Während des gesamten Projekts wurde das GBG (Stufe 1) eingesetzt, ergänzt durch die DBRC (Stufe 2). Zur Identifizierung von Kindern mit störendem Verhalten wurde eine Kurzversion der ITRF verwendet. Die ITRF ist ein universelles Verhaltensscreening, das die Ausprägung externalisierenden Verhaltens erhebt. Die Verlaufsdaten wurden mit Hilfe der DBR erhoben.

Die Einzelfallstudie wurde wie bereits erwähnt in Form eines Multiplen-Baseline-Designs geplant und durchgeführt. Die verschiedenen Schulen erhielten aufgrund dessen unterschiedliche Startzeiten für die Interventionen und die Baseline-Phasen variierten hinsichtlich ihrer Dauer. Um die Durchführung des Multiplen-Baseline-Designs und des Forschungsdesigns zu gewährleisten, enthielten die DBRs explizite Anweisungen und wurden durch das Projektteam kontinuierlich überwacht. Dadurch konnte sichergestellt werden, dass das Forschungsdesign problemlos in den Schulalltag integriert werden konnte.

5.3.3 Auswertung & Analyse

Die Daten wurden mittels Inferenzstatistik und deskriptiver Statistik für Einzelfalldaten untersucht. Dabei wurden Mehrebenen-Analysen und der NAP berechnet. Mithilfe der Mehrebenen-Analysen wurden Veränderungen über die gesamte Stichprobe hinweg erfasst,

wobei sowohl Level- als auch Slope-Effekte berechnet wurden. Insgesamt umfassten die drei Phasen 1.498 Messzeitpunkte. Für die fallübergreifende Analyse wurden zwei Modelle entwickelt: Das erste Modell vergleicht die Baseline-Phase (Phase A) mit der GBG-Interventionsphase (Phase B), das zweite Modell vergleicht das störende Verhalten in der Phase B mit dem in Phase C (zweite Interventionsstufe, Einsatz DBRC). Die Analysen wurden mit *R-Studio* (Posit-Team, 2022) und dem Paket *scan* (Wilbert & Lueke, 2022) durchgeführt.

5.3.4 Ergebnisse

Die Ergebnisse zeigen eine signifikante Reduzierung des störenden Verhaltens auf Stufe 1. Die Zuweisung eines Kindes zur Stufe 2 zeigt eine weitere Abnahme des störenden Verhaltens (im Trend-Effekt der B- und C-Phase). Die Reduktion auf der zweiten Stufe war jedoch nicht so stark wie die Reduktion des Verhaltens auf der ersten Stufe. Die Ergebnisse zeigen, dass GBG und DBRC zwar gemeinsam in einem MTSS implementiert werden können, um störendes Verhalten zu reduzieren, der Effekt der zweiten Stufe auf die Reduktion des störenden Verhaltens allerdings anderweitig überprüft werden muss. Der Gesamtansatz der Stufen 1 und 2 stellte sich jedoch insgesamt als effektiv zur Reduzierung des störenden Verhaltens innerhalb der Stichprobe dar.

5.3.5 Implikationen & Ausblick

Wie bereits kurz beschrieben ist zunächst festzuhalten, dass die Stufen 1 und 2 des Förderansatzes *Multimo* unterschiedlich effektiv darin sind, störendes Verhalten zu reduzieren. Die Kombination der Maßnahmen auf der zweiten Stufe lassen infolgedessen keine isolierten Aussagen über die Wirkung der DBRC auf der zweiten Stufe oder über eine möglicherweise fortlaufend bestehende Wirkung des GBG zu. Aus diesem Grund sind weitere Studien notwendig, um isolierte Effekte in einem ähnlichen Setting zu überprüfen. Darüber hinaus sind aufbauend auf diesen Ergebnissen keine Aussagen über die Wirkung der dritten Stufe möglich. Da die Ergebnisse auf Analysen über die Gesamtstichprobe hinweg beruhen und alle ausgewählten Kinder die 1. und 2. Förderstufe, aber nicht zwingend die 3. Stufe durchliefen, wären die Schlussfolgerungen hinsichtlich der Effekte der Intervention auf Stufe 3 möglicherweise verzerrt gewesen. Im Sinne der Validität der Daten in Bezug auf die Fragestellung wurde sich dementsprechend zunächst auf die Analyse der Daten auf Stufe 1 und 2 beschränkt, was im Folgenden um die Analysen der Kinder, die Stufe 1,2 und 3 durchliefen (also ein Subsample der Gesamtstichprobe), erweitert werden kann.

5.4 Wirksamkeit des Förderansatzes auf Stufe 3 anhand von Einzelfalldaten

Aufgrund des beschriebenen Fokus auf die Stufen 1 und 2 erschien eine Auswertung der Daten bezüglich der Stufe 3 in dem Kontext dieser Rahmenschrift sinnvoll. Als Fragestellung ergibt sich demnach: Wie effektiv ist die Förderung der Schüler*innen entlang der dritten Förderstufe hinsichtlich der Ausprägung des disruptiven Verhaltens? Hierfür wurde aus dem vorliegenden Datenmaterial ein Subsample erstellt. Es umfasst acht Schüler*innen aus fünf Schulen (C1, C7, C9, C12, C14, C17, C20, C24). Alle Schüler*innen hatten unterschiedliche Lehrkräfte, die die Förderung auf Stufe 3 (SCEP) durchliefen. Folglich war keines der Kinder mit einem anderen Kind in derselben Klasse. Alle weiteren Angaben zu den Schüler*innen entsprechen denen des Beitrags 2 und sind ergänzend in Tabelle 2 dargestellt.

5.4.1 Methodisches Vorgehen der ergänzenden quantitativen Datenauswertung

Die erhobene Variable beschreibt das disruptive Verhalten eines Kindes auf einer Skala von 0 bis 10 (Beitrag 2). Die Analyse der Daten erfolgte ebenfalls mit dem Paket *scan* in R-Studio (Posit-Team, 2022; Wilbert & Lueke, 2022). Aufbauend auf den deskriptiven Daten der Gesamtstichprobe wurde für das Subsample eine gleiche deskriptive Analyse durchgeführt. Anknüpfend daran wurde für die betreffenden Schüler*innen aufgrund der geringen Stichprobengröße eine intraindividuelle Auswertung in Form von fallbezogenen Regressionsmodellen vorgenommen und durch den NAP ergänzt. Für alle Fälle wurden individuelle Data-Plots erstellt, um die Verläufe zuletzt zu visualisieren. Dementsprechend unterscheidet sich die Auswertungsmethode von der des Gesamtsamples in einigen Bereichen (Beitrag 2).

5.4.2 Ergebnisse

Die deskriptiven Ergebnisse zeigen, dass in der Phase D (Stufe 3) im Durchschnitt 68,5 (Range: 58-77) Messzeitpunkte vorgesehen waren, wovon 32,2 (Range: 26-41) Missings pro Fall auftraten. In Tabelle 2 werden die deskriptiven Daten des Subsamples der ergänzenden Auswertung der Stufe 3 dargestellt. Sie zeigen ein durchschnittliches Alter der Stichprobe von 6,8 Jahren und, dass alle Kinder entlang mindestens einer Sub-Skala oder dem Gesamtproblemwert der ITRF auffälliges Verhalten zeigen. Die Abbildungen 4, 5, 6 und 7 zeigen die individuellen Entwicklungsverläufe der Schüler*innen.

Tabelle 2*Deskriptive Daten des Subsamples (Stufe 3)*

Fall ¹	Schule/Klasse	Geschlecht ²	Alter ³	ITRF _{APD}	ITRF _{OB}	ITRF _{Total}
C1	A/a	2	7	13	4	17
C7	B/c	2	7	8	6	14
C9	B/e	2	7	12	1	13
C12	C/f	2	6	11	18	29
C14	C/g	2	7	8	10	18
C17	D/h	2	7	5	7	12
C20	D/i	2	7	22	21	43
C24	E/j	2	7	16	16	32

¹ Fallnummern entsprechen denen der Gesamtstichprobe; ² = männlich; ³ = in Jahren; ITRF_{APD} = ITRF Academic Productivity/Disorganization sub-scale (cut off >=9); I; TRF_{OB} = Oppositional/Disruptive sub-scale (cut off >=4); ITRG Total = ITRF Total Score (cut off >=11)

Der NAP nahm Werte von 25.81 bis 78.54 an. Das Subsample zeigte somit für keine Kinder einen starken Effekt, für vier Kinder (C9, C12, C20, C24) einen mittleren und für ein Kind (C17) einen geringen Effekt.¹⁸ Drei Kinder zeigten keinen Effekt (C1, C7, C14), wobei der NAP von C14 nur knapp unter dem Cut-Off lag (Tabelle 3).

Tabelle 3*Effekte auf Fallebene des Subsamples (Stufe 3)*

Fall	n _D	mis _D	M _D (SD)	Level _{C-D}	Slope _{C-D}	NAP _{C-D}
				B	B	%
C1	61	30	3.68 (1.11)	3.33*	0.03	25.81
C7	75	39	1.81 (1.45)	1.04	-0.00	41.44
C9	62	41	0.71 (1.01)	2.07	0.06	66.22
C12	72	26	2.48 (1.09)	-2.17*	-0.00	72.07
C14	68	32	4.78 (1.12)	0.27	-0.13*	48.44
C17	75	30	1.93 (2.49)	-0.18	0.10	56.53
C20	77	32	0.38 (0.81)	-1.45*	-0.02	78.54
C24	58	27	7.35 (1.96)	-3.41*	0.09	71.52

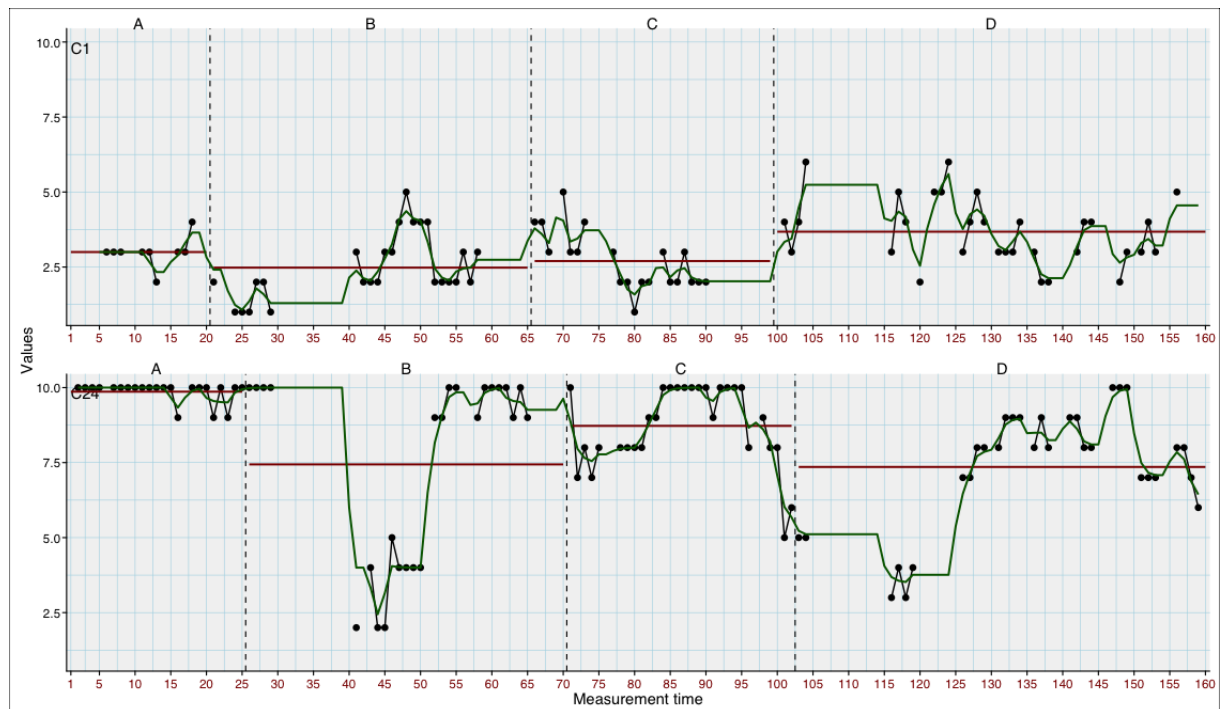
n_D = Messzeitpunkte der Phase D; mis_D = Missings der Phase D; M_D = Mittelwert der Phase D; SD = Standardabweichung; NAP = Non-Overlap of all Pairs in %

¹⁸ Ein Wert im Bereich von 50-65 gilt als kleiner Effekt, ein Wert im Bereich von 66-92 als mittlerer Effekt und ein Wert über 93 als starker Effekt (Parker & Vannest, 2009).

Die Analysen hinsichtlich der Level- und Slope-Effekte zeigen negative signifikante Effekte für vier Fälle (50 %; C1, C12, C20, C24). In drei Fällen (37,5 %; C12, C14, C20) ist der Level-Effekt signifikant und in einem Fall (12,5 %; C24) ist der Slope-Effekt signifikant. Ein Level-Effekt ist positiv signifikant (C1, 3.33). Die signifikanten Level-Effekte liegen zwischen -3.41 und 3.33 und der signifikante Slope-Effekt beträgt -0.13. Die folgenden Visualisierungen zeigen zusammengefasst alle Fälle des Subsamples. Die Fälle werden nach Schulen gruppiert. C1 und C24 werden in einer Grafik der Übersichtlichkeit halber zusammengefasst, obwohl es sich hierbei nicht um Fälle der gleichen Schule handelt. Die Fälle C7 & C9, C12 & C14, C17 & C20 werden entsprechend ihrer Schulzugehörigkeit zusammengefasst.

Abbildung 4

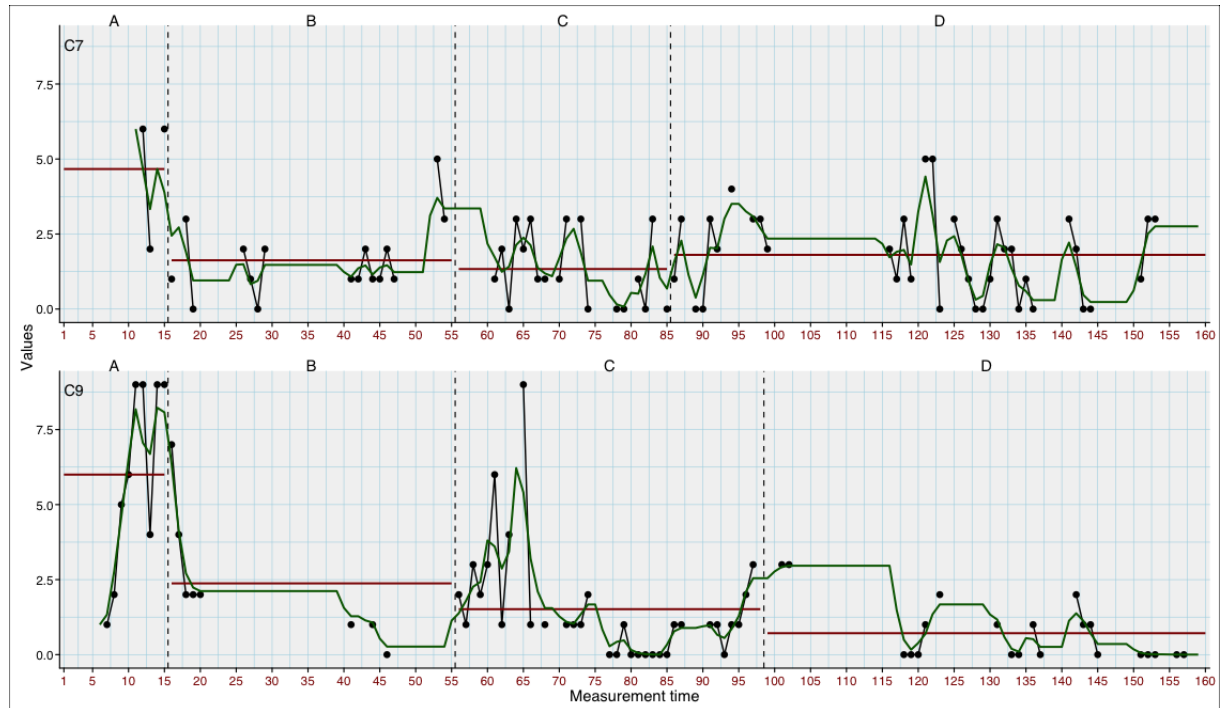
Visuelle Darstellung für die Fälle C1 & C24



Anmerkung: A = Baseline-Phase, B = 1. Stufe, C = 2. Stufe, D = 3. Stufe, Measurement time = Messzeitpunkt, Values = Werte

Abbildung 5

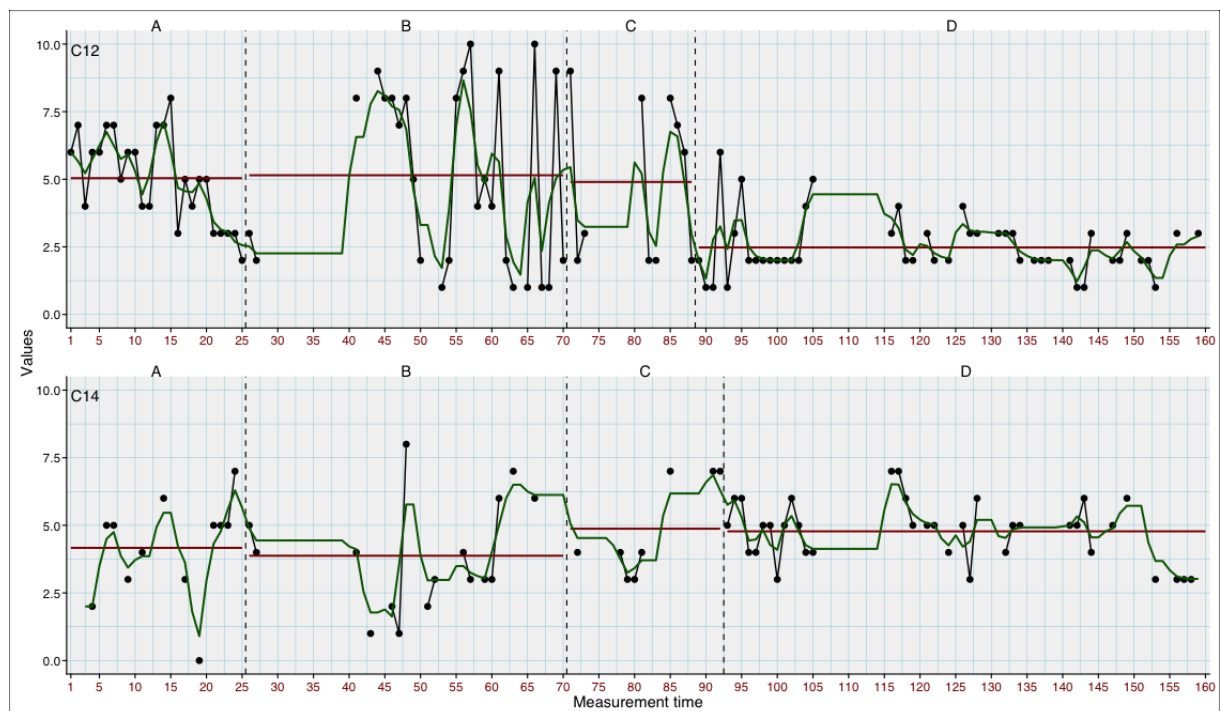
Visuelle Darstellung für die Fälle C7 & C9



Anmerkung: A = Baseline-Phase, B = 1. Stufe, C = 2. Stufe, D = 3. Stufe, Measurement time = Messzeitpunkt, Values = Wert

Abbildung 6

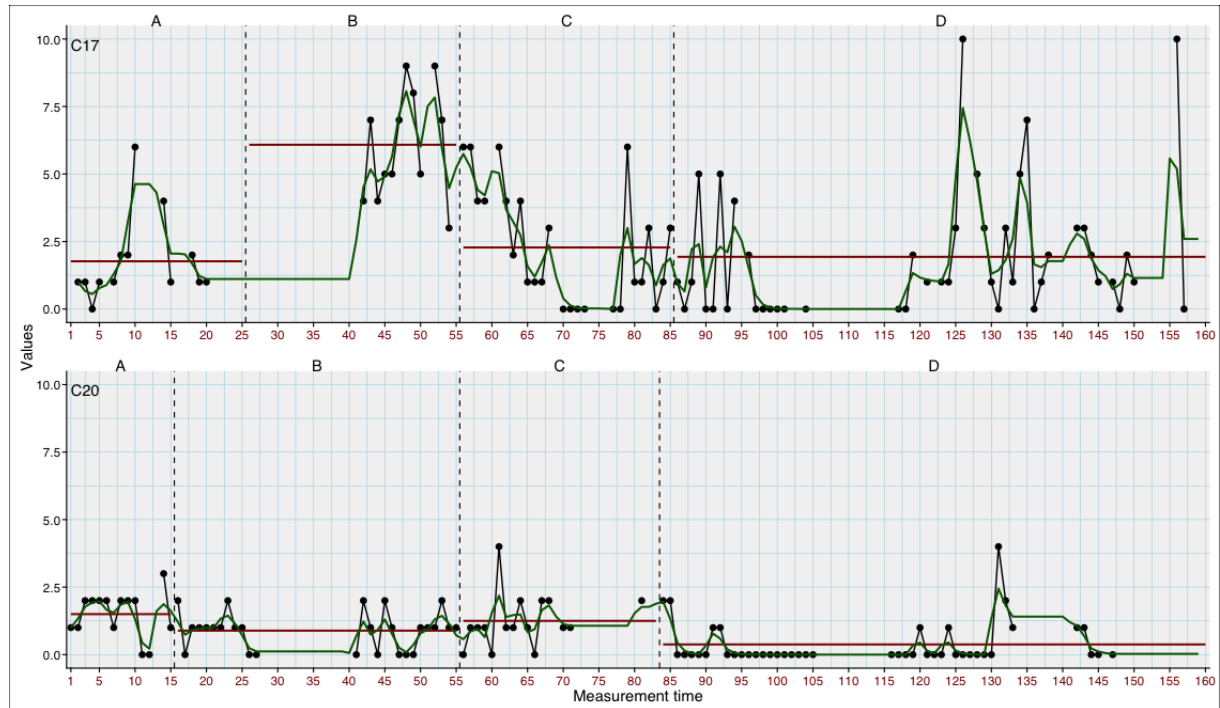
Visuelle Darstellung für die Fälle C12 & C14



Anmerkung: A = Baseline-Phase, B = 1. Stufe, C = 2. Stufe, D = 3. Stufe, Measurement time = Messzeitpunkt, Values = Werte

Abbildung 7

Visuelle Darstellung für die Fälle C17 & C20



Anmerkung: A = Baseline-Phase, B = 1. Stufe, C = 2. Stufe, D = 3. Stufe, Measurement time = Messzeitpunkt, Values = Werte

5.4.3 Diskussion

Nach Betrachtung der Ergebnisse wird offensichtlich, dass die Hälfte der Schüler*innen von der Förderung auf der dritten Stufe profitieren. Für ein Kind (C1) hat die Förderung nicht den gewünschten Effekt. Hier verstärkt sich das disruptive Verhalten entgegen den Erwartungen. Nach Betrachtung des Plots für das beschriebene Kind (C1) wird deutlich, dass das disruptive Verhalten sich lediglich in der ersten Phase nach der Baseline-Phase (Stufe 1) verringert. Sowohl die zweite Stufe als auch die dritte Stufe führen zu sehr variablen Verhaltensänderungen und lassen keine stabile Verhaltensveränderung erkennen. Das Verhalten wird zumeist zwischen 2 und 5 Punkten eingeschätzt und verstärkt sich in den Phasen 2 und 3 im Mittelwert. Die übrigen Daten sind erwartbar und lassen einen Interventionseffekt für die Hälfte der Gruppe erkennen. Wichtig bei der Betrachtung der Daten ist jedoch, dass die Stufe 3 eine Kombination verschiedener Interventionen darstellt: Hier werden das GBG, die DBRC, das SCEP (und die elternbezogenen Interventionen) gleichzeitig eingesetzt. Zudem ist das SCEP keine direkte Schüler*innen-Intervention, sondern verstärkt auf das Lehrkraftverhalten und unterrichtliche Adaptionen bezogen (Hanisch et al., 2020; Hanisch et al., 2018). Gemessen wird demnach ein auch ein indirekter Effekt bzw. *Spillover-Effekt* (z.B. Karatas et al., 2023) der Intervention(en),

welcher mitunter erschwert nachzuweisen ist. Welcher Bestandteil der Förderung hier nun explizit wirkt und welche Effekte die einzelnen Interventionen nach sich ziehen, ist nicht ermittelbar. Vielmehr sind diese Ergebnisse als Gesamtinterventionseffekte anzusehen.

5.4.4 Ausblick

Als Ausblick auf die vorliegende Untersuchung zur Wirksamkeit von *Multimo* auf Stufe 3 erscheint die Tatsache, dass es unerlässlich bleibt, die erzielten Effekte der Interventionen in Form weiterer Untersuchungen zu überprüfen. Aufbauend hierauf kann eine qualitative Erhebung in Form einer Interviewstudie wertvolle Einblicke bieten. Solch eine Studie würde es ermöglichen, tiefer in die individuellen Erfahrungen und Wahrnehmungen der beteiligten Lehrkräfte einzutauchen, was wiederum dazu beitragen kann, die quantitativen Ergebnisse zu ergänzen und zu erklären. In Ergänzung zu den Erkenntnissen dieser Studie erscheint es zudem sinnvoll, spezifische Komponenten des Förderansatzes wie die DBRC oder das GBG in zukünftigen Wirksamkeitsstudien genauer zu betrachten. Durch solche spezifischen Untersuchungen könnten die Wirkmechanismen der einzelnen Interventionen detaillierter identifiziert und deren Effizienz und Effektivität weiter optimiert werden.

5.5 Qualitative Analysen der Stufe 3: Das SCEP-Coaching (Beitrag 3)¹⁹

Anknüpfend an diese Befunde lassen sich die Ergebnisse des dritten Beitrags dieser kumulativen Dissertation anführen. In diesem Beitrag steht die Überprüfung des Förderansatzes in seiner Gesamtheit im Vordergrund. Hierzu wurden die bereits berichteten Einzelfallanalysen auf Stufe 1 und 2 samt ergänzender Interviewstudie zur Stufe 3 miteinander in Verbindung gebracht und die Ergebnisse gemeinsam diskutiert. Der Fokus liegt im Folgenden auf den Ergebnissen der Interviewstudie zur Stufe 3, da die anderen Ergebnisse bereits vorgestellt wurden (Beitrag 2).

5.5.1 Methodik

Für die qualitativen Interviews im Rahmen der Stufe 3 und des SCEP konnten sieben Grundschullehrerinnen befragt werden. Als Teilnahmevoraussetzung galt, dass sie zwischen

¹⁹ Nitz, J., Eiben, K., Hanisch, C., Hagen, T., Krull, J., Verbeck, L., Rauterkus, H., & Hennemann, T. (2024, im Druck). Mehrstufige Förderung in der inklusiven Grundschule – erste qualitative und quantitative Befunde zur Wirksamkeit und Umsetzung des Förderansatzes Multimo. In J. König, C. Hanisch, P. Hanke, T. Hennemann, K. Kaspar, M. Martens & S. Strauß (Hrsg.), *Teachers and their teaching matters – Auf die Lehrperson und ihren Unterricht kommt es an. 10 Jahre empirische Forschung im IZeF der Universität zu Köln*. Waxmann.

September und Dezember 2021 an einem SCEP mit einem Zielkind²⁰ teilgenommen und dieses abgeschlossen hatten. Die demografischen Daten der interviewten Lehrerinnen zeigen, dass die meisten von ihnen im Alter zwischen 31 und 40 Jahren (42 %) waren, gefolgt von den Altersgruppen 41 bis 50 (29 %) und 51 bis 60 (29 %). Ihre Berufserfahrung lag durchschnittlich bei 18 Jahren, mit einer Spanne von 14 bis 25 Jahren. Unter den Teilnehmenden befanden sich fünf Teilzeit- und zwei Vollzeitlehrkräfte. Die durchschnittliche Anzahl an Coachingsitzungen auf Stufe 3 pro Lehrkraft betrug 7,57, wobei die Anzahl der Sitzungen zwischen 6 und 8 variierte. In einigen Fällen wurden die Coachingsitzungen online durchgeführt.

5.5.2 Durchführung

Die Untersuchung der dritten Stufe bzw. des SCEP erfolgte durch halbstrukturierte Leitfadeninterviews, um die subjektive Perspektive der Lehrkräfte zu erfassen und deren subjektive Wirklichkeiten zu rekonstruieren. Der Interviewleitfaden basiert auf dem Wirkfaktorenmodell von Grawe (2000), das ursprünglich für die Psychotherapie entwickelt und später auf Beratung und Coaching angewandt wurde. Dieses Modell identifiziert Schlüsselfaktoren, die die Wirksamkeit dieser Interventionen beeinflussen, wie beispielsweise positive Beziehungen und Ressourcenaktivierung. Die Interviews wurden zwischen Februar und März 2022 durchgeführt, etwa 11 Wochen nach Ende der Interventionsphase der dritten Stufe. Insgesamt fanden Interviews mit sieben Lehrkräften statt, die durchschnittlich jeweils 29,4 Minuten dauerten (mit einer Spanne von 20 bis 42 Minuten). Sie wurden entweder telefonisch oder über die Online-Plattform *Zoom*²¹ von einer Projektmitarbeiterin und zwei am Projekt beteiligten Studierenden geführt. Alle Teilnehmerinnen hatten zuvor ihr schriftliches Einverständnis zur Durchführung und anonymisierten Verwendung der Interviewdaten gegeben. Die Audiodaten wurden gemäß dem semantisch-inhaltlichen Transkriptionssystem von Dresing und Pehl (2020) wörtlich und nach einfachen Transkriptionsregeln transkribiert.

5.5.3 Ergebnisse

Zusammenfassend zeigt sich, dass die Stufe 3 aus Sicht der Lehrkräfte als sehr positiv und wirkungsvoll wahrgenommen wird, wobei Veränderungen im Verhalten der Lehrkräfte, verbesserte Beziehungen zu den Schüler*innen, und eine gesteigerte Kompetenzwahrnehmung

²⁰ Das Kind, das während des Coachings von der Lehrkraft ausgewählt wird, um das Verhalten des Kindes (und der Lehrkraft) näher zu verstehen.

²¹ <https://zoom.us>

als Schlüsselemente erwähnt wurden. Die Qualität der Beziehung zum Coach sowie die Fähigkeit zur Selbstreflexion und Problembewältigung werden als wesentliche Wirkfaktoren identifiziert. Folgende Punkte lassen sich identifizieren:

1. **Positive Bewertung des Coachings:** Lehrkräfte beschreiben das SCEP durchweg positiv, etwa als „erfolgreich“, „gewinnbringend“ oder „effektiv“.
2. **Veränderung im Lehrkraftverhalten:** Ein Großteil der Äußerungen (67 %) hebt Veränderungen im Verhalten der Lehrkräfte und wiederholte Selbstreflexionen hervor.
3. **Verbesserte Beziehung und gesteigerte Kompetenz:** Fünf von sieben Lehrkräften berichten von einer verbesserten Lehrkraft-Kind-Beziehung und einem gesteigerten Gefühl eigener Kompetenz, wie in den Zitaten deutlich wird.
4. **Wirkfaktoren im Coachingprozess:** Die am häufigsten genannten Wirkfaktoren sind „Aktive Hilfe zur Problembewältigung“ und „Motivationale Klärung“, welche 81 % aller Textsegmente ausmachen. Die Lehrkräfte profitierten besonders von der Erarbeitung konkreter Strategien und Methoden.
5. **Beziehungsqualität als Schlüsselfaktor:** Alle Lehrkräfte betonen die Wichtigkeit einer vertrauensvollen Beziehung zum Coach für einen effektiven Coachingprozess, wobei die (zusätzliche) Fachkompetenz der Coaches und ein „neutraler Blick von außen“ als besonders relevant hervorgehoben werden.
6. **Verändertes Verständnis für problematische Situationen:** Mehr Verständnis für und eine veränderte Sicht auf problematische Situationen mit dem Zielkind werden häufig genannt, was oft mit einem selbstreflektierenden Blick auf das eigene Verhalten verknüpft ist.

5.5.4 Ausblick

Die dritte Stufe wird von den Lehrkräften entsprechend der Ergebnisse als effektiv eingeschätzt. Dies steht in Einklang mit einer früheren Wirksamkeitsstudie von [Hanisch et al. \(2020\)](#), die eine Abnahme des externalisierenden Verhaltens von Schüler*innen und eine Zunahme des Selbstwirksamkeitserlebens bei Lehrkräften feststellte. Darüber hinaus lassen sich damit die positiven Effekte der dritten Stufe im Bereich des disruptiven Verhaltens weiter erklären.

Die in den Interviews geäußerten potenziellen Wirkprozesse lassen sich hauptsächlich den drei Wirkfaktoren Problembewältigung, motivationale Klärung und Beziehung zuordnen. Besonders hervorgehoben wird die Vermittlung konkreter Interventions- und Umsetzungsmöglichkeiten, was mit Erkenntnissen aus Übersichtsarbeiten zu wirksamen

Schulinterventionen bei externalisierendem Problemverhalten übereinstimmt (z. B. Fabiano & Pyle, 2018; Waschbusch et al., 2018). Ebenso wird die hohe Bedeutung der Beziehungsqualität im Coachingprozess betont, was durch Studien aus der Psychotherapie- und Coachingforschung gestützt wird (z. B. Greif, 2008). Dennoch sind die Veränderungsprozesse auf der Stufe 3 nicht ausschließlich auf einzelne Wirkfaktoren zurückzuführen, sondern eher als Ergebnis eines komplexen Interaktionsprozesses zwischen allen Beteiligten – Lehrkräften, Coaches und Zielkindern – sowie deren Kontextfaktoren zu verstehen.

Dies Ergebnisse dieser Erhebung sind ein weiterer Beleg für die positive Erfahrung und Wirkweise einer Intervention und Stufe des *Multimo*-Ansatzes. Ergänzend zu diesen Ergebnissen sollen im Folgenden die eingesetzten Interventionen auf der Stufe 1 und der Stufe 2 einer isolierten Wirksamkeitsüberprüfung unterzogen werden.

5.6 Wirksamkeit des Good Behavior Games auf Stufe 1 (Beitrag 4)²²

Im Sinne der bereits erwähnten notwendigen Überprüfung einzelner Facetten bzw. Interventionen des *Multimo*-Ansatzes wurde somit die 1. Stufe hinsichtlich des Rückgangs disruptiver Verhaltensweisen und dem Aufbau lernbezogener Verhaltensweisen überprüft. Dies erfolgte in Form einer Einzelfallstudie mit einer Grundschule in Nordrhein-Westfalen. Die zentrale Forschungsfrage lautet, ob der regelmäßige Einsatz des GBG das Verhalten der Schüler*innen im Unterricht positiv beeinflussen kann. Die Ergebnisse dieser Studie liefern weitere Erkenntnisse bezüglich der Wirksamkeit der ersten Förderstufe und der hier implementierten Intervention, dem GBG.

5.6.1 Methodik

Die angesprochene Studie wurde als kontrollierte Einzelfallstudie in allen vier dritten Klassen einer Grundschule durchgeführt. Für die Stichprobe wurden von den Klassenlehrkräften jeweils fünf Schüler*innen ausgewählt, welche die stärksten externalisierenden Verhaltensprobleme aufwiesen, was zu einer Gesamtzahl von 20 teilnehmenden Kindern führte. Die Auswahlkriterien für die Kinder umfasste externalisierendes Verhalten und wurde durch den Einsatz der ITRF (Casale et al., 2019) erhoben. Die Stichprobe bestand aus einem Mädchen und 19 Jungen im Alter von acht bis neun Jahren. Von allen Kindern hatten 15 Kinder einen

²² Hagen, T., Nitz, J., Brack, F., Hövel, D. C., & Hennemann, T. (2023). Effekte des Good Behavior Game bei Grundschüler_innen mit externalisierenden Verhaltensproblemen. *Lernen und Lernstörungen*, 1-14. <https://doi.org/10.1024/2235-0977/a000410>

Migrationshintergrund. Es wurde zudem festgestellt, dass zwischen den Kindern Unterschiede in Bezug auf die Intensität der externalisierenden Verhaltensprobleme bestanden.

5.6.2 Durchführung

Das GBG wurde in vier Klassen fünf Tage pro Woche umgesetzt. Lehrkräfte erhielten hierzu eine vierstündige Schulung zu Präventionsstrategien und zum GBG, inklusive einer Einführung in die DBRs. Regelverstöße wurden nach jedem Einsatz des GBG erfasst und die Erhebungsbögen wurden anknüpfend hieran kontinuierlich zweimal täglich für ausgewählte Zielkinder ausgefüllt. Zusätzlich unterstützten regelmäßige Treffen mit Projektmitarbeiter*innen und Checklisten die Lehrkräfte bei der Durchführung des GBG. In den Klassen wurden einheitliche Regeln wie "Still sitzen" und "Leise arbeiten" angewandt und sowohl lernförderliches als auch störendes Schüler*innenverhalten wurde kontinuierlich und systematisch mittels DBR erfasst.

5.6.3 Auswertung & Analyse

Um die Auswirkungen des GBG auf das störende Verhalten (SV) und das lernförderliche Verhalten (LV) zu untersuchen, wurden verschiedene Analysemethoden angewandt. Auf Fallebene wurde der NAP verwendet, um die Daten aus Baseline und Interventionsphase miteinander zu vergleichen. Für die fallübergreifende Analyse kam eine Mehrebenen-Analyse zum Einsatz, um Level-, Trend- und Slope-Effekte zu berechnen (Wilbert et al., 2022). Die Mehrebenen-Analysen wurden als Random Intercept- und Random Slope-Modelle berechnet, wobei die Effekte als *Fixed-* und *Random-Effects* behandelt und mit *R-Studio* (Posit-Team, 2022) und dem Paket *scan* (Wilbert & Lueke, 2022) ausgewertet wurden.

5.6.4 Ergebnisse

Die Ergebnisse zeigen, dass die Effekte des GBG auf die Variable SV und die Variable LV der Schüler*innen unterschiedlich groß waren. Der NAP als Maß für die Effektstärke ergibt während der GBG-Phase eine Spanne von 32,33 % bis 97,45 % für SV und von 31,82 % bis 90,32 % für LV. Bei vier Kindern wurde keine Verbesserung im SV festgestellt, während neun Kinder mindestens mittlere Effekte zeigten. Die Auswertung der Variable LV zeigt, dass 13 Kinder profitierten (mindestens mittlere Effekte), aber für vier Kinder keine positive Tendenz in den Daten nachweisbar ist. Die allgemeine Verhaltensentwicklung über den Schultag hinweg zeigt ebenfalls eine große Variabilität im NAP, mit Werten zwischen 17,17 % und 91,11 % für SV und zwischen 17,30 % und 95,76 % für LV. Die Variable SV wies bei 17 Kindern

mindestens mittlere Effekte auf, während bei drei Kindern keine Verbesserung erkennbar ist. In Bezug auf LV zeigten 14 Kinder mindestens mittlere Effekte, aber bei einem Kind gab es keine positive Entwicklung.

Die Mehrebenen-Analysen, die durchgeführt wurden, um die Gesamteffekte des GBG auf SV und LV unter Berücksichtigung aller 20 Einzelfälle zu ermitteln, zeigen ähnliche Ergebnisse. Hier waren signifikante Level-Effekte für SV und LV nachweisbar, was bedeutet, dass zu Beginn der Interventionsphase von einer sofortigen und konstanten Verhaltensänderung ausgegangen werden kann. Im Durchschnitt sank das störende Verhalten um 0,404 Punkte während der GBG-Phase und um 0,401 Punkte über den gesamten Schultag. Das lernförderliche Verhalten stieg um 0,432 Punkte in der GBG-Phase und um 0,442 Punkte außerhalb der Spielphasen. Die Slope-Effekte, die eine kontinuierliche Entwicklung des Verhaltens ab dem Beginn der Intervention darstellen, erreichten keine statistische Signifikanz.

5.6.5 Ausblick

Die Studie zielte darauf ab, die Effektivität des GBG bei Grundschüler*innen mit externalisierenden Verhaltensproblemen zu überprüfen. Die Ergebnisse zeigen positive Effekte des GBG auf die Variablen SV und LV, sowohl während der GBG-Phasen als auch über den gesamten Schultag hinweg. Insgesamt profitierten mehr Schüler*innen im Kontext des lernförderlichen Verhaltens als hinsichtlich des störenden Verhaltens von der Intervention. Die Studie zeigt, dass das GBG sowohl im SV als auch im LV positive Effekte bei einem Großteil der Schüler*innen hat, insbesondere außerhalb der Spielphasen. Die Ergebnisse bestätigen und erweitern außerdem frühere Forschungen und deuten darauf hin, dass das GBG geeignet ist, um Verhaltensprobleme im Unterricht zu reduzieren (Flower et al., 2014; Leidig et al., 2022; Pennington & McComas, 2017). Die Intervention erscheint demnach wirksam, um störendes Verhalten in den Spielsituationen zu reduzieren und erzeugt gleichermaßen einen Effekt, der außerhalb der Spielsituationen anhält. Trotz gewisser Limitationen (z. B. große Variabilität der einzelnen Fälle) sind auf dieser Grundlage belastbare Aussagen möglich, die die Wirksamkeit der Intervention unterstreichen. Auch erfüllt sie die notwendigen Charakteristika einer universellen Fördermaßnahme im Sinne gestufter Förderung. Eine vergleichbare Studie von Leidig et al. (2022) unterstreicht diese Ergebnisse für den deutschsprachigen Raum. Anknüpfend hieran erscheint die Überprüfung der DBRC (Stufe 2) darüber hinaus notwendig, um valide Aussagen über den gesamten Förderansatz *Multimo* und die damit einhergehenden Interventionen treffen zu können.

5.7 Wirksamkeit der Daily Behavior Report Cards (Beitrag 5)²³

Wie die Ergebnisse aus den vorherigen Studien zeigen, kann eine weitere Überprüfung der Effekte der eingesetzten Intervention der zweiten Stufe (DBRC) zusätzlich Aufschluss über die Gesamteffektivität des Förderansatzes geben. Hierzu wurde ergänzend eine Einzelfallstudie mit folgender Hauptfragestellung geplant: Führt der Einsatz der DBRC in einer (inkluisiven) Grundschule zu einer Reduktion des disruptiven Verhaltens von Kindern mit externalisierendem Verhalten?

5.7.1 Methodik

Die Studie wurde als Einzelfallstudie im multiplen AB-Design geplant und umgesetzt. Zehn Zweitklässler*innen aus drei inklusiven Klassen einer inklusiven Grundschule in Nordrhein-Westfalen nahmen an der Studie teil. 60 % der Stichprobe waren Jungen und das Durchschnittsalter lag bei 7,70 Jahren. Die Schule lag am Stadtrand einer Großstadt in Nordrhein-Westfalen. Die Lehrkräfte wurden zu Beginn der Studie beauftragt, die auffälligsten Schüler*innen der Klasse zu bestimmen. Alle zehn Schüler*innen zeigten laut Lehrkraftbewertung störendes Verhalten, gemessen mit der ITRF (Casale et al., 2018).

5.7.2 Durchführung

Bezüglich der Implementation der DBRC erhielten die Lehrkräfte nach der Baseline-Phase eine halbtägige Schulung, um die Interventionsziele festzulegen und die konkrete Anwendung der Intervention im Schulalltag zu diskutieren. Regelmäßige Treffen zwischen dem Projektteam und den Lehrkräften sollten die Implementationsprozesse erleichtern. Die Implementationsgüte konnte daraufhin als hoch bewertet werden. Die Baseline-Phase umfasste 35-45 Messzeitpunkte, in denen zuvor festgelegte Verhaltensweisen der Kinder beobachtet und mittels DBR beurteilt wurden. Die Interventionsphase bestand ebenfalls aus 35-45 Messzeitpunkten. Erhoben wurden zwei verschiedene Variablen mit Hilfe der DBR: Allgemeines Störverhalten (GDB) und Spezifisches Zielverhalten (SGB), die von 0 (niedrig) bis 5 (hoch) eingeschätzt werden konnten. Die Variable GDB umfasst allgemein störendes Verhalten im Klassenraum, während die Variable SGB individuell auffälliges Verhalten umfasst, das vor der Intervention

²³ Nitz, J., Volpe, R., Hagen, T., Krull, J., Hennemann, T., Hanisch, C. (voraussichtlich 2024). The Effects of the Daily Behavior Report Cards (DBRC) on the Disruptive Behavior and Specific Goal Behavior of Elementary School Children: A Multiple Baseline Design Study: *Dieser Beitrag ist eingereicht und befindet sich im Begutachtungsverfahren.*

mit den Lehrkräften besprochen und als Zielverhalten für die DBRC festgelegt wurde. Vor der Interventionsphase wurden die Lehrkräfte in der Handhabung der DBR geschult, um das Schüler*innenverhalten präzise dokumentieren und die Intervention effektiv implementieren zu können.

5.7.3 Auswertung & Analyse

Die Daten wurden mit *R-Studio* (Posit-Team, 2022) und dem Paket *scan* (Wilbert & Lueke, 2022) ausgewertet. Jeder Fall wurde zunächst individuell untersucht, wobei Mittelwerte und der NAP für die Bewertung der Verhaltensentwicklung zwischen den Phasen berechnet wurden. Ergänzend wurden regressionsbasierte Methoden angewandt, um die Effekte der Intervention auf das Verhalten zu ermitteln, einschließlich Slope-, Trend- und Level-Effekten. Diese Effekte beschreiben sowohl die Gesamtveränderung der Stichprobe als auch die Entwicklung der einzelnen Kinder. Dabei wurden Modelle samt *Random-* und *Fixed-effects* berechnet, um die Unterschiede zwischen den Individuen zu analysieren.

5.7.4 Ergebnisse

Jeder Fall (d. h. jedes Kind) zeigte vor der Intervention ein hohes Maß an Störverhalten in den Ergebnissen der ITRF. Die mittleren Werte der OD- und APD-Skalen zeigten deutliche Auffälligkeiten mit entsprechenden Cut-off-Werten. Die Analyse des NAP ergab bei 40 % der Fälle einen starken und bei 60 % einen moderaten Effekt auf die Variable GDB. Ähnlich Effekte konnten für die Variabel SGB festgestellt werden. Hier zeigte sich in 40 % der Fälle ein starker und in 60 % ein moderater Effekt. Der NAP eines Falls war nicht signifikant. In den Gruppenanalysen wurde ebenfalls ein signifikanter Rückgang des externalisierenden Verhaltens festgestellt. Drei Modelle wurden für eine detaillierte Analyse entwickelt: Modell 1 zeigte, dass das Störverhalten während der B-Phase signifikant verändert wurde (Level-Effekt), jedoch ohne signifikante Trend- oder Slope-Effekte über die Zeit. Die zufälligen Effekte deuteten auf geringe Unterschiede zwischen den Fällen hin. Modell 2 ergab eine signifikante Veränderung in der Variable SGB während der B-Phase (Level-Effekt) ohne signifikante Trend- oder Slope-Effekte. Die Modelle zeigen zudem signifikante Unterschiede zwischen den Fällen. Ergänzend bestätigt Modell 3, dass die beiden erhobenen Variablen miteinander korrelieren. Zusammenfassend deuten die Modelle darauf hin, dass die Intervention signifikante Veränderungen im Verhalten bewirkte, mit geringen Unterschieden zwischen den Fällen und einem signifikanten Einfluss des spezifischen Zielverhaltens auf das allgemeine Störverhalten.

5.7.5 Ausblick

Diese Studie unterstreicht die Wirksamkeit von DBRC bei der Reduzierung von externalisierendem Verhalten. Methodisch zeichnet sich die Studie durch eine hohe Datenqualität aus, bedingt durch zahlreiche Messzeitpunkte. Die Implementationsqualität variierte zwischen den Klassen, wobei Klasse A die niedrigste und Klassen B und C höhere Umsetzungstreue zeigten. Die Beziehung zwischen der Implementationsqualität und dem Erfolg der Intervention ist jedoch nicht eindeutig.

Die zwei erhobenen Variablen können zudem Aufschluss über den Zusammenhang zwischen dem spezifischen Zielverhalten, welches als Ziel der Intervention definiert ist, und dem allgemeinen disruptiven Verhalten geben. Für weitere Studien können hier Forschungsdesigns überlegt werden, die diesen Zusammenhang ergänzend in anderer Form untersuchen. Für die Implementation gestufter Fördersysteme in der (inkluisiven) Grundschule in Deutschland lässt sich jedoch bereits festhalten, dass die Intervention wirksam ist und externalisierendes Verhalten reduzieren kann. Sie ist jedoch zeitaufwändig und bedarf der individuellen Anpassung, sodass sie als Intervention auf der selektiven und indizierten Ebene eingesetzt werden kann, allerdings keine Intervention der universellen Präventionsebene darstellt.

6 Implementation gestufter Fördersysteme in Deutschland

Die vorgestellten Forschungsergebnisse zum *Multimo*-Ansatz legen eine umfassende Wirksamkeit in Bezug auf die Reduktion von externalisierendem Verhalten nahe. Diese Effektivität zeigt sich entlang aller drei Stufen. Trotz der positiven Befunde bleibt jedoch eine Reihe von unklaren Einflüssen und Bedingungsfaktoren bestehen. Es wurde z. B. festgestellt, dass nicht alle Schüler*innen gleichermaßen von dem Ansatz profitieren. Die Gründe hierfür können vielfältig sein. Denkbar ist ein Einfluss bestimmter Rahmenbedingungen der Schulen oder individueller Bedürfnisse der Schüler*innen, die eine Rolle im Zuge der Umsetzbarkeit spielen. Zudem deuten die *Random-Effects* in den statistischen Analysen darauf hin, dass es unterschiedliche Fall- und Verhaltensverläufe gibt, was die Komplexität der Einflussfaktoren auch statistisch unterstreicht (Beitrag 2, 4, 5).

Vor diesem Hintergrund erscheint es im Folgenden sinnvoll, ergänzend die Implementation des gesamten Ansatzes und der genutzten Methoden und Interventionen näher zu untersuchen. Dies erscheint insbesondere gewinnbringend, da die Implementationsforschung im Kontext von Einzelinterventionen in der vergangenen Zeit belegen konnte, dass verschiedene Faktoren den Implementations- und Fördererfolg von MTSS maßgeblich beeinflussen (Fox et al., 2021). In diesem Zusammenhang hat sich ein breites Forschungsfeld entwickelt, welches sich unterschiedlicher Verfahren bedient, um kritische Implementationsfaktoren zu identifizieren. Ergänzend dazukönnen Studien belegen, dass die Implementationsqualität einen wichtigen Faktor für den Erfolg eines MTSS und einzelner pädagogischer Interventionen darstellen kann (Berkeley et al., 2009; Tyre et al., 2019). Hierzu zählen folgende Faktoren, die mit den jeweils beschriebenen Verfahren erhoben werden können (Grimshaw et al., 2006; Michie et al., 2009; Petermann, 2014):

1. **Akzeptanz (Acceptability):** Akzeptanz beschreibt die Zufriedenheit und qualitative Bewertung einer Maßnahme und steht in direkter Verbindung mit dem erlangten Wissen und den Erfahrungen mit einer Intervention. Die Akzeptanz von Implementationsprozessen kann durch Umfragen, qualitative oder semi-strukturierte Interviews sowie die Analyse administrativer Daten beurteilt werden.

2. **Übernahme (Adoption):** Übernahme bezieht sich auf die Absicht oder Entscheidung, eine Intervention in der Praxis einzusetzen. Zur Bewertung der Übernahme von Implementationsprozessen werden ebenfalls Umfragen und qualitative oder semi-strukturierte Interviews eingesetzt. Zusätzlich können die Beobachtung und die Auswertung administrativer Daten hilfreich sein.
3. **Angemessenheit (Appropriateness):** Dies beinhaltet die wahrgenommene Passung, Aktualität und Kompatibilität einer Intervention oder eines Verfahrens für ein spezifisches Setting oder eine*n spezifische*n Nutzer*in. Angemessenheit ist der Akzeptanz ähnlich, jedoch mit der Möglichkeit, dass eine Intervention als angemessen, aber nicht als akzeptabel empfunden wird und umgekehrt. Die Angemessenheit eines Implementationsprozesses lässt sich durch Umfragen, qualitative oder semi-strukturierte Interviews und Fokusgruppen ermitteln.
4. **Machbarkeit (Feasibility):** Machbarkeit beschreibt das Ausmaß, in dem eine neue Intervention in einem bestimmten Setting erfolgreich eingesetzt werden kann. Es kann rückwirkend als Erklärung für den Erfolg oder Misserfolg einer Implementierung herangezogen werden. Um die Machbarkeit jener Interventionen zu beurteilen, werden Umfragen und die Analyse administrativer Daten verwendet.
5. **Wiedergabetreue (Fidelity):** Die Wiedergabetreue, also das Ausmaß, in dem eine Implementation wie geplant umgesetzt wird, kann durch Beobachtungen oder Checklisten sowie Selbstberichte erfasst werden.
6. **Implementationskosten (Costs):** Dieser Faktor beschreibt die Kosten einer Implementation. Die Kosten der Implementation werden hauptsächlich durch die Analyse administrativer Daten erhoben.
7. **Durchdringung (Penetration):** Durchdringung bezieht sich auf das Ausmaß, in dem eine Maßnahme in eine Institution integriert wird, und kann durch das Verhältnis der Nutzung einer Intervention im Vergleich zur Gesamtzahl potenzieller Nutzer*innen erfasst werden. Zur Beurteilung der Durchdringung einer Implementation innerhalb einer Organisation können darüber hinaus Fallbesprechungen und Checklisten hilfreich sein.
8. **Nachhaltigkeit (Sustainability):** Die Nachhaltigkeit, also das Fortbestehen einer Implementation über einen längeren Zeitraum, wird durch Fallbesprechungen, semi-strukturierte Interviews, Fragebögen und Checklisten bewertet.

Die Terminologie der Implementationsforschung kann jedoch auch uneinheitlich sein, insbesondere bezüglich den Faktoren Akzeptanz, Angemessenheit und Machbarkeit. Verschiedene Studien verwenden unterschiedliche Begrifflichkeiten für vergleichbare Faktoren. Ergänzend zu den hier aufgelisteten Befunden konnte gezeigt werden, dass auch das Wissen und die Einstellung der Lehrkräfte sowie die verfügbaren Ressourcen und die organisatorischen Rahmenbedingungen entscheidend für die erfolgreiche Umsetzung von MTSS sind (Fox et al., 2021). Zusammenfassend kann demnach die Relevanz der Untersuchung der Implementationsprozesse im deutschsprachigen Raum abgeleitet werden. Die Orientierung an der US-amerikanischen Forschung zu diesem Themenbereich kann hierbei hilfreich sein, weswegen diese im Folgenden weiter ausgeführt wird.

6.1 Befunde US-amerikanischer Implementationsforschung

Die nachhaltige Implementation gestufter Fördersysteme stellt eine erhebliche und oft unterschätzte schul- und bildungspolitische Herausforderung dar (Berkeley et al., 2009; Fox et al., 2021; Don Kincaid, Karen Childs, et al., 2016). Implementation wird in diesem Zusammenhang verstanden als die langfristige Beibehaltung einer bestimmten pädagogischen Praxis und dem gleichzeitigen Erreichen der angestrebten Verhaltensänderungen auf Schüler*innenebene (Fullan, 2015). Ein wichtiger Aspekt dabei ist die kontinuierliche Überprüfung der Praxis durch Beobachtung und Bewertung (Petermann, 2014). Es wird schnell offensichtlich, dass für einen komplexen Vorgang wie diesen, eine detaillierte Betrachtung der beeinflussenden Faktoren eine hilfreiche und gleichzeitig notwendige Form von Grundlagenforschung darstellt.

Kent McIntosh et al. (2009) formulieren prominente Grundsätze für die erfolgreiche Implementation von MTSS. Zunächst ist der Begriff des *contextual fit* von Bedeutung, welcher die Notwendigkeit betont, das Fördersystem an die spezifischen Anforderungen und Gegebenheiten einer Schule anzupassen. Es geht darum, ein System zu schaffen, das sowohl zu den Bedürfnissen der Schüler*innen als auch zu den Rahmenbedingungen der Schule passt. Ein weiterer wichtiger Aspekt ist umschrieben mit *priority*, also die Priorisierung von Maßnahmen, um die festgelegten Förderziele effektiv zu erreichen. Dies kann auch eine Neuausrichtung bestehender Maßnahmen erfordern, um sicherzustellen, dass die wichtigsten Ziele im Vordergrund stehen und effizient verfolgt werden. Die *effectiveness* des Fördersystems ist ebenfalls entscheidend. Hierbei geht es darum, wie wirksam das System bei der Erreichung seiner Ziele ist. Dieser Punkt ist eng mit dem Begriff der *efficiency* verknüpft, welcher das

Verhältnis von Aufwand und erzieltm Nutzen beleuchtet. Ein effizientes Fördersystem nutzt Ressourcen so, dass ein maximaler Nutzen bei möglichst geringem Aufwand erzielt wird. Schließlich spielt die *continuous regeneration* eine zentrale Rolle, womit die ständige Weiterentwicklung und Anpassung des Fördersystems gemeint ist. Durch die kontinuierliche Analyse von Daten sowie die darauf basierende Anpassung der Maßnahmen und des Förderansatzes werden Entscheidungsprozesse bezüglich der Förderstufen der Schüler*innen ständig optimiert. Diese kontinuierliche Neuausrichtung sorgt dafür, dass das Fördersystem dynamisch bleibt und sich an verändernde Bedingungen und Bedürfnisse anpassen kann. Das Literaturreview von [Fox et al. \(2021\)](#) formuliert aufbauend hierauf weitere Einflussfaktoren, die in [Tabelle 4](#) zusammengefasst sind.

Tabelle 4

Variablen gelingender Implementation von MTSS nach Fox et al. (2021)

Variable	Beschreibung
Resourcing	Ausstattung mit finanziellen, materiellen und personellen Ressourcen
Fidelity of Implementation	Genauigkeit der Umsetzung des Systems entsprechend seiner Konzeption
SWPBIS Team Function	Koordination und Umsetzung des Systems durch das Team
Leadership	Führungsrolle der Schulleitung samt Unterstützung
Data Collection, Collation & Utilization	Erhebung, Auswertung und Nutzung von Daten zur Bewertung und Verbesserung des Systems
Buy-in	Akzeptanz von Lehrkräften und anderen Personengruppen
Ongoing Training and Support	Fortlaufende Schulungen und Unterstützung für das Personal
School Characteristics	Einfluss von Schuleigenschaften wie Größe, Standort und Schüler*innendemografie auf das System
Personal Belief(s)	Einfluss persönlicher Überzeugungen des Personals auf die Akzeptanz und Umsetzung des Systems
External Expertise/Support	Bereitstellung von zusätzlichem Fachwissen und Perspektiven durch externe Expert*innen
Staff Understanding about SWPBIS	Verständnis & Wissen des Personals über die Praktiken und Ziele des Systems

Family, Student, or Community Involvement	Einbindung von Familie, Schüler*innen und der Gemeinde zur Entwicklung des Systems
Prioritization/Alignment with Current Practices	Ausrichtung des Fördersystems auf aktuelle Praktiken und Prozesse
Communication	Erfolgreiche, barrierearme Einbindung aller Beteiligten
Adaptation	Fähigkeit des Systems, sich an veränderte Bedürfnisse und Umstände anzupassen
Staff Turnover	Auswirkungen von Personalwechsel auf Kontinuität und fortlaufende Umsetzung des Systems
School Climate and Stress	Einfluss des Schulklimas und Belastungserleben auf die Umsetzung
Community of Practice	Gemeinsamer Austausch & gemeinsames Lernen
District Characteristics	Einfluss von Merkmalen des Schulbezirks auf Implementation und Anpassung des Systems
Involving New Personnel	Konzepte für die Fortführung und Anpassung des Systems mit neuen Mitarbeiter*innen
Effectiveness	Fortlaufende Überprüfung der Effektivität
Speed of Initial Implementation	Einfluss der Geschwindigkeit der Implementation auf den Effekt dieser

Anmerkung: eine Übersetzung der Begriffe findet sich in [Beitrag 6](#)

Diese Faktoren haben sich entlang der internationalen Forschung als relevante Einflussvariablen herausgestellt. Sie bieten einen Rahmen für zukünftige Implementationsprozesse und helfen eventuelle Anpassungsmöglichkeiten der Systembedingungen samt ihrer Ausgestaltung zu identifizieren. Auch die Entwicklung quantitativer Fragebögen zur Analyse von Implementationsprozessen im deutschen Bildungssystem ist auf Grundlage jener Faktoren möglich. In den USA werden hierfür Instrumente wie das *School-Wide Evaluation Tool* ([SET, Horner et al., 2016](#)), das *Academic and Behavior Response to Intervention School Assessment* ([ASA, Scott et al., 2013](#)) oder das *SWPBIS Tiered Fidelity Inventory* ([TFI, Algozzine et al., 2014](#)) eingesetzt. Diese Fragebögen sind nützlich, um die Qualität und Nachhaltigkeit von gestuften Fördersystemen zu erheben. Für den deutschsprachigen Raum gibt es keine adaptierten und evaluierten Instrumente. Hier besteht folglich ein weiteres Desiderat im Kontext der MTSS-Forschung in Deutschland.

6.2 Herausforderungen und Gelingensbedingungen der Implementation von MTSS in der Grundschule (Beitrag 6)²⁴

Um die Implementationsbedingungen und den Kontext dieser Faktoren näher zu untersuchen, wurde aufbauend auf den zuvor erwähnten Faktoren gelingender Implementation nach Fox et al. (2021) eine qualitative Interviewstudie mit Lehrkräften geplant. Teilnehmende der Studie waren Lehrkräfte, die bereits die Implementation des *Multimo*-Ansatzes durchlaufen haben. Sie konnten somit zu den verschiedenen Herausforderungen und Gelingensbedingungen aus ihrer Sicht befragt werden. Die Ergebnisse jener Studie können zur weiteren Kontextualisierung der zuvor beschriebenen Studien und ihrer Implikationen genutzt werden. Darüber hinaus ist die weitere Überführung jener herausgearbeiteter Implementationsfaktoren in ergänzende Forschungsvorhaben (wie z. B. die Entwicklung eines Instruments zur Erfassung der Implementationsgüte) denkbar.

6.2.1 Methodik

Diese qualitative Interviewstudie umfasst neun Lehrkräfte aus acht verschiedenen Grundschulen im Kreis Mettmann, die den Förderansatz *Multimo* implementiert haben. Das Durchschnittsalter der Teilnehmenden liegt bei 43,6 Jahren, mit einer Berufserfahrung zwischen unter zehn und über 21 Jahren.

6.2.2 Durchführung

Die Interviews wurden etwa 1,5 Jahre nach Beginn der Implementation (zwei Monate nach Beendigung des Projektes) mit dem Videokonferenzprogramm *Zoom*²⁵ durch einen Projektmitarbeiter durchgeführt. Für die Durchführung wurde ein Interviewleitfaden für Expert*inneninterviews eingesetzt, der mithilfe der *Critical Incident Technique* (CIT, Flanagan, 1954) umgesetzt wurde und im Kern die Implementationsfaktoren der Studie von Fox et al. (2021) in Form deduktiver Kategorien enthielt. Diese Technik eignet sich besonders zur Untersuchung von Gelingensbedingungen und Herausforderungen im Kontext der Implementationsforschung, indem sie strukturierte Interviews nutzt, um spezifische Ereignisse und deren Bewertung der Proband*innen zu ermitteln.

²⁴ Nitz, J., Niederelz, A., Hanisch, C., & Hennemann, T. (2024). Herausforderungen und Gelingensbedingungen erfolgreicher Implementation eines mehrstufigen Förderansatzes an Grundschulen. *Zeitschrift für Heilpädagogik*, 75(1), 4-17.

²⁵ <https://zoom.us>

6.2.3 Auswertung & Analyse

Die Interviews wurden mittels der zusammenfassenden Inhaltsanalyse ausgewertet (Mayring & Fenzl, 2019). Hierfür wurden die Interviewdaten nach der Transkription durch das Projektteam im Programm *QCAmap.org* kodiert (Fenzl & Mayring, 2017). Zunächst wurden aus 30 % des Materials induktive Kategorien abgeleitet und anschließend der gesamte Inhalt erneut kodiert, um sowohl deduktive als auch induktive Kategorien anzuwenden. Zwei Projektmitarbeiter*innen führten die Kodierung durch, wobei eine hohe Interrater-Reliabilität von Cohens $\kappa = .80$ erreicht wurde (Landis & Koch, 1977; McHugh, 2012). Der Kodierleitfaden umfasst vier Oberkategorien mit insgesamt 14 Unterkategorien, darunter 12 deduktive und zwei induktive Kategorien. Diese deduktiven Kategorien wurden auf Basis der zuvor vorgestellten Analyse von Fox et al. (2021) entwickelt.

6.2.4 Ergebnisse

Die Ergebnisse der Interviews offenbaren, die von den Lehrkräften identifizierten Herausforderungen und Gelingensbedingungen (Tabelle 5).

Tabelle 5

Herausforderungen und Gelingensbedingungen zur Implementation von MTSS

Herausforderungen	Gelingensbedingungen
Mangelnde personelle Ressourcen erschweren Implementation	→ Ein höherer Personalschlüssel und externe Unterstützung erleichtern Implementation
Bisherige Fortbildung und Coaching oftmals nicht ausreichend, um MTSS zu implementieren	→ Entwicklung effektiver bedarfsorientierter Fortbildungskonzepte
Mangel an Zeit im alltäglichen Schulleben und Unterricht	→ Schaffung der erforderlichen zeitlichen Ressourcen
Unterschiedliche Bezahlung fällt mit gleichen Aufgabengebieten zusammen	→ Einheitliche und angemessene Vergütung für vergleichbare Positionen
Zu große Klassen, um gestufte Förderung effektiv umzusetzen	→ Kleinere Klassengrößen können Förderprozesse erleichtern
Erhebung und Nutzung von Daten erfordert einen Anstieg an weiteren administrativen Aufgaben	→ Fortbildungen im Kontext von schulischer Diagnostik, um die Effizienz zu erhöhen

Mangelnde Kooperation mit den Eltern erschwert die Implementation gestufter Fördersysteme	→ Kontakt und Kooperation mit Eltern initiieren, auch durch externe Partner*innen
Häufiges Auftreten von Verhaltensauffälligkeiten im schulischen Kontext	→ Gezielte Fortbildungen und Fachpersonal bzgl. Umgang mit Verhaltensauffälligkeiten
Überlastung der Lehrkräfte	→ Entlastung der Lehrkräfte
Maßnahmen in gestuften Fördersystemen lassen zu wenig Flexibilität zu	→ Flexibilität in der Anwendung von Interventionen
Akzeptanzprobleme im Kollegium	→ Akzeptanz durch Fortbildungen und Nutzen
Fehlende Unterstützung der Schulleitung	→ Überzeugung und Unterstützung durch die Schulleitung

Wichtige Herausforderungen sind laut Lehrkräften die personelle Ressourcen, Fortbildungen, die Klassengröße, die Datenerhebung und -nutzung sowie bestehende Verhaltensauffälligkeiten und Unterrichtsstörungen. Die Rolle von Führung und Schulcharakteristika erscheint weniger wichtig, möglicherweise wurde dies jedoch beeinflusst durch die Projektbegleitung und freiwillige Teilnahme an dem Forschungsprojekt. Die Studie konnte die Implementationsfaktoren *District characteristics*, *Effectiveness* und *Speed of initial implementation* nicht replizieren. Dies liegt vermutlich hauptsächlich an den Unterschieden zwischen den Bildungssystemen der USA und Deutschlands sowie der Fokussierung auf Nordrhein-Westfalen. Zwei zusätzliche induktive Kategorien wurden jedoch identifiziert: *sprachliche Barrieren* sowie *Überlastung und Überforderung der Lehrkräfte*. Gestufte Fördersysteme sollten demnach aus Sicht der Lehrkräfte sprachsensibel und an schulische Bedürfnisse angepasst sein, wobei Flexibilität zur Berücksichtigung individueller Schüler*innenbedürfnisse wichtig ist. Beim Betrachten der Ergebnisse fällt zudem auf, dass die Herausforderungen und Gelingensbedingungen auf unterschiedlichen Ebenen des Bildungssystems anzusiedeln sind. Einige sind systemisch (bspw. *Klassengröße* oder die *personellen Ressourcen*), andere Bereiche betreffen individuelle, personengebundene (bspw. *Überlastung*) oder schulbezogene Problembereiche (*Unterstützung der Schulleitung*). Die zu diskutierenden Ableitungen hieraus haben somit unterschiedliche Adressat*innen und sind nicht alleinig auf die Ebene der Schulen zu beziehen.

6.2.5 Ausblick

Der Beitrag schafft auf Grundlage US-amerikanischer Studien eine erste Datengrundlage für den Bereich gelingender Implementation gestufter Fördersysteme in Deutschland. Limitationen der Studie sind die regionale Begrenzung auf Nordrhein-Westfalen und die mögliche Beeinflussung der Ergebnisse durch die ständige Verfügbarkeit des Universitätsteams. Die Interviews zeigen jedoch, dass eine systematische Förderung als nützlich angesehen wird und die Zusammenarbeit im Team sowie die Unterstützung durch die Schulleitung hierfür wichtig sind. Die Grundschullehrkräfte fühlen sich jedoch oft nicht kompetent genug, effektive schulische Diagnostik durchzuführen. Diese Aufgabe sehen sie eher im Zuständigkeitsbereich der Sonderpädagog*innen. Vor dem Hintergrund bestehender psychischer und sozialer Problemlagen auf Seiten der Schüler*innen ist die erfolgreiche Umsetzung von Diagnostik jedoch unerlässlich für alle Lehrkräfte. Prozessbegleitende Fortbildungen und eine stärkere Ausrichtung auf Diagnostik könnten helfen, diese Herausforderung zu bewältigen. Für eine erfolgreiche Implementation von MTSS ist es darüber hinaus wichtig, dass Schulen angemessen mit diagnostischen Mitteln ausgestattet sind und Lehrkräfte den Nutzen schulischer Diagnostik erkennen. Sonderpädagog*innen können schulintern unterstützen und jede Schule sollte weiterführend ein auf ihre Bedürfnisse zugeschnittenes Fördersystem entwickeln, wobei umfassende Fortbildungen und die fachliche Zusammenarbeit mit internen und externen Expert*innen essenziell sind. Multiprofessionelle Teams, die innerhalb eines MTSS die Umsetzung der Förderung koordinieren, sind ein Beispiel für eine erfolgreiche Umsetzung.

7 Zusammenfassende Ergebnisdarstellung

Die durchgeführten Studien konnten bestehende Ergebnisse deutscher und internationaler Forschung zum Thema MTSS erweitern. Auf dieser Grundlage ist die Weiterentwicklung hinsichtlich inhaltlicher und methodischer Zugänge möglich. Eine abschließende synoptische Ergebnisdarstellung soll diesen Prozess erleichtern.

7.1 Eine systematische Literaturanalyse als Referenz- und Ausgangspunkt

Als Ausgangspunkt konnten die Ergebnisse von Beitrag 1 zeigen, dass die Einführung von MTSS unter Einbezug von verhaltensmodifizierenden Interventionen in internationalen Grundschulsystemen positive Schulerfahrungen hervorruft bzw. fördert und negative Bildungseinflüsse verringern kann. Insbesondere bei störendem Verhalten sowie emotionalen und verhaltensbezogenen Störungen ergeben sich studienübergreifend eine Vielzahl an Effekten. Der Großteil der Studien (72,5 %) stammt aus den USA, gefolgt von Ländern wie Australien, Kanada, Finnland, Deutschland, Neuseeland, Taiwan und Norwegen. Diese Verteilung unterstreicht die globale Relevanz von MTSS. Während einige Studien nur allgemein gestufte Förderansätze beschreiben, benennen andere spezifische Interventionen für verschiedene Stufen, darunter die Einführung schulweiter Regeln, das Absolvieren spezifischer Trainings, Beratungsangebote und intensivere, individuelle Maßnahmen. Die betroffenen Personengruppen umfassen Schüler*innen, Eltern, Lehrer*innen, angehende Lehrer*innen, das erweiterte Schulpersonal, Psycholog*innen und Sozialarbeiter*innen. Die meisten Studien verwenden Quasi-Experimentelle Design ohne Randomisierung. Weniger weit verbreitet sind Einzelfallstudien oder randomisierte Kontrollstudien. Trotz teilweise fehlender detaillierter deskriptiver Daten, aufgrund des Einbezugs von Informationen aus nationalen Schulregistern, zeigen viele Studien positive Auswirkungen von MTSS auf problematisches Schüler*innenverhalten. Wie bereits beschrieben, sollten sich zukünftige Forschungen auf eine effektive Anpassung der Systeme an die umgebenden Strukturen, auf die Entwicklung von Lernfortschrittsdiagnostik und die Beforschung verschiedener Grundsystematiken von MTSS konzentrieren. Es ist zudem wichtig, alle beteiligten Stakeholder in den Evaluations- und Entwicklungsprozess einzubeziehen.

7.2 Untersuchung der Stufen 1, 2 & 3

Die bereits beschriebenen Studien (Beitrag 2, 3, 4 & 5) knüpfen an diese Ergebnisse an. Die erste Untersuchung der Stufen 1 und 2 wurden dementsprechend als Einzelfalluntersuchung mit experimentellem Charakter geplant und durchgeführt (Beitrag 2). Die Grundlage war der Förderansatz *Multimo*, welcher in 18 Grundschulen in Nordrhein-Westfalen implementiert wurde. Es konnte eine signifikante Reduzierung des störenden Verhaltens auf der ersten Stufe beobachtet werden. Die Einführung der zweiten Stufe führte zu einer weiteren, jedoch weniger ausgeprägten Abnahme des Verhaltens. Insgesamt war der kombinierte Einsatz des GBG und der DBRC wirksam, um störendes Verhalten zu reduzieren. Zukünftige Analysen sollten sich entlang der Ergebnisse ergänzend auf die isolierten Effekte der Interventionen in ähnlichen Settings konzentrieren.

Eine Analyse der dritten Stufe erfolgte zunächst qualitativ (Beitrag 3). Im Nachgang wurde darüber hinaus eine Auswertung eines Subsamples der Hauptstichprobe vorgenommen (Kap. 5.4). Hierunter fielen jene Kinder, die alle drei Stufen des Förderansatzes durchliefen. Die Analyse von Level-, Slope- und Trendeffekten hinsichtlich der Förderung auf Stufe 3 ergab negative signifikante Effekte für die Hälfte der Fälle: In drei Fällen waren die Level-Effekte signifikant und in einem Fall der Slope-Effekt. Ein Level-Effekt war positiv signifikant. Die Ergebnisse deuten darauf hin, dass etwa die Hälfte der Schüler*innen von der Förderung auf der dritten Stufe profitierten. Für ein Kind zeigte die Förderung nicht den gewünschten Effekt, sondern das disruptive Verhalten verstärkte sich. Darüber hinaus wurde die dritte Stufe von den Lehrkräften in den Interviews positiv bewertet. Sie berichteten von Verhaltensänderungen der Schüler*innen und verbesserten Beziehungen zu diesen sowie einer gesteigerten eigenen Kompetenzwahrnehmung. Als Schlüsselemente wurden die Qualität der Beziehung zum Coach sowie die Fähigkeit zur Selbstreflexion und Problembewältigung identifiziert. Die Lehrkräfte betonten die Wichtigkeit einer vertrauensvollen Beziehung zum Coach und eines neutralen Blicks von außen. Zusammenfassend konnte gezeigt werden, dass die Stufen 1, 2 und 3 des Förderansatzes *Multimo* unterschiedlich effektiv bei der Reduzierung störenden Verhaltens sind.

Weiter ergänzt werden können diese Ergebnisse durch den Beitrag 4, der die Effekte des GBG auf störendes Verhalten und lernförderliches Verhalten bei Schüler*innen untersucht. Die Ergebnisse zeigen, dass der NAP während der Spielphase zwischen 32,33 % und 97,45 % für das störende Verhalten und zwischen 31,82 % und 90,32 % für das lernförderliche Verhalten variierte. Die allgemeine Verhaltensentwicklung über den Schultag hinweg zeigte ebenfalls eine

Reduktion des störenden Verhaltens und den Aufbau lernförderlichen Verhaltens. Eine Mehrebenen-Analyse bestätigte signifikante Level-Effekte für das störende Verhalten und das lernförderliche Verhalten, was auf eine sofortige und konstante Verhaltensänderung zu Beginn der Interventionsphase hinweist. Die Ergebnisse zeigen somit heterogene Effekte auf das störende Verhalten und das lernförderliche Verhalten, wobei mehr Schüler*innen hinsichtlich des lernförderlichen Verhaltens als im Hinblick auf störendes Verhalten von der Intervention profitierten. Das GBG zeigte zudem positive Effekte auch außerhalb der Spielphasen, was die Wirksamkeit der Intervention unterstreicht. Diese Ergebnisse bestätigen frühere Forschungen und deuten darauf hin, dass das GBG geeignet ist, Verhaltensprobleme im Unterricht zu reduzieren.

Beitrag 5 schließt an die bisher präsentierten Ergebnisse an und kann die Effektivität der DBRC (Stufe 2) durch eine weitere durchgeführte Studie bestätigen. Auch hier kam eine Einzelfallstudie zum Einsatz, die zeigen konnte, dass der Einsatz der DBRC das disruptive Verhalten von Schüler*innen von vier zweiten Klassen reduzieren kann und gleichzeitig eine Veränderung im spezifischen Zielverhalten, dem Interventionsziel der Intervention, erreichen kann.

7.3 Untersuchung implementationsbezogener Kontextbedingungen

Um die Studienergebnisse ergänzend zu kontextualisieren, wurde der Implementationsprozess zusätzlich aus Sicht der Lehrkräfte näher untersucht (Beitrag 6). Auf diese Weise konnten verschiedene Herausforderungen und Gelingensbedingungen zur Implementation von gestuften Fördersystemen in der Grundschule identifiziert werden: Zu den Herausforderungen gehören unter anderem mangelnde personelle Ressourcen, unzureichende Fortbildungen, Zeitmangel im Schulleben, zu große Klassen, mangelnde Kooperation mit Eltern, Verhaltensauffälligkeiten der Schüler*innen und Akzeptanzprobleme der Maßnahmen. Als Gelingensbedingungen werden höherer Personalschlüssel, bedarfsorientierte Fortbildungskonzepte, Schaffung zeitlicher Ressourcen, kleinere Klassengrößen, Fortbildungen in schulischer Diagnostik und die Unterstützung durch die Schulleitung genannt. Die Studie zeigt, dass Herausforderungen und Gelingensbedingungen in unterschiedlichen Bereichen des Bildungssystems zu finden sind, darunter auf der systemischen, der individuellen und der schulbezogenen Ebene. Alle Ergebnisse zusammengenommen bieten eine erste, umfassende Datengrundlage für die erfolgreiche Implementation gestufter Fördersysteme in Deutschland in der Grundschule.

7.4 Ein Ausblick auf Basis der Ergebnisse

Die Untersuchungen zu dem Themenkomplex MTSS im internationalen Kontext der Grundschule, speziell in Bezug auf externalisierendes Problemverhalten, hat sich als aufschlussreich erwiesen hinsichtlich der Entwicklung und Bewertung eingesetzter Maßnahmen und deren Implementation. Auf dieser Basis konnten belastbare Wirksamkeitsstudien für den Förderansatz *Multimo* durchgeführt werden, sodass Belege für die unterschiedliche Wirksamkeit auf jeder Stufe des Ansatzes gefunden und diskutiert werden konnten (Beitrag 2, 3, 4 & 5). Dies öffnete die Möglichkeit für eine Weiterentwicklung des bestehenden Konzepts auf Grundlage aller Einflussfaktoren und Ergebnisse.

Zudem konnten bestehende Kritikpunkte analysiert werden, wie z. B. die Aufnahme einer rassismuskritischen Diskussion rund um MTSS (Kap. 3.7.2.3). Aktuell gibt es in den USA eine populäre Diskussion um die weitverbreiteten, wenig kultursensiblen, mehrstufigen Fördersysteme, die spezifische Interventionen zur Förderung sozialer und emotionaler Kompetenzen vorsehen, aber z. B. nicht explizit antirassistisch ausgerichtet sind (z.B. Mayes & Byrd, 2022). Bettors-Bubon et al. (2022) kritisieren jene oft kulturneutrale Umsetzung und schlagen antirassistische Methoden innerhalb eines MTSS vor, die auf Reflexionstechniken, kultursensiblen und kulturkompetenten Maßnahmen sowie der Untersuchung der Schule und Schulgemeinde basieren. Konkret beschreiben sie z. B. die Introspektion des Selbst, eine tiefgehende, selbstreflektierende Untersuchung der eigenen Einstellungen und Vorurteile. Dieser Prozess ist verankert in der Theorie des transformativen Lernens und erfordert von der Schulaufsicht und der Schulleitung ein Verständnis für ihre eigene Identität, ihre Privilegien sowie die Auswirkungen dieser auf ihr Leben (Bettors-Bubon et al., 2022; Fallon et al., 2021; Mayes & Byrd, 2022).

In diesem Zusammenhang ist die Entwicklung von kultureller Kompetenz ein weiterer wichtiger Aspekt. Die Ausgestaltung dieser Kompetenz kann sehr unterschiedlich sein. Im Kern stehen jedoch das Wissen und Verstehen kultureller Praktiken sowie Werte und Lebensrealitäten. Dies knüpft an die Ergebnisse des Beitrags 6 an. Zusätzlich sollen Schulleitungen und Lehrkräfte ihre eigenen expliziten und impliziten (rassistischen) Vorurteile reflektieren, um Schüler*innen effektiv zu unterstützen. Dieser Prozess gilt allgemein hin als notwendig, um Rassismus in Bildungskontexten entgegenzuwirken (Edirmanasinghe et al., 2022; Fallon et al., 2021). Schließlich wird auch die Untersuchung des Schulumfelds und der Gemeinde aus einer antirassistischen Perspektive betont, was auch bei Fallon et al. (2021) zur Sprache kommt. Hier wird neben der Anerkennung und Thematisierung systemischer

Rassismus-Effekte im Bildungsbereich empfohlen, dass Schulpsycholog*innen die Verantwortung übertragen wird, gemeinsam mit Schüler*innen und ihren Familien klare Verhaltenserwartungen und Anforderungen an die Lernumgebung zu definieren, um auf diese Weise eine produktive und sichere Lernumgebung zu schaffen. Schüler*innen sollen explizit lernen, diese Erwartungen zu erfüllen, wobei eine positive Verstärkung durch Lob und eine instruktive Herangehensweise anstatt einer disziplinarisch orientierten pädagogischen Haltung empfohlen wird. Zusammenfassend kann dies antirassistischen Strukturen vorbeugen oder helfen, diese zu erkennen. Die Ansätze zielen darauf ab, durch kontinuierliche Selbstreflexion, die Untersuchung von Schulsystemen und -praktiken sowie durch gezielte Maßnahmen ein inklusiveres und gerechteres Bildungsumfeld zu schaffen. Jene Diskussion muss auch unabhängig der US-amerikanischen Forschung verfolgt und zusätzlich in den Entwicklungsprozess miteinbezogen werden, um dem Grundsatz gestufter Förderung, alle Kinder in höchstem Maß inklusiv zu beschulen und zu fördern, umzusetzen (Marraccini et al., 2023). Für die Entwicklung und Adaption deutscher MTSS sollte dieses Thema, aufgrund vergleichbarer struktureller Rassismen und damit verbundenen Benachteiligungen, mitgedacht werden (Knauer, 2019; Molela Moukara, 2023).

Ein Kernergebnis aller Überlegungen ist somit, dass die Entwicklung, Implementation und Durchführung von gestufter Förderung in Deutschland, speziell im inklusiven Setting der Grundschule und bezogen auf Verhaltensaspekte, möglich und effektiv ist. Die gezielte Beforschung des MTSS *Multimo* offenbart gewisse Adaptionprozesse. Zudem wird deutlich, dass weitere Studien zu verschiedenen Settings und Systemen sinnvoll wären. Darüber hinaus wird offensichtlich, dass auch methodisch flexible und vielfältige Zugänge gewählt werden müssen, um die Effektivität und Umsetzung solcher Förderprogramme kontinuierlich zu verbessern und an die spezifischen Bedürfnisse und Rahmenbedingungen der Schüler*innen, Lehrkräfte, Eltern und weiterer beteiligter Personen anzupassen. Im Folgenden soll die kritische Einordnung der gewählten Methodik und eine Erweiterung dieser im Vordergrund der Betrachtung stehen. Hieran schließt eine abschließende Diskussion der Ergebnisse an.

8 Methodenkritik

Im Laufe der Untersuchungen und der Diskussion der Ergebnisse wurde deutlich, dass die methodische Herangehensweise bzw. der methodische Zugang einen bedeutenden Einfluss auf die Qualität der Ergebnisse im Kontext der Untersuchung gestufter Fördersysteme haben kann. Die Fragestellung nach der Untersuchung der Qualität der Förderung (z. B. die Reduktion disruptiven Verhaltens) kann auf unterschiedlichen Wegen erfolgen. Beitrag 1 konnte zeigen, dass international diverse Studiendesigns mit den gleichen Zielvariablen und Fragestellungen geplant wurden und allesamt zu evidenten Ergebnissen kommen. Im Folgenden sollen nun jeweilige methodische Zugänge vor dem Hintergrund der potenziell beantwortbaren Fragestellungen und ihren Vor- und Nachteilen auf Grundlage der präsentierten Daten und Informationen systematischer Literaturanalysen beleuchtet werden.

8.1 Kritische Analyse durchgeführter Einzelfallstudien

Experimentelle Einzelfall-Studiendesigns haben in der jüngeren Vergangenheit neben traditionellen experimentellen Gruppendesigns eine weit verbreitete Anwendung gefunden. Diese Art der Forschung, beschrieben z.B. durch Kratochwill et al. (2012) und Smith (2012), bietet die Möglichkeit, tiefgreifende Einblicke in individuelle Veränderungsprozesse zu gewinnen, was besonders für die Interventionsforschung relevant ist (Epstein & Dallery, 2022; Wilbert et al., 2022). Solche experimentellen Einzelfallstudien können vor allem in der Schulforschung nützlich sein, um die Wirksamkeit von Interventionen zur Bewältigung von Verhaltensproblemen zu untersuchen (Lipien et al., 2023; Moeyaert et al., 2023). Im Folgenden sollen die Ergebnisse und die Durchführung der Studien (Beitrag 2 & 4) samt ihren Auswertungsmethoden kritisch analysiert und eingeordnet werden. Zusätzlich sollen die bereits präsentierten Vor- und Nachteile sowie empfohlene Auswertungsmethoden und das Studiendesign in eine Diskussion miteinbezogen werden.

Der Beitrag 2, der Beitrag 4 und der Beitrag 5 untersuchen die Stufen 1 und 2 sowie die eingesetzten Interventionen des gestuften Fördersystems *Multimo* (Kap. 4) mit Hilfe von drei Einzelfallstudien²⁶. Alle Studien wurden in Form eines Multiplen-Baseline-Designs geplant und erheben den Interventionseffekt bezüglich disruptiven Verhaltens (Beitrag 2, 4 & 5) und

²⁶ Die ergänzenden Analysen zu Stufe 3 basierten auf der Stichprobe von Beitrag 2 und wurden vergleichbar ausgewertet (Kap. 5.4).

lernförderlichen Verhaltens (Beitrag 4) einzelner Schüler*innen aus mehreren Schulklassen mittels täglicher Verhaltensbeurteilung. Im Kern der Erhebung steht in Beitrag 2 die Wirksamkeitsüberprüfung der Stufen 1 und 2 (GBG & DBRC), in Beitrag 4 die isolierte Überprüfung des GBG, und in Beitrag 5 die isolierte Wirkung der DBRC.

8.1.1 Betrachtung der verschiedenen Studiendesigns

Das Studiendesign und die Datenerhebung für Beitrag 2 zur Untersuchung der Stufe 1 und 2 des Förderansatzes *Multimo* (Kap. 4) wurde gemeinsam mit den jeweiligen Schulleitungen und Klassenlehrkräften vorbereitet und geplant. Hierbei wurden Schulferien und weitere freie Zeiten in die Planung einbezogen, sodass eine Baseline-Phase im Durchschnitt 21,55 und zwei Interventionsphasen mit 43,45 und 25,9 Messzeitpunkten pro Kind erreicht werden konnten. Die Instruktionen des Projektteams, die Fortbildungen und die permanente Begleitung waren im Sinne einer hohen Datenqualität notwendig, da auf diese Weise die Wichtigkeit einer umfangreichen Baseline-Phase verdeutlicht werden konnte. Dies hielt die Lehrkräfte dazu an, möglichst gewissenhaft das Verhalten der Kinder zu beurteilen.

Ein vergleichbares Bild zeigt sich für Beitrag 4, für dessen Erhebung ebenfalls eine enge partizipative Begleitung während der Studie, die die Wirksamkeit des GBG untersuchte, vorgesehen war. Auch hier wurden das Studiendesign und die Datenerhebungsprozesse sowie die Bedeutung eines vollständigen und umfangreichen Datensatzes den Lehrkräften erläutert. Dies führte bei den Lehrkräften zu einem Bewusstsein für die Sorgfältigkeit der zu erhebenden Daten, sodass für die Baseline-Phase im Durchschnitt 17,5 und die Interventionsphase 36,25 Messzeitpunkte pro Kind erreicht werden konnte.

Die in Beitrag 5 besprochene Einzelfallstudie zur Untersuchung der Wirksamkeit der DBRC erweitert die Studienkonzepte der Beiträge 2 und 4 durch weitere partizipative Faktoren. Konkret bedeutet dies, dass die Variablen und Ausprägungen der Skalen zusammen mit den Lehrkräften und dem Projektteam entwickelt und ausformuliert wurden. Zusätzlich wurden die Hauptproblembereiche der Kinder in individuelle Items überführt, was zu einer zusätzlichen Engführung im Studiendesign beitrug. Dieses Prozedere wurde im Nachgang von den Lehrkräften und vom Projektteam als gewinnbringend beurteilt, da auf diese Weise eine hohe Proximalität der erhobenen Daten aus Sicht der Lehrkräfte und eine valide Datengrundlage aus Sicht des Projektteams erreicht werden konnte. Hier wurde ein durchschnittlicher Wert von 41 Messzeitpunkten in der Baseline-Phase und 39 Messzeitpunkten in der Interventionsphase erreicht.

8.1.2 Analyse der Stichprobenqualität

Die Stichproben der jeweiligen Studien sind allesamt im Vergleich anderer Einzelfallstudien groß angelegt und reichen von $N = 10$ (Beitrag 5) über $N = 20$ (Beitrag 4) bis hin zu $N = 29$ (Beitrag 2). Der Grund hierfür sind lange Interventionszeiträume, die ein erhöhtes Risiko des frühzeitigen Drop-Outs²⁷, eine potenzielle Einflussnahme Erziehungsberechtigter und weitere Faktoren wie unerwartete Lebensereignisse der Familien oder Lehrkräfte mit sich bringen. Auch können organisatorische Umstrukturierungen hinsichtlich der Beschulung der teilnehmenden Schüler*innen nicht ausgeschlossen werden. Auf diese Weise konnten die Fragestellungen der verschiedenen Studien, die sowohl die Entwicklung mehrerer Individuen zusammen als auch einzelner Schüler*innen über die Zeit hinweg adressierten, beantwortet werden. Zusätzlich haben alle Studien gemein, dass Lehrkräfte (zumeist die Klassenlehrkraft) das Verhalten der Schüler*innen beurteilten. Das Alter der Schüler*innen bewegte sich studienübergreifend in einem Bereich von 7,14 (Beitrag 2) über 7,7 Jahre (Beitrag 5) bis hin zu 8,3 Jahre (Beitrag 4). Wichtig anzumerken ist, dass das Verhalten der Schüler*innen neben der täglichen Verhaltensbeurteilung mittels DBR schon vor dem Start der Baseline-Phase zusätzlich in Form der ITRF (Casale et al., 2018) eingeschätzt wurde. Auf diese Weise wurde ausgeschlossen, dass die ausgewählten und bewerteten Schüler*innen kein oder zu geringes externalisierendes Verhalten zeigten. Somit konnte sichergestellt werden, dass die Lehrkräfte in dem zumeist ohnehin von Zeitmangel bestimmten Schulalltag keine täglichen Verhaltensbewertungen durchführten, die schlussendlich nicht hätten verwertet werden können.

8.1.3 Analyse der Datenqualität

Eine der Herausforderungen in Bezug auf die Durchführung und Auswertung von Einzelfallstudien ist die Erhebung belastbarer und vollständiger Daten (Smith, 2012). Da die Datengrundlage einen erheblichen Einfluss auf die Belastbarkeit der Ergebnisse hat (Ninci, 2023; Wilbert et al., 2022), soll diese im Folgenden für alle zuvor erwähnten Einzelfallstudien analysiert und kritisch diskutiert werden.

In der Baseline-Phase (A-Phase) in Beitrag 2 waren im Durchschnitt 22 (15-25)²⁸ Messzeitpunkte geplant, in der 1. Interventionsphase mit dem GBG (B-Phase) 44 (40-45) und in der 2. Interventionsphase (C-Phase, Einsatz der DBRC) 26 (15-43). Hierbei fehlten im

²⁷ Studienabbruch bzw. Nichtbeendigung der Studie eines Falls

²⁸ Alle Angaben in diesem Abschnitt beziehen sich auf den minimalen Wert und den größten Wert (min. – max.).

Durchschnitt in der A-Phase etwa sechs, in der B-Phase etwa 21 und in der C-Phase etwa zehn Datenpunkte. Somit fehlten 28 % der Daten in der A-Phase, 48 % in der B-Phase und 37 % in der C-Phase (Tabelle 6). Für jeden Fall konnte dennoch eine ausreichend hohe Datengrundlage von mindestens zehn Messzeitpunkten pro Phase (im Durchschnitt mindestens 15,5 Datenpunkte) erhoben werden. Die interne Validität hätte jedoch durch mehrere Messungen in dieser Studie weiter erhöht werden können. Die externe Validität kann unter den genannten Einschränkungen aufgrund der großen Stichprobe und des Studiensettings dennoch als hoch bewertet werden²⁹(Woolcock, 2013).

Ein ähnliches Bild zeigt sich für Beitrag 4. Hier waren in der A-Phase für jede Variable im Durchschnitt 17,5 Messzeitpunkte (10-20) geplant, wovon im 1. Setting „Spielphase“ für die Variable störendes Verhalten und für die Variable lernförderliches Verhalten im Durchschnitt ca. 4 Datenpunkte fehlten. In der B-Phase waren im Durchschnitt 36 Messzeitpunkte geplant (30-42), wovon für beide Variablen durchschnittlich 14 Datenpunkte fehlten. Im Setting „Schultag“ war die gleiche Anzahl an Messzeitpunkten geplant und es fehlten für beiden Variablen in der A-Phase etwa vier Datenpunkte und in der B-Phase etwa 15 Datenpunkte. Die Ergebnisse von Beitrag 4 zeigen, dass für beide Variablen (in beiden Settings) gemittelt 21 % der Daten in der A-Phase und 41 % in der B-Phase fehlten (Tabelle 6). Zwei Fälle (S5 & S15) erreichten hier in der Baseline-Phase und in der Interventionsphase im Setting „Schultag“ eine geringe Anzahl an Messzeitpunkten in beiden Phasen (vier und sechs Datenpunkte). Der Fall S15 erreichte zudem auch in dem Setting „Spielphase“ einen geringen Wert von nur vier Datenpunkten. Alle anderen Schüler*innen ($n = 27$) weisen höhere Werte auf, sodass zusammenfassend auch hier von einer guten Datenqualität und validen Ergebnissen ausgegangen werden kann. Es wird aber deutlich, dass die Aussagekraft durch eine verbesserte Datenqualität, insbesondere für die Fälle S5 und S15, stark erhöht hätte werden können.

In Beitrag 5 wurden die Phasen aus den Erfahrungen der vorherigen Studien deshalb verlängert, was die Datenqualität steigern und stabilisieren konnte. Hier zeigt sich folgendes Bild: Für beide Variablen waren im Durchschnitt 41 Messzeitpunkte in der A-Phase (35-45) und 39 Messzeitpunkte in der B-Phase (35-45) geplant. Über alle Fälle hinweg fehlten hier in der A-Phase im Durchschnitt 18 Messwerte und in der B-Phase 11. Der Datensatz des Beitrag 5 zeigt, dass 44 % der Daten in der A-Phase und 29 % in der B-Phase fehlten (Tabelle 6). In der Interventionsphase konnte, verglichen mit den vorherigen Studien, somit eine höhere

²⁹ Da die externe Validität von Einzelfallstudien grundsätzlich als niedriger im Vergleich zu größer angelegten randomisierten Studiendesigns beschrieben wird (Woolock, 2013), ist dies im Verhältnis zu anderen Einzelfallstudien zu sehen.

Datenqualität sichergestellt werden, die zu validen Ergebnissen führt. Die absolute Anzahl an Messzeitpunkten war in dieser Studie insgesamt höher als die von Beitrag 2 und Beitrag 4.

In Tabelle 6 wird die Anzahl an durchschnittlich fehlenden Werten in den jeweiligen Phasen der Studien über alle Variablen und Messzeitpunkte hinweg in Prozent zusammengefasst. Es zeigt sich, dass Beitrag 5 die höchste Anzahl an Messzeitpunkten pro Phase aufweist, gefolgt von Beitrag 2 und Beitrag 4. Die fehlenden Werte entstehen vor allem aufgrund von Schulferien und weiteren freien Zeiten, Ausfällen durch das Lehrpersonal, Fortbildungen, Krankheit und selten aufgrund mangelnder Möglichkeit der Umsetzung. Über alle Studien hinweg zeigt sich jedoch, dass mit einem Ausfall von mindestens 30-40 % gerechnet werden kann. Dieses Ergebnis bzw. dieser Richtwert kann für weitere Studien in die Planung einbezogen werden, sodass ausreichend viele Datenpunkte im Vorfeld angelegt werden, um eine mangelnde Aussagekraft durch fehlende Daten zu verhindern.

Tabelle 6

Übersicht über vorhandene und fehlende Daten der Einzelfallstudien (Beitrag 2, 4 & 5)

Studie	M_{nA}	M_{nB}	M_{nC}	mis_A	mis_B	mis_C
<u>Beitrag 2</u>	15,5	22,5	16,4	28 %	48 %	37 %
<u>Beitrag 4</u>	13,8	21,5		21 %	41 %	
<u>Beitrag 5</u>	23,0	27,8		44 %	29 %	

Anmerkung: M_{nA} = Durchschnittliche Anzahl an Messzeitpunkten aller Variablen in der A-Phase; M_{nB} = Durchschnittliche Anzahl an Messzeitpunkten aller Variablen in der B-Phase, M_{nC} = Durchschnittliche Anzahl an Messzeitpunkten aller Variablen in der C-Phase, mis_A = Missings der A-Phase (Baseline-Phase), mis_B = Missings der B-Phase (1. Interventionsphase), mis_C = Missings der C-Phase (2. Interventionsphase)

Verglichen mit Ergebnissen einer Studie von Wilbert et al. (2022) erscheint die Anzahl an Datenpunkten in den Studien der Beiträge 2, 4 und 5 als größtenteils ausreichend. Sie liegen jedoch in Teilen unter der empfohlenen Mindestanzahl an Messungen. Die Autor*innen stellen fest, dass die Teststärke stark von den spezifischen Designmerkmalen der Studie abhängt, sodass empfohlen wird, in Phase A mindestens neun Messzeitpunkte anzulegen, wenn eine Reduktion des Problems durch die Intervention von 60 % oder mehr erwartet wird. Bei einer erwarteten Reduktion von 40-60 % sollten mindestens 15 Messungen in Phase A und 30 in Phase B erfolgen. Um die Werte im Folgenden zu kontextualisieren, wurde der prozentuale Rückgang des Verhaltens entlang der Mittelwerte der verschiedenen Studien geschätzt. Hierzu wurden die durchschnittlichen Mittelwerte aller Fälle in den jeweiligen Phasen aufsummiert und durch ihre Anzahl geteilt. Der durchschnittliche Mittelwert einer Phase und einer Variable

wurde dann in einen prozentualen Wert umgerechnet. Die Ergebnisse der verschiedenen Beiträge sind in Tabelle 7 dargestellt.

Tabelle 7

Übersicht über die Verhaltensveränderung der Fälle der Einzelfallstudien (Beitrag 2, 4 & 5)

Studie	Verhaltensveränderung	
	A-B	B-C
<u>Beitrag 2</u>	23 %	17,7 %
<u>Beitrag 4</u>	26,3 % (SV) -31,7 % (LV)	
<u>Beitrag 5</u>	70,1 %	

Anmerkung: A-B / B-C= Prozentuale Veränderung des bewerteten Verhaltens über alle Fälle auf Basis der Mittelwerte der Fälle hinweg von der A- zur B-Phase und von der B- zur C-Phase, SV = störendes Verhalten, LV = lernförderliches Verhalten

Weitere Studien, z. B. Hopman et al. (2017); Leidig et al. (2022); Spilles et al. (2019) zum GBG oder auch Ackley (2022); Miller et al. (2017); Vannest et al. (2019) zur DBRC, bestätigen die Ergebnisse und ergänzen die Angaben bzgl. des erwarteten Rückgangs des Verhaltens. Hier wird erneut deutlich, dass in Beitrag 2 und Beitrag 4 eine höhere Anzahl an Messzeitpunkten validere Ergebnisse hervorgebracht hätten, was auch durch Analysen von Wilbert et al. (2022) bestätigt werden kann. Für die weitere Planung und Umsetzung von Einzelfallstudien im Kontext von MTSS sollte dies, wie in Beitrag 5, berücksichtigt werden.

Abschließend lässt sich nicht eindeutig bewerten, inwiefern die Ergebnisse durch fehlende Daten beeinflusst wurden. Da die Stufe 2 des Förderansatzes *Multimo* dieselbe Intervention enthielt, sich die Effekte der jeweiligen Studien allerdings unterscheiden, könnte dies eine der Erklärungen für die unterschiedlichen Effekte sein. Hinzu kommt eine unterschiedliche Art der Fortbildung (Multiplikator*innenkonzept in Beitrag 2; Fortbildung direkt durch die Universität in Beitrag 4 und 5) sowie eine andere, kleinere Stichprobe mit weniger Lehrkräften in Beitrag 5. Hierdurch kann eine engere Betreuung ermöglicht werden, die bei der Interpretation der unterschiedlichen Ergebnisse mitbedacht werden muss. Außerdem kann keine Aussage über die Implementationsqualität gemacht werden, da diese in Beitrag 2 nicht erhoben wurde.

8.1.4 Gängige Verfahren zur Auswertung von Einzelfalldaten

Um im Folgenden die eng mit den Daten verknüpften Auswertungsmethoden zu diskutieren, soll kurz ein Überblick über gängige Verfahren gegeben werden. Neben der weitverbreiteten Diskussion um die Datenqualität stehen Auswertungsmethoden und ihre Anwendung ebenfalls

im Fokus der Forschung (Smith, 2012). Die Verwendung der Auswertungsmethoden hängt von den spezifischen Eigenschaften der Daten und den Zielen der Analyse ab: Als Beispiel ist der Einsatz regressionsbasierter Auswertungsmethoden besonders nützlich, wenn das Ziel der Analysen ist, die Veränderungen in der Steigung oder das sofortige Niveau der Veränderungen des Verhaltens zu quantifizieren (Dowdy et al., 2021). Die stückweise Regression (*Piecewise Regression*) kann somit verwendet werden, um sofortige Veränderungen im Niveau zu quantifizieren (Wilbert et al., 2022; Wilbert & Lueke, 2022). Das Ermitteln der Slope- und Level-Effekte quantifiziert hierbei den durchschnittlichen Unterschied im Niveau der Phasen. Diese Methoden sind besonders geeignet, wenn die visuelle Inspektion der Daten eine sofortige Veränderung nicht eindeutig anzeigt oder wenn eine feinere Quantifizierung der Veränderungen pro Messzeitpunkt erforderlich ist. Ergänzend hierzu vergleichen Jamshidi et al. (2020) verschiedene Multilevel-Modelle zur Analyse von Daten aus Einzelfallstudien, insbesondere in Bezug auf ihre Fähigkeit, abhängige Regressionskoeffizienten zu ermitteln. Es wurden drei Arten von Modellen getestet: ein univariates Modell, das Abhängigkeiten der Daten ausschließt, ein multivariates Modell, das Kovarianzen auf höheren Ebenen ignoriert und ein multivariates Modell, das Kovarianzen zwischen den Variablen einbezieht. Die Ergebnisse zeigen, dass multivariate Modelle, unabhängig davon, ob und wie sie Kovarianzen einbeziehen, bessere Schätzungen der Gesamteffekte und der Varianz liefern können als univariate Modelle. Die Studie hebt auch die erhöhte Robustheit der multivariaten Modelle hervor. Die Empfehlung der Autor*innen ist demnach, ein multivariates Mehrebenen-Modell zu berechnen, um die bestehende Abhängigkeit zwischen mehreren Regressionskoeffizienten (d. h. die Effektgrößen der Intervention) zu modellieren. Ergänzend wird in Einzelfallstudien im schulischen Bereich häufig der NAP (Parker & Vannest, 2009) berechnet (Smith, 2012). Die Berechnung eignet sich, wenn es keine klare Tendenz in den Baseline-Daten gibt und visuell nicht klar erkennbar ist, wie viel Überlappung in den Daten vorhanden ist (Parker & Vannest, 2009). Dieser Index kann jedoch in bestimmten Fällen weniger informativ sein, insbesondere weil er keine Quantifizierung in Bezug auf die abhängige Variable (z. B. Anzahl der Wörter oder Sätze pro Minute) bietet (Chen et al., 2016; Pustejovsky et al., 2023). Weitere potenziell anwendbare Indizes sind z. B. der *Tau-U* (Parker, Vannest, et al., 2011), der PND (Scruggs & Mastropieri, 2012), der *Percentage of Zero Data* (PZD, Scotti et al., 1991), die *Robust Improvement Rate Difference* (RIRD, Parker et al., 2009) und die *Mean Baseline Reduction* (MBLR, Campbell, 2004). Die verschiedenen Indizes haben allesamt Vor- und Nachteile bezüglich der Interpretation und Berechnung sowie der benötigten Datengrundlage (Manolov & Moeyaert,

2017). Im Folgenden soll die Auswahl der Auswertungsmethoden in den vorgestellten Beiträgen diskutiert werden.

8.1.5 Analyse der verwendeten Auswertungsmethoden

Für die vorliegenden Studien (Beitrag 2, 4 & 5) wurde, basierend auf den beschriebenen Regressionsmodellen, neben den Level-, Slope und Trend-Effekten der NAP (Parker & Vannest, 2009) berechnet. Dies erscheint vor dem Hintergrund der Datengrundlage als sinnvoll und aussagekräftig. Insbesondere die Gruppenanalysen bzw. das Berechnen der verschiedenen Modelle gilt im Kontext der verfügbaren Auswertungsmethoden und ihren Anwendungsbereichen als die belastbarste Methode bezüglich der Quantifizierung der Veränderung des Verhaltens (Wilbert et al., 2022). Hierdurch konnten die Datenstruktur und die nicht normalverteilten Daten bestmöglich berücksichtigt werden. Zusätzlich können regressionsbasierte Analysen mit einer Vielzahl von Datenmustern umgehen, einschließlich solcher mit hoher Variabilität oder nicht-linearen Trends (Campbell, 2004; Jamshidi et al., 2020), die in den hier präsentierten Studien vorkamen. Die Berechnung standardisierter Maße für Level-, Slope- und Trend-Effekte erleichtert zudem den Vergleich von Ergebnissen über die verschiedenen Studien hinweg. Zusätzlich konnten jene Analysen kleine, subtile Effekte aufdecken, die bei herkömmlichen deskriptiven Analysen möglicherweise übersehen worden wären.

Für den NAP (Parker & Vannest, 2009) ergibt sich ein ähnliches Bild: Einer der Hauptvorteile des NAP ist seine einfache Berechnung und Interpretation. Es handelt sich um ein direktes, leicht verständliches Maß, das nicht auf komplexen statistischen Berechnungen basiert (Parker & Vannest, 2009). Dies macht den Index für Praktiker*innen und Forscher*innen zugänglich, die möglicherweise keine tiefgreifenden Kenntnisse in Statistik besitzen. Zusätzlich ist der NAP besonders nützlich in Studien mit kleinen Stichprobengrößen (Chen et al., 2016). In solchen Fällen bietet der NAP eine solide Alternative zur Beurteilung von Interventionseffekten. Darüber hinaus ist der NAP ein nicht-parametrisches Maß, was keine spezifischen Anforderungen über die Verteilung der Daten voraussetzt, sodass er gut geeignet ist für Daten, die eine hohe Variabilität aufweisen.

Die Auswahl der Auswertungsmethoden erscheint vor dem Hintergrund wissenschaftlicher Einordnung, der Datenstruktur, der untersuchten Variablen sowie der Komplexität der Studien als passend und ist an folgenden Stellen zu diskutieren: Zum einen wird durch die Analysen von Jamshidi et al. (2020) deutlich, dass multivariate Mehrebenen-Analysen die höchste Test-Güte aufweisen. Die Mehrebenen-Struktur der Daten (Kinder in

Klassen und z. T. in Schulen) hätte auf der Basis der vorhandenen Datenstruktur jedoch nicht weiter berücksichtigt werden können. Dies hätte zu Teststärke-Problemen (Power-Probleme)³⁰ geführt und eine valide Berechnung erschwert (Cohen, 2013). Darüber hinaus sind die Stichproben und die Anzahl der Fälle auf den verschiedenen Ebenen (Klassen und Schulen) zu gering, um valide Ergebnisse einer erweiterten gestuften Auswertung mit Hilfe von Dummy-Kodierungen zu erwarten. Somit könnten wichtige Varianzen übersehen worden sein, die durch die Klassenzugehörigkeit hätten erklärt werden können. Durch die ergänzenden Analysen auf Fallebene konnte aber dennoch ein umfassendes Bild der Daten und ihrer Effekte erzeugt werden. Zudem wurden in allen Studien (mehrere) multivariate Modelle berechnet, die im Sinne der Analysen von Jamshidi et al. (2020) als aussagekräftiger gelten als univariate Modelle. Die Beziehung zwischen den Fällen wurde hier (sofern möglich) durch die Berechnung der *Random-Effects* (Zwischensubjekt-Effekte) vollzogen und diskutiert. Zusätzlich hätten neben dem NAP weitere Overlap-Indizes berechnet werden können, was jedoch unter Berücksichtigung der Diskussion der verschiedenen Indizes und ihrer Belastbarkeit keine erhöhte Aussagekraft nach sich gezogen hätte. Dennoch wären Indizes wie der RIRD eine robuste Methode zur Bewertung von Verhaltensinterventionen, da er Sensitivität gegenüber kleinen, aber bedeutenden Verhaltensänderungen zeigt. Daneben ist er robust gegenüber Datenvariabilität, was ihn in der angewandten Verhaltensanalyse besonders wertvoll macht. Diese Faktoren erfüllt der NAP allerdings in einem gleich hohen Maße (Lenz, 2017; Parker & Vannest, 2009; Parker et al., 2009). Neben dem RIRD wären noch der PZD und der MBLR infrage gekommen. Diese Indizes sind geeignet, wenn der Fokus auf der Reduzierung bestimmter Verhaltensweisen liegt, was in den Studien der Fall war. Der PZD wird definiert als der Prozentsatz der Datenpunkte, bei denen das untersuchte Verhalten nicht auftritt, sobald die erste problemfreie Messung auftritt (bspw. Wert für disruptives Verhalten = 0). Der PZD misst somit den Anteil der Zeiten, in denen das Zielverhalten (ausbleibendes störendes Verhalten) nicht mehr beobachtet wird, nachdem es zum ersten Mal ausblieb (das erste Mal kein störendes Verhalten auftrat). Abschließend kann jedoch nach einer intensiven Gegenüberstellung verschiedener Auswertungsmethoden ein hohes Maß an wissenschaftlicher Güte und Belastbarkeit der verwendeten Analysemethoden abgeleitet werden. Insbesondere die

³⁰ Eine Power-Analyse ist ein statistisches Verfahren, das verwendet wird, um die erforderliche Stichprobengröße zu bestimmen. Sie berücksichtigt den erwarteten Effekt, die Variabilität der Daten, das gewünschte Signifikanzniveau und die angestrebte statistische Power einer Analyse (Cohen, J., 2013).

Kombination eines Overlap-Indexes, wie dem NAP, mit den regressionsbasierten Auswertungen kann als valide Auswertungsmethodik angesehen werden.

8.1.6 Zusammenfassende Darstellung

Es lässt sich also folgende Zusammenfassung im Kontext der umgesetzten Einzelfallstudien ableiten: Die drei Studien (Beitrag 2, 4 & 5) analysieren die Wirksamkeit von Interventionen bei externalisierendem Verhalten im Kontext des Förderansatzes *Multimo*. Sie nutzen ein multiples Baseline-Design und fußen auf der Datenerhebung durch tägliche Verhaltensbeurteilungen des Schüler*innenverhaltens. Das Studiendesign und die Datenerhebung waren eng mit den Schulen abgestimmt, um eine umfassende und vollständige Datenerfassung zu ermöglichen. Die Stichproben waren (vergleichsweise) groß angelegt, um verschiedene Risikofaktoren wie Drop-Outs oder unvorhergesehene Ereignisse zu berücksichtigen. Bei der Analyse der Datenqualität zeigten sich jedoch Herausforderungen bezüglich der Vollständigkeit der Daten. Trotz fehlender Daten konnte eine ausreichende Datenbasis für valide Ergebnisse gesichert werden. Die Studien zeigen, dass für den schulischen Bereich mit einem Prozentsatz von 30-40 % an fehlenden Daten kalkuliert werden sollte. Für zukünftige Studienplanungen kann dies berücksichtigt werden. Zur Auswertung der erhobenen Einzelfalldaten wurden verschiedene Methoden angewandt, darunter regressionsbasierte Analysen und der NAP (Parker & Vannest, 2009). Diese Methoden ermöglichten eine differenzierte Bewertung der Interventionseffekte und konnten subtile Verhaltensveränderungen aufdecken, die in deskriptiven Analysen möglicherweise übersehen worden wären. Die verwendeten Auswertungsmethoden wurden verglichen mit weiteren potenziell einsetzbaren Methoden als wissenschaftlich fundiert und valide betrachtet. Die Kombination von Overlap-Indexes und regressionsbasierten Auswertungen wurde vor dem Hintergrund wissenschaftlicher Metaanalysen und methodischer Diskussionen als valide Auswertungsmethode bewertet.

8.2 Qualitative Forschung im Kontext MTSS

Neben der quantitativen Untersuchung des Förderansatzes *Multimo* wurden zwei qualitative Studien bezüglich der Wirkfaktoren der Stufe 3 und des Implementationsprozesses durchgeführt. Um die Studien im Gesamtkontext der bereits besprochenen Beiträge und der globalen Diskussion rund um die methodische Triangulation dieser Arbeit kritisch diskutieren zu können, erfolgt zunächst ein Blick auf bestehende qualitative Forschung im Kontext von

MTSS. Abschließend sollen die durchgeführten Studien kritisch diskutiert und eingeordnet werden.

8.2.1 Analyse bestehender qualitativer Forschung im Kontext MTSS

Der Fokus der qualitativen Studien im Bereich MTSS liegt vor allem auf Fragestellungen bezüglich Implementationsprozessen und der konkreten Umsetzung und Ausführung gestufter Förderung. Eine im Vergleich zu den quantitativen Studien geringere, aber dennoch umfassende Anzahl verschiedener Werke beschäftigt sich mit diesen Themen, was in der US-amerikanischen Forschung politische Hintergründe hat. Die Einführung gestufter Fördersysteme erfolgt dort einer politischen Vorgabe und Ausrichtung des Bildungswesen betreffend, sodass eine flächendeckende Überprüfung auch vor dem Hintergrund der Wirtschaftlichkeit und Nachhaltigkeit der implementierten Systeme und Vorgänge notwendig erscheint (Kap. 3).

Die systematische Literaturübersicht von Fox et al. (2021), die bereits Gegenstand dieser Rahmenschrift war, fasst neben diversen quantitativen Erhebungen durch standardisierte Erhebungsbögen 16 Interviewstudien (einschließlich fünf Mixed-Method-Ansätzen) zusammen, die sich größtenteils auf die Herausforderungen und Gelingensbedingungen der Implementation gestufter Fördersysteme in den USA konzentrieren. Das Thema mehrerer Studien ist hier die Bedeutung von administrativer Unterstützung und Führung als Schlüsselfaktor. Andreou et al. (2014), Flannery et al. (2009) und Horner et al. (2013) betonen z. B. die Rolle der Schulleitungen und Bezirkskoordinatoren in der erfolgreichen Implementation und Aufrechterhaltung von gestufter Förderung und beforschen dies qualitativ. Ein weiteres wiederkehrendes Thema ist die Bedeutung von Lehrkraftengagement und -einbindung. Studien wie die von Don Kincaid, Karen Childs, et al. (2016) und Pinkelman et al. (2015) zeigen auf, dass die Akzeptanz und Unterstützung der Lehrkräfte und des erweiterten Schulpersonals entscheidend für eine gelingende Implementation sind. Die Nutzung von Daten für das *decision making* ist ein weiterer wichtiger Aspekt, der in mehreren Studien, einschließlich der von George et al. (2018) und Savage et al. (2011), qualitativ erhoben wird. Studien von Andreou et al. (2014) und Don Kincaid, Karen Childs, et al. (2016), ergänzen diese Befunde ebenfalls durch qualitative Analysen.

Jene Erkenntnisse lassen auf ein vielfältiges methodisches Vorgehen in Bezug auf die Generierung von qualitativen Interviewdaten durch die Forschenden schließen. Insbesondere wurden semi-strukturierte Interviews als zentrales Instrument verwendet, um die Interviews durchzuführen. Diese Methode wurde in z.B. den Studien von Andreou et al. (2014), Bambara

et al. (2009) und Lohrmann et al. (2008) eingesetzt, um tiefgehende Einblicke in die Erfahrungen und Meinungen von Lehrkräften, Schuladministrator*innen und anderen Beteiligten zu gewinnen. Diese Herangehensweise ermöglicht es, detaillierte und subjektive Perspektiven zu ergründen, die für das Verständnis der Implementationsprozesse von MTSS wesentlich sind. Ein weiteres verbreitetes Interviewformat sind offene Fragen (Feuerborn et al., 2016; Flannery et al., 2009; Tyre et al., 2019). Diese Methodik erlaubt es den Teilnehmenden, ihre Gedanken und Bedenken frei zu äußern, was besonders wertvoll für das Sammeln von unvoreingenommenen und umfassenden Informationen ist. Ergänzend wählten Studien strukturierte Interviews oder schriftliche Antworten (George et al., 2018; Horner et al., 2013). Savage et al. (2011) wählten eine Kombination aus semi-strukturierten Interviews und Dokumentenanalyse. Dies ermöglichte sowohl qualitative Daten aus persönlichen Gesprächen als auch objektive Informationen aus bestehenden Dokumenten auszuwerten. Eine etwas andere Methodik wurde in einer Studie von Don Kincaid, Karen Childs, et al. (2016) eingesetzt, um Implementationsprozesse aus der Sicht verschiedener Personen zu ergründen. Dort kam eine Gruppeninterviewmethodik namens *Nominal Group Process* (NGP, Delbecq et al., 1986) zum Einsatz. Diese Technik ist besonders geeignet, um einen Konsens bzw. Dissens einer Gruppe bezüglich eines Themas zu identifizieren.

Die in den Studien genannten Gruppen reichen von Lehrkräften und Schuladministrator*innen bis hin zu spezialisierten Coaches und weiteren externen Personen (Hoover et al., 2019; Horner et al., 2013; Marlowe, 2021; McKenna et al., 2021). Diese Vielfalt in der Stichprobenwahl liefert ein umfassendes Bild der verschiedenen Perspektiven und Erfahrungen in Bezug auf die Implementation und Nachhaltigkeit von MTSS und den damit verbundenen Prozessen. Hierbei waren Lehrkräfte und Schuladministrator*innen die am häufigsten befragten Teilnehmer*innengruppen (z.B. Andreou et al., 2014; Bambara et al., 2009; Savage et al., 2011). Einige Studien, wie die von Pinkelman et al. (2015), bezogen zudem Eltern und andere Mitglieder der Schulgemeinschaft mit ein. Diese Einbindung bot zusätzliche Einblicke in die Auswirkungen von MTSS auf die Schüler*innen und ihre Familien und half, ein ganzheitlicheres Bild der Implementationsprozesse zu zeichnen. In einigen Fällen wurden spezifischere Gruppen wie Distrikt-Koordinator*innen (George et al., 2018) oder staatliche Koordinator*innen samt Delegierten (Horner et al., 2013) befragt. Diese Gruppen gaben Aufschluss über die Implementation von MTSS auf einer breiteren, organisatorischen Ebene, was für das Verständnis der systemweiten Herausforderungen und Erfolgsfaktoren wichtig ist.

Zusätzlich zu den Implementationsprozessen kann auch die Effektivität hinsichtlich Einstellungen, Verfahrensweisen, Umgangsarten oder individuellen Handlungsmustern

Gegenstand qualitativer Interviewstudien sein (Braun et al., 2018). Diese Methodik ist somit eine wirksame und aufschlussreiche Erhebungsform, um proximale Eigenschaften pädagogischen Handelns zu ergründen und es zu bewerten.

8.2.2 Einordnung der durchgeführten Interviewstudien (Beitrag 3 & 6)

Um die Wirkfaktoren und intraindividuellen Veränderungsprozesse hinsichtlich der Wahrnehmung und Regulation von externalisierendem Problemverhalten zu untersuchen, setzt die Studie an den Zielpersonen des Coachings, den Lehrkräften an. Diese konnten durch semi-strukturierte Interviews einen Einblick in ihr pädagogisches Handeln und ihre Einstellungen sowie ihre erlebten Veränderungsprozesse geben. Zusätzlich konnte die Effektivität der Maßnahmen mit diesen qualitativen Ergebnissen kontextualisiert werden (Kap. 5.4). Vergleichbare Studien wenden eine ähnliche Methodik sowohl in Bezug auf die Datenerhebung als auch auf die Auswertung an und konnten valide Ergebnisse erzielen (Castillo et al., 2022; Gerzel-Short, 2018). Die angewandte Methodik der Interviewstudie (Beitrag 3) kann vor dem Hintergrund bestehender Forschung somit grundsätzlich als aussagekräftig beurteilt werden.

Die Methodik der Interviewstudie (Beitrag 6) fügt sich in das Bild bestehender Forschung ein. Die Stichprobe und die verwendeten Auswertungsmethoden erfüllen die Gütekriterien qualitativer Forschung und führen zu belastbaren Ergebnissen. Für zukünftige Studien könnte in Betracht gezogen werden, weitere Stakeholder, die in den Implementationsprozess eingebunden sind, in die Stichprobe einzubeziehen und somit das Gesamtbild der Daten zu erweitern. Zusätzlich ist die Orientierung an der US-amerikanischen Forschung (deduktive Kategorien) und die Erweiterung dieser um induktive Kategorien als gängig und passend zu bewerten. Auch wenn sich die Strukturen auf verschiedenen Ebenen des Bildungssystem zwischen den USA und Deutschland stark unterscheiden, lassen sich auf prozeduraler Ebene der Anwendung und Durchführung gestufter Förderung Gemeinsamkeiten erkennen. Insbesondere die grundlegenden Methoden, wie das *decision making*, die datengeleitete Diagnostik oder auch die angepassten Interventionen entlang der Förderstufen, können schulsystemunabhängige Herausforderungen darstellen. Darüber hinaus wurde mit der CIT (Flanagan, 1954) eine Technik angewandt, die auch in anderen Studien zur Implementationsforschung in den USA zum Einsatz kam (McIntosh et al., 2015). Diese Technik ermöglicht es, kritische Ereignisse samt ihre Gründe zu erkennen und somit pädagogische Verfahrensweisen und Einstellungen offenzulegen.

Für weitere Studien zeigt sich, dass eine Erweiterung der bisher eingesetzten Methoden sinnvoll wäre. Denkbar wären hier z. B. Ansätze wie die NGP-Methode (Delbecq et al., 1986),

um ein tiefgreifendes Verständnis der gruppebezogenen Arbeitsprozesse herauszustellen (Don Kincaid, Karen Childs, et al., 2016). Die NGP-Methode ist eine strukturierte Gruppentechnik, die sich auf die Generierung und Priorisierung von Ideen konzentriert. Sie beginnt mit einer Phase, in der die Teilnehmer*innen individuell ihre subjektiven Gedanken zu einer zentralen Frage notieren. Anschließend teilen sie ihre Ideen in einer Austauschrunde, wobei jede Idee aufgezeichnet wird, ohne dass eine Diskussion oder Kommentierung stattfindet. Danach folgt eine zeitlich begrenzte Diskussionsphase, in der Klarstellungen zu den Ideen erfragt werden, um sicherzustellen, dass alle Ideen verstanden werden. Die Teilnehmer*innen bewerten dann die Ideen und es wird ein Durchschnittsrang für jede Idee berechnet. Nach einer Diskussion der Rangfolge erfolgt eine Neubewertung der „Top-Ideen“. Abschließend bewertet jede*r Teilnehmer*in jede Idee auf einer 7-Punkte-Skala zur weiteren Priorisierung. Der Zweck dieser Methode liegt in der Sammlung eines breiten Spektrums an Ideen sowie Meinungen und deren Priorisierung. Sie wird besonders in Situationen eingesetzt, in denen die gleichberechtigte Einbeziehung aller Teilnehmer*innen wichtig ist, um ein umfassendes Verständnis von Ansichten oder Lösungen für ein bestimmtes Problem zu erlangen. Die Methode wird für ihre Fähigkeit geschätzt, bessere Ergebnisse als unstrukturierte Gruppeninteraktionen zu liefern, weil sie mehr Ideen fördert und die Kreativität der Lösungen erhöht (Delbecq et al., 1986; Don Kincaid, Karen Childs, et al., 2016). Allerdings ist sie zeitintensiv und erfordert eine sorgfältige Moderation, um ihre Effektivität zu gewährleisten. Im Kontext der Forschung zur Implementation von MTSS könnte diese Technik auf zwei Weisen nützlich sein: Zum einen um Implementationsprozesse und Ideen für die eigene Schule und das Umfeld zu ergründen und eine Möglichkeit abzuleiten, gestufte Förderung in der eigenen Schule zu implementieren. Zum anderen kann sie helfen, Daten zu gewinnen, die im weiteren Prozess ausgewertet werden können, um Verfahrensprozesse und Möglichkeiten der Teamarbeit, Datengewinnung oder Interventionsumsetzung zu erforschen. Zusammenfassend kann der Ausblick gegeben werden, dass die Anwendung gruppenbezogener Auswertungsverfahren eine sinnvolle und passende Möglichkeit in Bezug auf vorrangige Fragestellungen der Implementationsprozesse von MTSS darstellen kann.

8.2.3 Zusammenfassende Darstellung

Die existierende qualitative Forschung konzentriert sich vornehmlich auf Implementationsprozesse verschiedener MTSS (Fox et al., 2021). Der vorrangige Schwerpunkt liegt hierbei auf den Herausforderungen und Gelingensbedingungen jener Systeme (vor allem in den USA). Erkannte Schlüsselfaktoren sind administrative Unterstützung und Führung,

Lehrkraftengagement und -einbindung, effektive Datennutzung sowie Ressourcenverfügbarkeit (z.B. [Bambara et al., 2009](#); [Fallon et al., 2022](#); [Tyre et al., 2019](#)). Verschiedene Studien betonen die Bedeutung dieser Faktoren für die erfolgreiche Umsetzung von MTSS. Methodisch dominieren semi-strukturierte Interviews und offene Fragen die Methodik ([Fox et al., 2021](#)). Die befragten Gruppen reichen von Lehrkräften bis Schuladministrator*innen, was ein umfassendes Bild der verschiedenen Perspektiven auf MTSS liefern kann. [Beitrag 6](#) fügt sich in das Spektrum der bestehenden Forschung ein. Es handelt sich um die erste qualitative Interviewstudie in Deutschland zum Themenbereich MTSS und sie orientiert sich an US-amerikanischer Forschung. Zukünftige Forschungen könnten jedoch von der Einbeziehung weiterer Stakeholder profitieren. Methoden wie die NGP ([Delbecq et al., 1986](#)), eine strukturierte Gruppentechnik, könnten hierbei hilfreich sein, um ein tiefgreifendes Verständnis von Implementationsprozessen zu erlangen und gruppenbezogene Arbeitsprozesse zu erforschen. Diese Methode fördert die Ideengenerierung und -Priorisierung und könnte nützlich sein, um MTSS effektiv in Schulen zu implementieren und relevante Daten für die Forschung zu gewinnen.

8.3 Bewertung der methodischen Triangulation

Die angewandte methodische Triangulation der verschiedenen methodischen Zugänge soll abschließend bewertet werden. Ausgangspunkt der methodischen Verbindung von quantitativen Studien ([Beitrag 2, 4 & 5](#)) ist ein systematischer Literaturüberblick ([Beitrag 1](#)), der die unterschiedlichen Forschungsschwerpunkte samt methodischen Schwerpunkten zusammenfasst und eine Empfehlung für die Beforschung von gestuften Fördersystemen gibt. Hierdurch wurde ersichtlich, dass die Effektivität einer Förderung mittels Einzelfalluntersuchungen gut zu messen ist. Die Implementationsprozesse können durch dieses Studiendesign jedoch nicht effektiv nachgezeichnet werden, sodass im Kontext des Erkenntnisinteresses hinsichtlich etwaigen Barrieren und Gelingensbedingungen eine qualitative Interviewstudie umgesetzt wurde. Die dritte Stufe der Förderung wurde ebenfalls durch eine qualitative Interviewstudie erhoben. Der Hintergrund hierfür war das Erkenntnisinteresse bezüglich der intrapersonellen Veränderungsprozesse, die Lehrkräfte hinsichtlich der Wahrnehmung und Regulation von Verhaltensproblematiken (bzw. externalisierendem Verhalten) durch das SCEP erfahren. Diese Verknüpfung verschiedener Methodiken kann ein umfassendes Gesamtbild der Prozesse rund um die Effektivität, die Implementation, die direkte Umsetzung sowie Ausgestaltung der Interventionen und der

Förderung liefern. Zusätzlich kann die partizipative Studienplanung durch den Einbezug der Schulleitungen und Lehrkräfte als Gelingensbedingung hervorgehoben werden. In jeder Studie wurde dies positiv erwähnt. Kritisch zu erwähnen ist, dass alle angewandten quantitativen und qualitativen Methoden kein besonders hohes Maß an Generalisierbarkeit der Erkenntnisse gewährleisten. Sie sind somit im Gegensatz zu größeren Studienformaten, wie z. B. Wartekontrollgruppen-Designs, limitierter bezüglich der externen Validität³¹. Im folgenden Kapitel (Kap. 8.4) sollen aufbauend auf den präsentierten Ergebnissen Überlegungen zur methodischen Weiterentwicklung der Beforschung gestufter Fördersysteme abgeleitet werden.

8.4 Denkbare Studiendesigns & methodische Zugänge im Kontext MTSS

In der Forschung zu gestuften Fördersystemen im Schulbereich konnte gezeigt werden, dass, speziell zur Untersuchung externalisierenden Verhaltens, ist der methodische Ansatz entscheidend mit der Qualität und Aussagekraft der Ergebnisse zusammenhängt. Einzelfallstudien zeigen zusätzlich die Bedeutung sorgfältiger Planung und Durchführung unter Berücksichtigung schulischer Rahmenbedingungen, um valide Ergebnisse erzielen zu können. Herausforderungen wie die Auswahl und Größe der Stichproben, der Umgang mit fehlenden Daten sowie die Auswahl der Auswertungsmethoden sind kritische Faktoren, die die Validität der Ergebnisse beeinflussen können. Zusätzlich spielt qualitative Forschung, vor allem in Bezug auf Implementationsprozesse von MTSS, eine wichtige Rolle. Die methodische Triangulation ermöglicht ein tieferes Verständnis der Effektivität und Implementation. Diese Herangehensweise offenbart jedoch auch die Grenzen der Generalisierbarkeit der Ergebnisse, was die Notwendigkeit einer kontinuierlichen methodischen Weiterentwicklung in der Forschung zu MTSS unterstreicht. Letztlich erscheint es sinnvoll und hilfreich im Kontext zukünftiger Forschung, die Ergebnisse und die methodische Analyse sowie Diskussion in eine mögliche praktische Anwendung zu übersetzen und hierbei weitere potenzielle Forschungsvorhaben zu skizzieren.

8.4.1 Gruppenbasierte empirische Forschungsdesigns

Zu den prominenten Forschungsdesigns im Kontext der Effektivitätsforschung im Bereich MTSS gehören Randomized Controlled Trials (RCTs), Regression Discontinuity Designs

³¹ Übertragbarkeit der Ergebnisse auf andere Settings

(RDDs) und Quasi-Experimental Designs (QEDs), die jeweils spezifische Stärken und Einschränkungen aufweisen (z.B. Lee & Gage, 2020).

RCTs gelten oft als höchst valide Designform in der Interventionsforschung, basierend auf einer zufälligen Zuweisung von Teilnehmer*innen zu Interventions- oder Kontrollgruppen (Stanley, 2007). Diese Randomisierung sorgt für vergleichbare Gruppen und minimiert systematische Verzerrungen, was in der MTSS-Forschung eine klare Einschätzung der Wirksamkeit von Interventionen ermöglicht. Allerdings können ethische und praktische Herausforderungen, insbesondere in realen schulischen Umgebungen, die Durchführung von RCTs erschweren. RDDs bieten eine Alternative, wenn eine Randomisierung nicht möglich oder ethisch bedenklich ist (Cattaneo & Titiunik, 2022). Sie nutzen einen vordefinierten Schwellenwert zur Gruppeneinteilung, was besonders für die Untersuchung spezieller Förderprogramme in MTSS nützlich ist. Der Hauptvorteil liegt in ihrer Fähigkeit, kausale Schlüsse ohne Zufallszuweisung zu ziehen. In ihrer externen Validität könnten sie jedoch eingeschränkt sein. Bezüglich der Forschung zum Themenbereich MTSS bieten beide Designs wertvolle Einblicke. Die Wahl des Designs hängt von den Forschungsfragen, den Rahmenbedingungen der Schule und ethischen Überlegungen ab. QEDs kommen ins Spiel, wenn RCTs aus praktischen oder ethischen Gründen nicht durchführbar sind (Miller et al., 2020). Sie unterscheiden sich von RCTs durch eine kontrollierte Gruppenzuweisung. Dies birgt das Risiko von Verzerrungen, da die Gruppen möglicherweise nicht in allen relevanten Aspekten vergleichbar miteinander sind. QEDs könnten beispielsweise eingesetzt werden, um die Effekte von Fördermaßnahmen eines MTSS (oder auch gesamte Fördersysteme und ihre Wirkung) zu untersuchen, indem Schüler*innen (oder Schulen), die diese Maßnahmen erhalten, mit denen verglichen werden, die sie nicht erhalten. QEDs bieten den Vorteil der Flexibilität und Anwendbarkeit in natürlichen Bildungsumgebungen. Allerdings erfordern sie eine sorgfältige Planung und Analyse, um Verzerrungen zu minimieren. Sie ermöglichen die Untersuchung von Interventionen in realen Bildungskontexten und bieten eine praktikable Alternative zu RCTs, wenn eine zufällige Zuweisung nicht möglich ist.

Durch Beitrag 1 ist bekannt, dass die Anwendung von RCTs und QEDs weltweit erfolgte und valide Ergebnisse erzielen konnte. Insbesondere in den USA sind jene größeren Gruppenvergleiche (mit gewissen angesprochenen Limitationen) durchgeführt worden. Hierbei wurden verschiedene Outcome-Measures im Vergleich zwischen einer randomisierten Kontrollgruppe (RCT) oder einer nicht randomisierten Kontrollgruppe (QED) überprüft, sodass die Wirksamkeit des eingeführten MTSS in der Interventionsgruppe quantifiziert werden konnte. Jene Designs sind sich in der Planung und Umsetzung oftmals sehr ähnlich. Bei

entsprechend großer Stichprobe sind sie in einem deutschen Bildungskontext ebenfalls umsetzbar, wie die Studien von [Voß et al. \(2014\)](#) und [Voß et al. \(2016\)](#) zeigen konnten. Weitaus weniger weit verbreitet sind RDD-Ansätze, deren Verwendung im kommenden Kapitel skizziert werden soll.

8.4.2 Skizzierung eines Regression Discontinuity Designs

Neben RCT- und QED-Designs sind RDD-Ansätze eine weitere Möglichkeit, die Effektivität von verhaltensbezogenen Interventionen im Kontext eines MTSS zu überprüfen ([z.B. Losinski et al., 2021](#)). Diese gestalten sich jedoch aufgrund gewisser Anforderungen in der Planung komplexer.

In einer hypothetischen Studie, welche die Wirksamkeit eines MTSS zur Reduktion externalisierenden Verhaltens bei Schüler*innen untersuchen würde, könnte die Interventionsgruppe z. B. durch einen Cut-Off-Wert einer Skala der ITRF ([Casale et al., 2018](#)) bestimmt werden³². Potenzielle weitere Variablen wären frei zu wählen, um so den Fördererfolg von z.B. einer Leseintervention oder einem Training zum Sozial-Emotionalen Lernen zu überprüfen. Schüler*innen über dem Cut-Off Wert würden folglich der Interventionsgruppe zugeordnet werden und somit die Förderung durch die Stufen 2 und 3 des jeweiligen Ansatzes (Leseintervention oder Training zum Sozial-Emotionalen Lernen) erhalten. Die erste Förderstufe (universelle Maßnahmen) erhalten jedoch auch alle anderen Kinder. Auf diese Weise wäre die ethische Vertretbarkeit des Forschungsvorhabens höher einzuschätzen als bei Kindern, die keine Förderung (Kontrollgruppe in einem RCT z. B.) erhalten. Eine Einschränkung wäre jedoch, dass ein solches Design vorrangig den Fördererfolg der Stufen 2 und 3 abbilden würde und die Stufe 1 aufgrund der Intervention in Kontrollgruppe sowie Interventionsgruppe nicht erheben könnte. Gewisse *Spillover*-Effekte ([z.B. Karatas et al., 2023](#)) der Interventionen könnten ebenfalls nicht ausgeschlossen werden. Dennoch könnte auf diese Art ein Großteil eines MTSS auf Schulebene durch eine Erhebung in Form eines gruppenbasierten empirischen Forschungsdesign überprüft werden. Ein weiterer Vorteil wäre, dass keine weitere Schule in den Prozess einbezogen werden müsste, die nicht durch die angewandten Interventionen profitieren würde (Kontrollgruppe). Auch können andere Setting-Bedingungen den Vergleich der Gruppen erschweren. Als Outcome-Measure für eine solche

³² Bspw. alle Kinder mit einem auffälligen Wert im Gesamtproblemwert

Studie könnte die Erhebung von z. B. störendem Verhalten im Unterricht, aggressivem Verhalten oder On-Task-Verhalten³³ dienen.

Auszuwerten wären die Daten mittels RDD-Analysen ([Baird et al., 2024](#); [Eggers et al., 2017](#)). Diese Auswertungsmethode basiert auf einem Graphen der abhängigen Variable mit Pretest-Scores auf der X-Achse und Posttest-Scores auf der Y-Achse. Der Graph wird dann verwendet, um die Daten auf Hinweise auf eine Diskontinuität oder einen Unterschied zwischen den Steigungen um den Cut-Off-Wert herum zu überprüfen. Statistische Tests und Effektgrößen werden aufbauend hierauf geschätzt, um die Diskontinuität zu bewerten. Ergänzend könnte ein nichtparametrisches RDD-Modell mit lokaler linearer Regression und Bandbreitenbeschränkung berechnet werden. Auf diese Weise wäre die Annahme, dass die Gruppen direkt unterhalb und oberhalb des Cut-Offs hinsichtlich ihrer Eigenschaften ähnliche Werte aufzeigen, zu überprüfen. Basierend auf diesen Ergebnissen können kausale Schlussfolgerungen über die Wirksamkeit des MTSS (Stufen 2 und 3) abgeleitet werden. Außerdem könnte überprüft werden, ob die Beziehung zwischen der Zuweisungsvariablen (z.B. ITRF-Score) und den Outcome-Maßen (z.B. Verhalten) konsistent und ohne Diskontinuitäten durch den Cut-Off-Wert, verlief. Diese Überprüfungen wären entscheidend, um sicherzustellen, dass die Ergebnisse gültig und valide sind.

8.4.3 Weitere vielversprechende methodische Zugänge

Wie bereits dargelegt, haben sich im Kontext der Forschung zum Bereich MTSS Forschungsdesigns wie RCTs oder QEDs als Standard etabliert. Die Möglichkeit der Anwendung von RDD-Designs wurde im vorherigen Kapitel ergänzend erläutert. Um jedoch ein umfassenderes Verständnis des MTSS-Forschungsprozesses zu gewährleisten, ist es empfehlenswert, diese Ansätze durch qualitative Ansätze zu ergänzen. Die qualitative Forschung bietet ergänzend durch die (direkte) Analyse der Einstellungen, der Haltung oder des Wissens vertiefte Einblicke in Prozesse, die jene Outcome-Measures aus größeren quantitativen Ansätzen unbemerkt moderieren könnten. Solche Ansätze ermöglichen es somit, Verständnisse der individuellen und kollektiven Erfahrungen, Wahrnehmungen und Einstellungen zu gewinnen und dadurch einen ganzheitlicheren Blick auf die Wirksamkeit und Umsetzbarkeit von gestufter Förderung zu bekommen.

³³ Lernbezogenes Arbeitsverhalten

Denkbare Forschungsthemen wären dementsprechend qualitative Analysen des bestehenden Bildungs- und Schulsystems samt passenden Interventionsmöglichkeiten im Kontext externalisierenden Problemverhaltens aus Sicht sonderpädagogischer Lehrkräfte, Psycholog*innen oder beispielsweise dem Jugendamt. Auf diese Weise könnten potenzielle Kooperationsprozesse verschiedener an der Förderung beteiligter Gruppen identifiziert werden. Darüber hinaus wäre eine Untersuchung bezüglich der Einstellung und Haltung der Schulleitungen zur Einführung gestufter Fördersysteme an ihrer Schule sowie bezüglich konkreter Umsetzungswege ein relevanter Forschungsbereich. Hieraus könnten weitere Bedingungsfaktoren für die Implementation abgeleitet werden. Zusätzlich konnte durch Beitrag 6 identifiziert werden, dass Lehrkräfte in Regelschulsystemen weitere Schulungen im Umgang mit schulischer Diagnostik und der Auswertung und Bewertung von Daten benötigen. Vertiefte Analysen hierzu könnten die schulinterne, nachhaltige Umsetzung von MTSS stärken. Im Falle bereits bestehender, implementierter MTSS-Ansätze könnten erweiterte Gruppeninterviewtechniken zum Einsatz kommen, um konkurrierende Sichtweisen und Einstellungen zwischen verschiedenen Stakeholdern einer Schule zu ergründen und somit Prozesse innerhalb eines gestufter Fördersystems anzupassen.

9 Abschließende Diskussion

Die präsentierten Ergebnisse erweitern das Wissen über MTSS in der Grundschule in Deutschland. Die Forschung zeigt, dass verhaltensmodifizierende Interventionen positive Schulerfahrungen fördern und negative Bildungseinflüsse verringern können. Dies steht im Einklang mit den Ergebnissen der internationalen Forschungslandschaft (Gage, Whitford, & Katsiyannis, 2018; Lee & Gage, 2020; Solomon et al., 2012).

Besonders hervorzuheben ist, dass kein MTSS-Ansatz auf Schulebene in den Recherchen der hier präsentierten Arbeit mit der gleichen Intensität überprüft wurde. Dies schließt auch die verschiedenen Forschungszugänge ein, die für die Analysen gewählt wurden. Die gesammelten Daten konnten zusätzlich methodisch trianguliert werden, was als eine Besonderheit des Vorgehens und der Auswertung, selbst im internationalen Gesamtkontext bestehender Analysen, betrachtet werden kann. Dementsprechend stellen die Ergebnisse in ihrem Umfang auch in Deutschland eine einzigartige Datengrundlage dar, die in der Form und in dem Umfang bisher nicht erhoben wurde. Sie erweitern die deutschen Ergebnisse und bilden Evidenz für die Effektivität eines MTSS mit vorrangigem Fokus auf verhaltensmodifizierenden Maßnahmen. Neben der zuvor erfolgten kritischen Analyse und Diskussion bezüglich der eingesetzten Methodik sollen die Kernergebnisse der Beiträge kurz zusammenfassend dargelegt werden, um nachfolgend eine mögliche Weiterentwicklung des Ansatzes *Multimo* zu skizzieren.

9.1 Beantwortung der Leitfragen 1-5

Im Folgenden werden die Kernergebnisse der Beiträge kurz eingeordnet und mit bestehender Forschung in Verbindung gesetzt, um eine Grundlage für die Weiterentwicklung des *Multimo*-Ansatzes abzuleiten. Zunächst werden die Leitfragen 1-5 kurz auf Grundlage der präsentierten Ergebnisse beantwortet, woran die Beantwortung und Diskussion der Leitfrage 6 anschließt.

9.1.1 Beantwortung der Leitfragen 1 & 2: Der internationale Vergleich von MTSS

Beitrag 1 bietet einen umfassenden Überblick über MTSS in Grundschulen mit dem Schwerpunkt auf verhaltensmodifizierenden Interventionen. Er zeigt, dass MTSS hauptsächlich in westlichen Ländern verbreitet sind und sich vorrangig auf Schüler*innen konzentrieren, wobei eine Erweiterung auf Lehrkraftinterventionen vorgeschlagen wird. Studien von Lee und

Gage (2020), Gage, Whitford und Katsiyannis (2018) und Fox et al. (2021) bestätigen diese Ergebnisse für den US-amerikanischen Raum und stützen die Erkenntnis, dass die Implementation von MTSS erfolgreicher bei engagierten Lehrkräften und Schulleitungen ist. Zusätzlich wird deutlich, dass die Wirksamkeit von MTSS als Gesamtsystem selten untersucht wird und dass die Notwendigkeit besteht, verschiedene Ebenen und Wirkmechanismen von MTSS zu erforschen. Gleichermaßen kann die Bedeutsamkeit der Partizipation aller Beteiligten an der Entwicklung der Ansätze abgeleitet werden, was andere Studien ebenfalls bestätigen konnten (Bettors-Bubon et al., 2022; Gerzel-Short, 2018). Insgesamt bestätigt die Studie die Effektivität von MTSS, zeigt jedoch Forschungslücken insbesondere in Bezug auf die detaillierte Untersuchung der Wirksamkeit einzelner MTSS-Komponenten auf.

9.1.2 Beantwortung der Leitfragen 3 & 4: Effektivität von MTSS in der Grundschule

Beitrag 2 legt den Schwerpunkt auf die Überprüfung des GBG und der DBRC im Kontext der Reduktion von externalisierendem Verhalten. Es zeigt sich, dass das GBG das disruptive Verhalten signifikant reduziert. Die Implementation des GBG bringt somit positive Effekte, die ähnlich groß sind wie in anderen Studien (Leidig et al., 2022). Die Stufe 2 des Ansatzes, in Form der Einführung der DBRC, zeigt eine weitere leichte Reduzierung des disruptiven Verhaltens, welches aber, verglichen mit vorherigen Studien, weniger stark verringert werden konnte (Ackley, 2022; Iznardo et al., 2020). Der Beitrag betont die Notwendigkeit der weiteren Überprüfungen der Maßnahmen.

Beitrag 3 konzentriert sich auf die Wirksamkeit des *Multimo*-Konzepts und untersucht Wirkfaktoren des SCEP auf der Stufe 3. Der Fokus der Stufe 3 liegt verstärkt auf der Entwicklung von Reflexionsprozessen bei Lehrkräften und hebt hervor, dass Coachingprozesse nicht nur das Schüler*innenverhalten beeinflussen (können), sondern auch die Selbstwirksamkeit der Lehrkräfte steigern können. Die Ergebnisse der Studie korrespondieren mit Ergebnissen von Hanisch et al. (2020). Zusätzlich kann die Reduktion des disruptiven Verhaltens für vier von acht Fällen auf Stufe 3 nachgewiesen werden (Kap. 5.4).

Beitrag 4 und 5 befassen sich mit der Effektivität des GBG und der DBRC in Form von Einzelfallstudien. Sie zeigen, dass beide Interventionen in der Lage dazu sind, disruptives Verhalten (beide Studien) sowie spezifische Zielverhaltensweisen (Beitrag 5) zu reduzieren und bestätigen damit Ergebnisse internationaler Forschung (Ackley, 2022; Joslyn et al., 2019; Miller et al., 2017; Rubow et al., 2018; Spilles et al., 2019; Vannest et al., 2019; Wiskow et al., 2021). Interessant ist hierbei die Beobachtung, dass sich die Effekte des GBG nicht nur auf die Spielphasen beschränken, sondern den gesamten Schultag positiv beeinflussen und dass die

Reduktion spezifischer Verhaltensweisen, die als Interventionsziel angegeben werden, und das allgemeine disruptive Verhalten eines Kindes miteinander in Verbindung stehen.

9.1.3 Beantwortung der Leitfrage 5: Implementationsbedingungen von MTSS in der Grundschule

Beitrag 6 thematisiert abschließend die praktische Umsetzung und Herausforderungen von MTSS in Grundschulen. Lehrkräfte erkennen den Nutzen der systematischen Förderung, sehen sich jedoch mit Herausforderungen konfrontiert, insbesondere in Bezug auf schulische Diagnostik und die Notwendigkeit einer kompetenten Durchführung und Auswertung. Der Beitrag erweitert die Ergebnisse US-amerikanischer Implementationsforschung (Fox et al., 2021).

9.2 Diskussion des Förderansatzes Multimo

Auf der Grundlage der Ergebnisse konnten verschiedene Bereiche identifiziert werden, die eine Erweiterung und Adaption des Konzeptes sinnvoll erscheinen lassen. Dies dient in erster Linie dazu, den allgemeinen Fördererfolg des Ansatzes zu erhöhen und das System an die Bedürfnisse der Schüler*innen und Lehrkräfte anzupassen. Grundlage für die Überarbeitung des Konzeptes sind die bereits vorgestellten Ergebnisse samt Kontextualisierungen durch weitere Studien und Projekte samt Konzept-Ideen (z.B. Farmer et al., 2020; Lee & Gage, 2020; Simonsen, 2021; Sugai & Horner, 2019).

Insbesondere die Daten und Ergebnisse der Interviewstudien (Beitrag 3 & 6) können zusätzlichen Aufschluss über potenziell zu verbessernde Bereiche des Ansatzes liefern. Ziel ist somit auf Kritikpunkte, die im Rahmen der Interviewstudien bei einigen beteiligten Personen anklingen, zu reagieren und gepaart mit den Ergebnissen der Wirksamkeitsstudien eine konzeptuelle Überarbeitung von *Multimo* anzuregen. Ein erster erkennbarer Faktor ist die mangelnde Flexibilität des Systems. Dies bezog sich vorrangig auf die Auswahl der Maßnahmen und die Implementation dieser in den Schul- bzw. Klassenalltag (Beitrag 3 & 6). Häufig waren zudem die Implementationszeitpunkte, die durch die äußeren Umstände der wissenschaftlichen Begleitstudie terminiert wurden, ungünstig und führten zu einer Mehrbelastung. Überdies erscheint der Ansatz auch im wissenschaftlichen und internationalen Vergleich ähnlicher Fördersysteme als weniger flexibel und adaptiv, was allerdings auch mit der Erst-Implementation dieses Ansatzes zusammenhängen kann.

Grundsätzlich kann die erste Förderstufe jedoch als effektiv und wirksam eingeschätzt werden (Beitrag 2, 4 & 5). Diese Befunde decken sich mit internationalen Studienergebnissen (z.B. Leidig et al., 2022; Pennington & McComas, 2017; Rubow et al., 2018). Kritisch anzumerken ist, dass das GBG mit der Zeit an Reiz verlieren kann und Lehrkräfte das Spiel im Unterrichtsalltag nach einiger Zeit weniger erfolgreich einsetzen. Die Einzelfallstudien und die qualitative Analyse können dies in Teilen bestätigen (Beitrag 2, 4 & 6). Vereinzelt klang zudem das Interesse an weiteren Fortbildungen an, die im Sinne eines effektiven Classroom Managements Strategien bereithalten, um flexibel auf externalisierendes Verhalten reagieren zu können.

Die zweite Förderstufe wurde im Vergleich zu der ersten Förderstufe als weniger wirksam und effektiv beschrieben. Dies widerspricht den isolierten Befunden aus Beitrag 5³⁴, deckt sich allerdings mit den Befunden aus Beitrag 2. Im Rahmen der Implementation trat zudem das Problem auf, dass die Eltern nicht erfolgreich eingebunden werden konnten. Darüber hinaus stellte die erfolgreiche Zielformulierung im Kontext der DBRC in Teilen ein Problem für die Lehrkräfte dar (Beitrag 6). Die Multimodalität wurde jedoch insgesamt als hilfreich und sinnvoll bewertet.

Zusätzlich zeigte sich, dass die pädagogische Geschlossenheit, die für die erfolgreiche Umsetzung eines gestuften Fördersystems erforderlich ist, an einigen Schulen nicht umfänglich gegeben war (Beitrag 6). Hierunter fällt auch die Tatsache, dass der Ansatz zwar in einigen Klassen erfolgreich implementiert wurde, dass aber eine schulweite Auseinandersetzung bezüglich der Umsetzungen und Adaptionen nicht gegeben war. Nicht immer waren sich die verschiedenen Kollegien innerhalb der Schulen einig über pädagogische Vorgehensweisen. Ebenso waren sie an Ausgestaltung der gestuften Förderung nicht gleichermaßen involviert. Dies stellt schlussendlich vorrangig eine methodische Limitation dar, die durch diverse Implementationsstudien gestützt wird und bei zukünftigen Studien beachtet werden sollte (Fox et al., 2021).

Als weiterer Kritikpunkt, der bei einer Neuausrichtung des Ansatzes mitbedacht werden sollte, sind fehlende Kenntnis bzw. fehlendes Wissen über die Wirkweise von mehrstufigen Fördersystemen (Beitrag 6). Dieses ist für die erfolgreiche Implementation notwendig, konnte allerdings durch das Multiplikator*innenkonzept in Teilen nicht erfolgreich transportiert werden. Ebenfalls fehlte das Verständnis des diagnostischen Prozesses, welcher als Grundlage

³⁴ Die Stichproben von Beitrag 2 und Beitrag 5 sind nicht identisch. Zusätzlich wurde in Beitrag 2 das gesamte Fördersystem und in Beitrag 5 lediglich die Wirksamkeit der DBRC untersucht.

für die Förderentscheidung zu sehen ist. Praktisch bedeutet dies, dass die Lehrkräfte die verschiedenen Ebenen der gestuften Förderung oftmals nicht überblicken konnten und gewisse Prozesse als unnütz oder Belastung empfanden – wie z. B. eine fortlaufende diagnostische Überprüfung (Beitrag 6).

9.3 Adaptionmöglichkeiten des Förderansatzes *Multimo*

Aus allen Ergebnissen der Beiträge lassen sich mögliche Weiterentwicklungen und Adaptionmöglichkeiten des Förderansatzes *Multimo* ableiten. Die zunächst bemängelte Flexibilität könnte im Sinne des *Adaptive Support* (Farmer et al., 2020) erhöht werden, sodass den Lehrkräften im Rahmen der gestuften Förderung verschiedene Interventionen und pädagogische Handlungsmöglichkeiten bereitgestellt werden, um auf externalisierendes Verhalten zu reagieren. Dies umfasst über die drei Stufen hinweg die Optimierung von Unterrichtsroutinen, die Förderung des Engagements und der akademischen Leistung der Schüler*innen unter Berücksichtigung der sozialen Beziehungen im Klassenzimmer, gepaart mit präventiven und interventiven pädagogischen Methoden. Der Ansatz unterstreicht die Bedeutung jener individualisierten Ausrichtung, die auf die spezifischen Bedürfnisse und Umstände jedes Kindes eingeht. Eine systematische, aber flexible Herangehensweise ist laut der Autor*innengruppe erforderlich, die sowohl regelmäßiges (datenbasiertes) Monitoring als auch individuelle Anpassungen beinhaltet, um die Interventionen effektiv umzusetzen.

Ergänzend kann der Einsatz der *funktionalen Verhaltensanalyse* (Functional Behavioral Assessments, FBA Gresham et al., 2001; O’Neill et al., 1997) hilfreich sein, um Problembereichen möglichst umfassend und nachhaltig zu begegnen. Die Einbeziehung jenes Verfahrens in die Planung von Verhaltensinterventionen gilt in den USA bereits als bewährte Praxis (Quinn et al., 2001). Der Einsatz der FBA kann z. B. Umweltfaktoren, die problematisches Verhalten beeinflussen, bestimmen und durch direkte Beobachtungen oder Interviews sowie Verhaltensbewertung dazu beitragen, jene Faktoren abzuschwächen oder sie zu verändern. Die Umsetzung sieht vor, dass das problematische Verhalten zunächst identifiziert und anschließend systematisch analysiert wird. In der Phase der Analyse werden die gesammelten Informationen ausgewertet, um Muster zu erkennen und die Funktion des problematischen Verhaltens zu verstehen. Auf der Grundlage dieser Analyse werden dann Interventionsstrategien entwickelt, die speziell auf die individuellen Bedürfnisse des Kindes zugeschnitten sind (Gresham et al., 2001). Studien zeigen, dass Interventionen, die auf der Anwendung von FBA basieren, signifikant effektiver sind als solche, die es nicht tun (Miller &

Lee, 2013). Im Rahmen einer Adaption des *Multimo*-Ansatzes könnte die Einführung systematischer FBA bereits auf der selektiven Stufe hilfreich sein.

Werden solche Maßnahmen schulweit festgelegt (und gemeinsam diskutiert), könnte dies zusätzlich Auswirkungen auf die pädagogische Geschlossenheit und die Akzeptanz des Systems haben. Hiermit verbunden ist auch die gesamte Konzeptentwicklung, die laut der Studien von Sadjadi et al. (2022) und Fox et al. (2021) im besten Fall auch von den Lehrkräften und der Schulleitung ausgeht, um ein möglichst passendes System für die jeweiligen Schulen zu entwickeln. Ergänzend hierzu sollten schulinterne Fortbildungen zum Thema „schulische Diagnostik“ und „Wirkweise & Funktion gestufter Förderung“ implementiert werden (Beitrag 6).

Eine Möglichkeit *Multimo* zu erweitern und gleichzeitig eine nachhaltige und effektive Förderung sicherzustellen, wäre demnach die Einführung von Kompetenzteams (MTSS-Teams) in der Schule. Diese Teams sind in der US-amerikanischen Schullandschaft bereits fester Bestandteil eines jeden MTSS (Schaffer, 2023). Sie sind verantwortlich für die Ausgestaltung, Adaption sowie Umsetzung und koordinieren die Entscheidungsprozesse im Kontext der Diagnostik. Den MTSS-Teams kommen nach der Analyse möglicher Adaptionsbereiche umfassende Aufgaben zu. Sie sollten aus ca. drei bis fünf Personen bestehen, die jegliche Entscheidungen rund um die Förderung und die Anpassung der Interventionen gemeinsam treffen und abstimmen. Denkbar wäre die jahrgangsbezogene Aufstellung dieser Teams, sodass sich für jeden Jahrgang (z. B. alle 2. Klassen) ein Team formieren würde, das für weitere Lehrkräfte, pädagogische Fachkräfte und andere Mitarbeitende als beratende Instanz zur Verfügung steht. Zusätzlich könnte die Verantwortung der MTSS-Teams darin bestehen, die erfolgreiche Umsetzung und Auswertung der Verlaufsdiagnostik zu koordinieren. Auf diese Weise könnte sichergestellt werden, dass eine kollegiumsübergreifende Zusammenarbeit samt intensiver Betrachtung des Fördererfolgs erfolgt. Die mangelnden zeitlichen und personellen Ressourcen, die in Beitrag 6 angesprochen wurden, könnten jedoch ein starkes Hindernis für die tatsächliche praktische Umsetzung jener MTSS-Teams darstellen.

Solche Teams wurden in der ersten Implementation des Ansatzes nicht integriert. Stellvertretend hierfür fungierte das Projektteam der Universität. Als Gelingensbedingung für die erfolgreiche Implementation wird jedoch ein schulinternes Team mit Kenntnis der genauen schulischen Rahmenbedingungen angesehen (Fox et al., 2021). Zusammenfassend ergeben sich folgende Aspekte, die bei der Überarbeitung des *Multimo*-Ansatzes mitbedacht werden sollten:

- 1) Bildung von MTSS-Teams an den Schulen, die für die Umsetzung und Implementation des MTSS verantwortlich sind.
- 2) Einführung weiterer universeller Strategien wie Classroom Management und Strategien guten Unterrichts.
- 3) Einführung der FBA (auf selektiver Ebene), um Verhaltensauffälligkeiten besser zu verstehen und ausgehend davon Verhaltensziele sowie Interventionen abzuleiten.
- 4) Entwicklung vertiefender Fortbildungen zum Umgang mit schulischer Diagnostik und datengeleiteten Förderentscheidungen.

10 Fazit & Ausblick

Die Kernergebnisse des *Multimo*-Ansatzes betonen die Effektivität des Ansatzes hinsichtlich der Reduktion von externalisierendem Verhalten in der Grundschule. Auf Grundlage der Ergebnisse konnten kritische Bereiche identifiziert werden, die durch eine Überarbeitung adressiert werden können. Ein Kernergebnis ist die Forderung bzw. der Wunsch nach einer Erhöhung der Flexibilität (*Adaptive Support*), wodurch Lehrkräften vielfältige Interventionen und Methoden zur Verfügung gestellt werden, um besser auf problematisches Verhalten reagieren zu können. Ergänzend kann die FBA zur umfassenden Analyse und Entwicklung individueller Interventionsstrategien genutzt werden. Die Bildung von MTSS-Teams in Schulen könnte diesen Prozess ergänzen und stützen. Aufbauend hierauf können und sollen Weiterbildungen und schulinterne Fortbildungen das Verständnis für schulische Diagnostik und gestufte Förderung stärken. Schließlich sollte die gesamte Schule in die Entwicklung und Anwendung des Konzepts einbezogen werden, um die Akzeptanz des MTSS und die pädagogische Geschlossenheit zu erhöhen. Diese Maßnahmen zielen darauf ab, *Multimo* an die spezifischen Bedürfnisse der Schüler*innen und Schulen anzupassen, wobei die Herausforderungen begrenzter Ressourcen berücksichtigt werden müssen.

Zusätzlich ist im Rahmen der Umsetzung der Studien der Themenbereich Theorie-Praxis-Transfer in den Fokus des Projektteams gerückt. Viele Schulen und Lehrkräfte erreicht die wissenschaftliche Forschung nicht, sodass die Einführung jener Maßnahmen und Systeme in die Schule ausbleibt. Maßgeblichen Einfluss hat eine in sich kohärente, aber in Teilen exklusive Sprechart von Wissenschaft, die eine leicht zugängliche und uneingeschränkte Distribution der Ergebnisse verhindert oder zumindest erschwert (Schemme et al., 2017). Eine Möglichkeit, solche Übersetzungsproblematiken zu verhindern bzw. ihnen entgegenzukommen, können praxisnahe Fortbildungen darstellen, die jenen Transfer der Ergebnisse hin zu ihrer Anwendung erleichtern. Die Komplexität der Forschungsergebnisse wird auf diese Weise reduziert und anwendungsnah aufbereitet.

Um flächendeckend MTSS-Systeme und ihre Wirkweisen in Deutschland zu implementieren und sie einzusetzen, bedarf es bei einer Anzahl von etwa 15.000 Grundschulen (KMK, 2024) in Deutschland einen immensen zeitlichen und personellen Aufwand. Im Feld der Professionalisierung von Lehrkräften hat sich unabhängig hiervon, sondern vielmehr angetrieben durch die Einsparung finanzieller Mittel und der COVID-19 Pandemie, die Umsetzung online-basierter Fortbildungen etabliert. Nicht alle pädagogisch-schulischen

Themen eignen sich für die Umsetzung in Form einer Online-Fortbildung, jedoch können *blended-learning* Formate ([Hrastinski, 2019](#)) ebenfalls eine hohe Effektivität nachweisen und eignen sich daher, um niederschwellige Inhalte flächendeckend zu distribuieren. Die Vorbereitung solcher Formate findet durch die Erstellung digital aufbereiteter Inhalte statt, die entweder in einzelnen Dokumenten oder über eigene Fortbildungs-Homepages bereitgestellt werden. Im Sinne der Ergebnisse bietet es sich an, eine eigene onlinebasierte Fortbildungs-Homepage im Kontext MTSS-Implementation und Ausgestaltung aufzubauen. Um diesem praktischen Nutzen zu entsprechen, sind bereits erste Fortbildungsmodule erarbeitet und erstellt worden, die auf der Website mtss.info im Jahr 2024 veröffentlicht werden sollen. Die Website soll einen übersichtlichen und verständlichen Einblick in die Vorgänge und Verfahrensweisen der gestuften Förderung, ihrer Anwendung und Evidenz samt Beispielen und Umsetzungsideen geben. Von den präsentierten Inhalten ausgehend können Fortbildungsformate in den Schulen angeboten werden, die auf bereits vorhandenes und in Teilen erprobtes Wissen (sowie dessen Umsetzungen) zurückgreifen. Im Zuge dessen findet auch die Entwicklung und Erprobung einer App-basierten Verhaltensmessung (Verlaufsmessung) statt, deren Quellcode und Installationsdateien in den gängigen App-Stores sowie auf der Website zur Verfügung gestellt werden.

Diese Schritte der Transformation und Übertragung der Ergebnisse in die schulische Praxis sind jedoch letztlich von dem Willen und der Bereitschaft der Schulen, Eltern, Gemeinden und vor allem der Politik abhängig. Es konnte gezeigt werden, dass gestufte Förderung eine Möglichkeit darstellt, auf verhaltensbezogenen Problematiken im schulischen Kontext zu reagieren und effektiv jene Verhaltensdimensionen zu adressieren. Auch wenn eine flächendeckende Einführung jener Systeme deutschlandweit in weiter Ferne scheint, so können Schulen und Lehrkräfte mit eigenem Engagement dazu beitragen, die Ergebnisse dieser Arbeit in das schulpraktischen Feld zu übertragen. Die Weiterentwicklung der bestehenden Systeme ist hierbei ein wesentlicher Faktor gestufter Förderung und darf nicht als abgeschlossen betrachtet werden. Es wird folglich nie *das* MTSS geben, das perfekt auf jede Schule oder Schulform passt. Allerdings es gibt Verfahrensweisen und (notwendige) Bedingungen, die pädagogisches (Zusammen)Arbeiten stärken und fördern können, um somit besser auf externalisierendes Problemverhalten reagieren zu können. Dieser langfristig positive Effekt auf Ausgrenzung- und Belastungsfaktoren von Kindern und Jugendlichen mit externalisierendem Verhalten sowie die positiven Auswirkungen auf die psychische Belastung und Gesundheit der Lehrkräfte sollte alleinig schon Antrieb sein, gestufte Förderung zu implementieren.

Literaturverzeichnis

- Aasheim, M., Drugli, M. B., Reedtz, C., Handegård, B. H., & Martinussen, M. (2018). Change in teacher–student relationships and parent involvement after implementation of the Incredible Years Teacher Classroom Management programme in a regular Norwegian school setting. *British Educational Research Journal*, 44(6), 1064-1083. <https://doi.org/10.1002/berj.3479>
- Ackley, M. M. (2022). *Meta-Analysis of Daily Behavior Report Cards* (Publication Number 29261938) [Ph.D., The University of Southern Mississippi]. United States - Mississippi.
- Agyapong, B., Obuobi-Donkor, G., Burbach, L., & Wei, Y. (2022). Stress, Burnout, Anxiety and Depression among Teachers: A Scoping Review. *Int J Environ Res Public Health*, 19(17), 1-42. <https://doi.org/10.3390/ijerph191710706>
- Ahnert, L., Harwardt-Heinecke, E., Kappler, G., Eckstein-Madry, T., & Milatz, A. (2012). Student-teacher relationships and classroom climate in first grade: how do they relate to students' stress regulation? *Attach Hum Dev*, 14(3), 249-263. <https://doi.org/10.1080/14616734.2012.673277>
- Albrecht, S. F., Mathur, S. R., Jones, R. E., & Alazemi, S. (2015). A School-wide Three-tiered Program of Social Skills Intervention: Results of a Three-year Cohort Study. *EDUCATION AND TREATMENT OF CHILDREN*, 38(4), 565-586. <https://doi.org/10.1353/etc.2015.0023>
- Aldabbagh, R., Glazebrook, C., Sayal, K., & Daley, D. (2022). Systematic Review and Meta-Analysis of the Effectiveness of Teacher Delivered Interventions for Externalizing Behaviors. *J Behav Educ*, 1-42. <https://doi.org/10.1007/s10864-022-09491-4>
- Algozzine, B., Barrett, S., Eber, L., George, H., Horner, R., Lewis, T., Putnam, B., Swain-Bradway, J., McIntosh, K., & Sugai, G. (2014). School-wide PBIS Tiered Fidelity Inventory. OSEP Technical Assistance Center on Positive Behavioral Interventions and Supports. *pbis.org*.
- Algozzine, K., & Algozzine, B. (2007). Classroom Instructional Ecology and School-Wide Positive Behavior Support. *Journal of Applied School Psychology*, 24(1), 29-47.
- Allen-Meares, P., Montgomery, K. L., & Kim, J. S. (2013). School-based social work interventions: a cross-national systematic review. *Soc Work*, 58(3), 253-262. <https://doi.org/10.1093/sw/swt022>
- Alresheed, F., Hott, B. L., & Bano, C. (2013). Single Subject Research: A Synthesis of Analytic Methods. *The Journal of Special Education Apprenticeship*, Vol. 2, 1-18.

- Andreou, T. E., McIntosh, K., Ross, S. W., & Kahn, J. D. (2014). Critical Incidents in Sustaining School-Wide Positive Behavioral Interventions and Supports. *The Journal of Special Education, 49*(3), 157-167. <https://doi.org/10.1177/0022466914554298>
- APA. (2018). *Diagnostisches und Statistisches Manual Psychischer Störungen DSM-5®* (P. Falkai, H.-U. Wittchen, M. Döpfner, W. Gabel, W. Maier, L. Diatara, W. Rief, H. Saß, & M. Zaudig, Hrsg.). Hogrefe. <https://doi.org/10.1026/02803-000>
- Arco, L. (2003). Review of Positive Behavioural Support in the Classroom: Principles and Practices. *Behaviour Change, 20*(4), 233-234.
- Avant, D. W. (2016). Using Response to Intervention/Multi-Tiered Systems of Supports to Promote Social Justice in Schools. *Journal for Multicultural Education, 10*(4), 507-520.
- Ayano, G., Abraha, M., Tsegay, L., & Gizachew, Y. (2024). Umbrella Review of the Global Prevalence of Conduct Disorder in Children and Adolescents. *Psychiatr Q, 95*(1), 173-183. <https://doi.org/10.1007/s11126-023-10060-9>
- Babineau, J. (2014). Product Review: Covidence (Systematic Review Software). *Journal of the Canadian Health Libraries Association / Journal de l'Association des bibliothèques de la santé du Canada, 35*(2), 68-71. <https://doi.org/10.5596/c14-016>
- Bachmann, C. J., Lehmkuhl, G., Petermann, F., & Scott, S. (2010). Evidenzbasierte psychotherapeutische Interventionen für Kinder und Jugendliche mit aggressivem Verhalten. *Kindheit und Entwicklung, 19*(4), 245-254. <https://doi.org/10.1026/0942-5403/a000032>
- Baird, K., Byrne, A., Cockayne, S., Cunningham-Burley, R., Fairhurst, C., Adamson, J., Vernon, W., Torgerson, D. J., & trial, R. (2024). Can routine assessment of older people's mental health lead to improved outcomes: A regression discontinuity analysis. *PloS one, 19*(3). <https://doi.org/10.1371/journal.pone.0300651>
- Bambara, L. M., Nonnemacher, S., & Kern, L. (2009). Sustaining School-Based Individualized Positive Behavior Support. *Journal of Positive Behavior Interventions, 11*(3), 161-176. <https://doi.org/10.1177/1098300708330878>
- Barrish, H. H., Saunders, M., & Wolf, M. M. (1969). Good behavior game: effects of individual contingencies for group consequences on disruptive behavior in a classroom. *J Appl Behav Anal, 2*(2), 119-124. <https://doi.org/10.1901/jaba.1969.2-119>
- Bartling, A., Spilles, M., Kluge, J., Gottfried geb. Neroznikova, K., Huber, C., Hennemann, T., König, J., Fussangel, K., Kaspar, K., Gräsel, C., Melzer, C., Strauß, S., & Grosche, M. (2021). Partizipation in einem

- Response-to-Intervention-Modell für den Förderschwerpunkt Emotionale und soziale Entwicklung (PARTI): Beschreibung und Evaluation einer praxisorientierten Fortbildungsreihe. *Zeitschrift für Heilpädagogik*, 72, 664–675.
- Batsche, G. (2014). *Multi-Tiered System of Supports for Inclusive Schools*. Routledge. <https://doi.org/10.4324/9780203102930.ch14>
- Baumgardt, I. (2016). Vielfalt im Klassenzimmer – Chancen und Herausforderungen auf dem Weg zur inklusiven Grundschule. In *Vielfalt als Chance* (S. 281-288). <https://doi.org/10.14220/9783737006880.281>
- Baving, L., Laucht, M., & Schmidt, M. H. (2003). Frontal EEG correlates of externalizing spectrum behaviors. *Eur Child Adolesc Psychiatry*, 12(1), 36-42. <https://doi.org/10.1007/s00787-003-0307-5>
- Beelmann, A., Stemmler, M., Lösel, F., & Jaurisch, S. (2007). Zur Entwicklung externalisierender Verhaltensprobleme im Übergang vom Vor- zum Grundschulalter. *Kindheit und Entwicklung*, 16(4), 229-239. <https://doi.org/10.1026/0942-5403.16.4.229>
- Behr, J., Leidig, T., Krull, J., Spilles, M., & Hennemann, T. (2020). Multiplikatorenkonzepte zur Professionalisierung von Lehrkräften – ein systematisches Review empirischer Studien. *Lehrerbildung auf dem Prüfstand*, 13, S. 151-176.
- Bellmer, S., Scheidtweiler, B., & Borries, N. (2020). Starke Schulen stärken Kinder - der RTI-Ansatz in der Praxis. Inklusionsprojekte im Schulamt für den Märkischen Kreis. *Zeitschrift für Heilpädagogik*, 71(7), 339-353.
- Berger, E. (2019). Multi-tiered Approaches to Trauma-Informed Care in Schools: A Systematic Review. *School Mental Health*, 11(4), 650-664. <https://doi.org/10.1007/s12310-019-09326-0>
- Berkeley, S., Bender, W. N., Gregg Peaster, L., & Saunders, L. (2009). Implementation of response to intervention: a snapshot of progress. *J Learn Disabil*, 42(1), 85-95. <https://doi.org/10.1177/0022219408326214>
- Besa, K., Biehl, A., Möllmann, B., & Rothland, M. (2023). Inklusion ohne Kooperation? Eine Untersuchung zu Kooperationserfahrungen und Einstellungen zur (multiprofessionellen) Kooperation von Lehramtsstudierenden unterschiedlicher Schulformen. In K. t. Poel, P. Gollub, C. Siedenbiedel, S. Greiten, & M. Veber (Hrsg.), *Heterogenität und Inklusion in den Schulpraktischen Studien* (Bd. 8, S. 77-96). Waxmann.
- Bettters-Bubon, J., Pianta, R., Sweeney, D., & Goodman-Scott, E. (2022). Antiracism Starts With Us: School Counselor Critical Reflection Within an Multitiered Systems of Support Framework. *Professional School Counseling*, 26(1a). <https://doi.org/10.1177/2156759x221086747>

- Bilz, L., & Fischer, S. M. (2020). Interventionsstrategien und Interventionserfolg von Lehrkräften bei Cybermobbing und traditionellem Mobbing aus Schülersicht. *Kindheit und Entwicklung, 29*(2), 84-91. <https://doi.org/10.1026/0942-5403/a000306>
- Bjorkqvist, K. (2018). Gender differences in aggression. *Curr Opin Psychol, 19*, 39-42. <https://doi.org/10.1016/j.copsyc.2017.03.030>
- Blair, K. S. C., Park, E.-Y., & Kim, W.-H. (2021). A Meta-Analysis of Tier 2 Interventions Implemented within School-Wide Positive Behavioral Interventions and Supports. *Psychology in the Schools, 58*(1), 141-161.
- Blumenthal, Y., & Voß, S. (2016). Effekte des Response to Intervention-Ansatzes auf die soziale und emotionale Situation von Grundschulern der vierten Jahrgangstufe. *Empirische Pädagogik, 30*, 81-97.
- Bohanon, H. S., Wu, M.-J., Kushki, A., & LeVesseur, C. (2023). A Preliminary Study Connecting School Improvement and MTSS With Student Outcomes. *Australasian Journal of Special and Inclusive Education, 1-15*. <https://doi.org/10.1017/jsi.2023.15>
- Bolton, T. (2010). Review of Bullying prevention and intervention: Realistic strategies for schools. *Educational Psychology in Practice, 26*(2), 213-214.
- Borckardt, J. J., Nash, M. R., Murphy, M. D., Moore, M., Shaw, D., & O'Neil, P. (2008). Clinical practice as natural laboratory for psychotherapy research: a guide to case-based time-series analysis. *Am Psychol, 63*(2), 77-95. <https://doi.org/10.1037/0003-066X.63.2.77>
- Borgen, N. T., Kirkeboen, L. J., Ogden, T., Raaum, O., & Sorlie, M. A. (2020). Impacts of school-wide positive behaviour support: Results from National Longitudinal Register Data. *Int J Psychol, 55 Suppl 1*, 4-15. <https://doi.org/10.1002/ijop.12575>
- Bradley, R., Danielson, L., & Doolittle, J. (2005). Response to intervention. *J Learn Disabil, 38*(6), 485-486. <https://doi.org/10.1177/00222194050380060201>
- Bradshaw, C. P., & Pas, E. T. (2011). A Statewide Scale Up of Positive Behavioral Interventions and Supports: A Description of the Development of Systems of Support and Analysis of Adoption and Implementation. *School Psychology Review, 40*(4), 530-548.
- Bradshaw, C. P., Pas, E. T., Goldweber, A., Rosenberg, M. S., & Leaf, P. J. (2012). Integrating school-wide positive behavioral interventions and supports with tier 2 coaching to student support teams: The PBISplus model. *Advances in School Mental Health Promotion, 5*(3), 177-193.
- Bradshaw, C. P., Schaeffer, C. M., Petras, H., & Ialongo, N. (2010). Predicting negative life outcomes from early aggressive-disruptive behavior trajectories: gender differences in maladaptation across life domains. *J Youth Adolesc, 39*(8), 953-966. <https://doi.org/10.1007/s10964-009-9442-8>

- Braun, G., Kumm, S., Brown, C., Walte, S., Hughes, M. T., & Maggin, D. M. (2018). Living in Tier 2: educators' perceptions of MTSS in urban schools. *International Journal of Inclusive Education*, 24(10), 1114-1128. <https://doi.org/10.1080/13603116.2018.1511758>
- Breeman, L. D., van Lier, P. A. C., Wubbels, T., Verhulst, F. C., van der Ende, J., Maras, A., Struiksma, A. J. C., Hopman, J. A. B., & Tick, N. T. (2015). Effects of the Good Behavior Game on the Behavioral, Emotional, and Social Problems of Children With Psychiatric Disorders in Special Education Settings. *Journal of Positive Behavior Interventions*, 18(3), 156-167. <https://doi.org/10.1177/1098300715593466>
- Breitenstein, S. M., Hill, C., & Gross, D. (2009). Understanding disruptive behavior problems in preschool children. *J Pediatr Nurs*, 24(1), 3-12. <https://doi.org/10.1016/j.pedn.2007.10.007>
- Briesch, A. M., Chafouleas, S. M., Nissen, K., & Long, S. (2019). A Review of State-Level Procedural Guidance for Implementing Multitiered Systems of Support for Behavior (MTSS-B). *Journal of Positive Behavior Interventions*, 22(3), 131-144. <https://doi.org/10.1177/1098300719884707>
- Brodbeck, J., Botschi, S. I. R., Vetsch, N., Berger, T., Schmidt, S. J., & Marmet, S. (2022). Investigating emotion regulation and social information processing as mechanisms linking adverse childhood experiences with psychosocial functioning in young swiss adults: the FACE epidemiological accelerated cohort study. *BMC Psychol*, 10(1), 99. <https://doi.org/10.1186/s40359-022-00798-5>
- Brown-Chidsey, R. (2016). *Practical handbook of multi-tiered systems of support: building academic and behavioral success in schools*. The Guilford Press.
- Bruhn, A. L., Lane, K. L., & Hirsch, S. E. (2013). A Review of Tier 2 Interventions Conducted Within Multitiered Models of Behavioral Prevention. *Journal of Emotional and Behavioral Disorders*, 22(3), 171-189. <https://doi.org/10.1177/1063426613476092>
- Brundrett, M., Beauchamp, G., Latham, P., Mistry, M., Murray, M., Taylor, B., & Wood, P. (2022). *Contemporary Issues in Primary Education*. <https://doi.org/10.4324/9781003316862>
- Buchanan-Pascall, S., Gray, K. M., Gordon, M., & Melvin, G. A. (2018). Systematic Review and Meta-analysis of Parent Group Interventions for Primary School Children Aged 4-12 Years with Externalizing and/or Internalizing Problems. *Child Psychiatry Hum Dev*, 49(2), 244-267. <https://doi.org/10.1007/s10578-017-0745-9>
- Bundschuh, K. (2023). Heilpädagogische Diagnostik als Förderdiagnostik – Herausforderungen und Leistung. In K. Bundschuh (Hrsg.), *Heilpädagogische Psychologie* (5 Aufl.). Ernst Reinhardt.

- Burns, M. K., Jimerson, S. R., & VanDerHeyden, A. M. (2016). *Handbook of Response to Intervention: The Science and Practice of Multi-Tiered Systems of Support* (2 Aufl.). Springer US. <https://doi.org/10.1007/978-1-4899-7568-3>
- Calvete, E., & Orue, I. (2012). The role of emotion regulation in the predictive association between social information processing and aggressive behavior in adolescents. *International Journal of Behavioral Development, 36*(5), 338-347. <https://doi.org/10.1177/0165025412444079>
- Campbell, J. C., Harris, M. J., & Lee, R. K. (1995). Violence research: an overview. *Sch Inq Nurs Pract, 9*(2), 105-126.
- Campbell, J. M. (2004). Statistical comparison of four effect sizes for single-subject designs. *Behav Modif, 28*(2), 234-246. <https://doi.org/10.1177/0145445503259264>
- Casale, G., Volpe, R. J., Daniels, B., Hennemann, T., Briesch, A. M., & Grosche, M. (2018). Measurement Invariance of a Universal Behavioral Screener Across Samples From the USA and Germany. *European Journal of Psychological Assessment, 34*(2), 87-100. <https://doi.org/10.1027/1015-5759/a000447>
- Casale, G., Volpe, R. J., Hennemann, T., Briesch, A. M., Daniels, B., & Grosche, M. (2019). Konstruktvalidität eines universellen Screenings zur unterrichtsnahen und ökonomischen Diagnostik herausfordernden Verhaltens von Schüler_innen – eine Multitrait-Multimethod-Analyse. *Zeitschrift für Pädagogische Psychologie, 33*(1), 17-31. <https://doi.org/10.1024/1010-0652/a000232>
- Castillo, J. M., Scheel, N. L., Wolgemuth, J. R., Latimer, J. D., & Green, S. M. (2022). A scoping review of the literature on professional learning for MTSS. *J Sch Psychol, 92*, 166-187. <https://doi.org/10.1016/j.jsp.2022.03.010>
- Cattaneo, M. D., & Titiunik, R. (2022). Regression Discontinuity Designs. *Annual Review of Economics, 14*(1), 821-851. <https://doi.org/10.1146/annurev-economics-051520-021409>
- Chalfant, J. C., VanDusen Pysh, M., & Moultrie, R. (1979). Teacher Assistance Teams: A Model for Within-Building Problem Solving. *Learning Disability Quarterly, 2*, 85 - 96.
- Chang, T. F., & Qin, D. B. (2018). Maternal Monitoring Knowledge Change and Adolescent Externalizing Behaviors in Low-Income African American and Latino Families. *Psychol Rep, 121*(4), 635-668. <https://doi.org/10.1177/0033294117738022>
- Charlton, C. T., Moulton, S., Sabey, C. V., & West, R. (2020). A Systematic Review of the Effects of Schoolwide Intervention Programs on Student and Teacher Perceptions of School Climate. *Journal of Positive Behavior Interventions, 23*(3), 185-200. <https://doi.org/10.1177/1098300720940168>

- Chen, M., Hyppa-Martin, J. K., Reichle, J. E., & Symons, F. J. (2016). Comparing Single Case Design Overlap-Based Effect Size Metrics From Studies Examining Speech Generating Device Interventions. *Am J Intellect Dev Disabil*, 121(3), 169-193. <https://doi.org/10.1352/1944-7558-121.3.169>
- Cheney, D., & Jewell, K. (2012). Positive Behavior Supports and Students with Emotional and Behavioral Disorders. In *Behavioral Disorders: Practice Concerns and Students with EBD* (S. 83-106). Advances in Special Education (MS). [https://doi.org/10.1108/s0270-4013\(2012\)0000023008](https://doi.org/10.1108/s0270-4013(2012)0000023008)
- Cicchetti, D., & Lynch, M. (1993). Toward an ecological/transactional model of community violence and child maltreatment: consequences for children's development. *Psychiatry*, 56(1), 96-118. <https://doi.org/10.1080/00332747.1993.11024624>
- Clark, A. G., & Dockweiler, K. A. (2020). *Multi-tiered systems of support in elementary schools: the definitive guide to effective implementation and quality control*. Routledge.
- Clifford, M. E., Nguyen, A. J., & Bradshaw, C. P. (2020). Both/and: Tier 2 Interventions with Transdiagnostic Utility in Addressing Emotional and Behavioral Disorders in Youth. *Journal of Applied School Psychology*, 36(2), 173-197. <https://doi.org/10.1080/15377903.2020.1714859>
- Clonan, S. M., McDougal, J. L., Clark, K., & Davison, S. (2006). Use of office discipline referrals in school-wide decision making: A practical example. *Psychology in the Schools*, 44(1), 19-27. <https://doi.org/10.1002/pits.20202>
- Cohen, J. (2013). *Statistical Power Analysis for the Behavioral Sciences*. Routledge. <https://doi.org/10.4324/9780203771587>
- Cohn, J. R., & Cicero, D. C. (2024). Cognitive and social-cognitive factors as mediators of aggression using structural equation modeling. *Psychology, Crime & Law*, 1-24. <https://doi.org/10.1080/1068316x.2024.2316858>
- Congress, U. S. (2004). *Individuals with Disabilities Education Improvement Act of 2004*. U.S. Government Publishing Office.
- Congress, U. S. (2015). *Every Student Succeeds Act of 2015*. U.S. Government Publishing Office.
- Couban, R. (2016). Covidence and Rayyan. *Journal of the Canadian Health Libraries Association / Journal de l'Association des bibliothèques de la santé du Canada*, 37(3). <https://doi.org/10.5596/c16-025>
- 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), 74-101. <https://doi.org/10.1037/0033-2909.115.1.74>

- Daniels, B., Volpe, R. J., Briesch, A. M., & Fabiano, G. A. (2014). Development of a problem-focused behavioral screener linked to evidence-based intervention. *Sch Psychol Q*, 29(4), 438-451. <https://doi.org/10.1037/spq0000100>
- de Jong, S. R. C., van den Hoofdakker, B. J., van der Veen-Mulders, L., Veenman, B., Twisk, J. W. R., Oosterlaan, J., & Luman, M. (2023). The efficacy of a self-help parenting program for parents of children with externalizing behavior: a randomized controlled trial. *Eur Child Adolesc Psychiatry*, 32(10), 2031-2042. <https://doi.org/10.1007/s00787-022-02028-0>
- Debnam, K. J., Pas, E. T., & Bradshaw, C. P. (2011). Factors Influencing Staff Perceptions of Administrator Support for Tier 2 and 3 Interventions. *Journal of Emotional and Behavioral Disorders*, 21(2), 116-126. <https://doi.org/10.1177/1063426611410571>
- Deighton, J., Humphrey, N., Belsky, J., Boehnke, J., Vostanis, P., & Patalay, P. (2018). Longitudinal pathways between mental health difficulties and academic performance during middle childhood and early adolescence. *Br J Dev Psychol*, 36(1), 110-126. <https://doi.org/10.1111/bjdp.12218>
- Delbecq, A., Ven, A., & Gustafson, D. (1986). Group Techniques for Program Planning: A Guide to Nominal Group and Delphi Processes. *Glenview, Illinois: Scott Forman and Co.*
- Demirtaş-Madran, H. A. (2023). Violent Video Games and Aggression. In C. R. Martin, V. R. Preedy, & V. B. Patel (Hrsg.), *Handbook of Anger, Aggression, and Violence* (S. 149-170). Springer International Publishing. https://doi.org/10.1007/978-3-031-31547-3_21
- Demmer, C., & Hopmann, B. (2020). Multiprofessionelle Kooperation in inklusiven Ganztagschulen. In P. Bollweg, J. Buchna, T. Coelen, & H.-U. Otto (Hrsg.), *Handbuch Ganztagsbildung* (S. 1467-1477). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-23230-6_108
- Denham, S. A., & Bassett, H. H. (2020). "You Hit Me! That's Not Nice and it Makes Me Sad!!!": Relations of Young Children's Social Information Processing and Early School Success. *Early Child Dev Care*, 190(6), 791-805. <https://doi.org/10.1080/03004430.2018.1491562>
- Deno, E. (2018). Special Education as Developmental Capital. *Exceptional Children*, 37(3), 229-237. <https://doi.org/10.1177/001440297003700306>
- Deno, S. L., & Mirkin, P. K. (1977). Data-based program modification: A manual.
- Diehm, I. (2020). Differenz – die pädagogische Herausforderung in der Schule für alle Kinder. In *Diversität und soziale Ungleichheit* (S. 9-19). https://doi.org/10.1007/978-3-658-27529-7_2

- Dobbins, M., & Bieber, T. (2014). Bildungspolitik in den USA. In C. Lammert, M. Siewert, & B. Vormann (Hrsg.), *Handbuch Politik USA* (S. 1-18). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-04125-0_24-1
- Dobrosavljevic, M., Kuja-Halkola, R., Li, L., Chang, Z., Larsson, H., & Du Rietz, E. (2023). Attention-deficit/hyperactivity disorder symptoms and subsequent cardiometabolic disorders in adults: investigating underlying mechanisms using a longitudinal twin study. *BMC Med*, *21*(1), 452. <https://doi.org/10.1186/s12916-023-03174-1>
- Docherty, M., Decrop, R., McManamon, B., Boxer, P., Dubow, E. F., & Huesmann, L. R. (2023). Exposure to violence predicts callous-unemotional traits and aggression in adolescence in the context of persistent ethnic-political conflict and violence. *Aggress Behav*, *49*(6), 655-668. <https://doi.org/10.1002/ab.22103>
- Donaldson, J. M., Fisher, A. B., & Kahng, S. (2017). Effects of the Good Behavior Game on individual student behavior. *Behavior Analysis: Research and Practice*, *17*(3), 207-216. <https://doi.org/10.1037/bar0000016>
- Döpfner, M. (2022). Kategoriale versus dimensionale Klassifikation psychischer Störungen in der Kinder- und Jugendpsychiatrie und -psychotherapie. In J. Fegert, F. Resch, P. Plener, M. Kaess, M. Döpfner, K. Konrad, & T. Legenbauer (Hrsg.), *Psychiatrie und Psychotherapie des Kindes- und Jugendalters* (S. 1-11). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-49289-5_64-2
- Döpfner, M., & Banaschewski, T. (2022). Aufmerksamkeitsdefizit-/Hyperaktivitätsstörung (ADHS) im Kindes- und Jugendalter. In J. Fegert, F. Resch, P. Plener, M. Kaess, M. Döpfner, K. Konrad, & T. Legenbauer (Hrsg.), *Psychiatrie und Psychotherapie des Kindes- und Jugendalters* (S. 1-23). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-49289-5_124-1
- Dowdy, A., Peltier, C., Tincani, M., Schneider, W. J., Hantula, D. A., & Travers, J. C. (2021). Meta-analyses and effect sizes in applied behavior analysis: A review and discussion. *J Appl Behav Anal*, *54*(4), 1317-1340. <https://doi.org/10.1002/jaba.862>
- Dresing, T., & Pehl, T. (2020). Transkription. In *Handbuch Qualitative Forschung in der Psychologie* (S. 1-20). https://doi.org/10.1007/978-3-658-18387-5_56-2
- Drummond, T. (1994). *The Student Risk Screening Scale (SRSS)*. Josephine County Mental Health Program.
- Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., & Schellinger, K. B. (2011). The impact of enhancing students' social and emotional learning: a meta-analysis of school-based universal interventions. *Child Dev*, *82*(1), 405-432. <https://doi.org/10.1111/j.1467-8624.2010.01564.x>

- Edirmanasinghe, N., Goodman-Scott, E., Smith-Durkin, S., & Tarver, S. Z. (2022). Supporting All Students: Multitiered Systems of Support from an Antiracist and Critical Race Theory Lens. *Professional School Counseling, 26*(1). <https://doi.org/10.1177/2156759x221109154>
- Eggers, A. C., Freier, R., Grembi, V., & Nannicini, T. (2017). Regression Discontinuity Designs Based on Population Thresholds: Pitfalls and Solutions. *American Journal of Political Science, 62*(1), 210-229. <https://doi.org/10.1111/ajps.12332>
- Eklund, G., Sundqvist, C., Lindell, M., & Toppinen, H. (2020). A study of Finnish primary school teachers' experiences of their role and competences by implementing the three-tiered support. *European Journal of Special Needs Education, 36*(5), 729-742. <https://doi.org/10.1080/08856257.2020.1790885>
- Epstein, L. H., & Dallery, J. (2022). The Family of Single-Case Experimental Designs. *Harv Data Sci Rev, 4*(S13). <https://doi.org/10.1162/99608f92.ff9300a8>
- Erchul, W. P., & Martens, B. K. (2010). Problem Solving and Response to Intervention. In W. P. Erchul & B. K. Martens (Hrsg.), *School Consultation* (S. 17-30). Springer New York. https://doi.org/10.1007/978-1-4419-5747-4_2
- Ernst, S. (2023). Zwischen Dauerreflexion und Wegschauen? Diversität und Vielfalt in der Schule. In *Diversitäts- und Organisationsforschung* (S. 411-432). <https://doi.org/10.5771/9783748934684-411>
- Erskine, H. E., Norman, R. E., Ferrari, A. J., Chan, G. C., Copeland, W. E., Whiteford, H. A., & Scott, J. G. (2016). Long-Term Outcomes of Attention-Deficit/Hyperactivity Disorder and Conduct Disorder: A Systematic Review and Meta-Analysis. *J Am Acad Child Adolesc Psychiatry, 55*(10), 841-850. <https://doi.org/10.1016/j.jaac.2016.06.016>
- Ervin, R. A., Schaughency, E., Goodman, S. D., McGlinchey, M. T., & Matthews, A. (2006). Merging research and practice agendas to address reading and behavior school-wide. *School Psychology Review, 35*(2), 198-223.
- Ervin, R. A., Schaughency, E., Matthews, A., Goodman, S. D., & McGlinchey, M. T. (2006). Primary and secondary prevention of behavior difficulties: Developing a data-informed problem-solving model to guide decision making at a school-wide level. *Psychology in the Schools, 44*(1), 7-18. <https://doi.org/10.1002/pits.20201>
- Estrapala, S., Rila, A., & Bruhn, A. L. (2020). A Systematic Review of Tier 1 PBIS Implementation in High Schools. *Journal of Positive Behavior Interventions, 23*(4), 288-302. <https://doi.org/10.1177/1098300720929684>

- Evans, A. L., Bulla, A. J., & Kieta, A. R. (2021). The Precision Teaching System: A Synthesized Definition, Concept Analysis, and Process. *Behav Anal Pract*, 14(3), 559-576. <https://doi.org/10.1007/s40617-020-00502-2>
- Evers, W. J. G., Tomic, W., & Brouwers, A. (2016). Burnout among Teachers. *School Psychology International*, 25(2), 131-148. <https://doi.org/10.1177/0143034304043670>
- Fabiano, G. A., & Pyle, K. (2018). Best Practices in School Mental Health for Attention-Deficit/Hyperactivity Disorder: A Framework for Intervention. *School Mental Health*, 11(1), 72-91. <https://doi.org/10.1007/s12310-018-9267-2>
- Fallon, L. M., Cathcart, S. C., Johnson, A. H., Minami, T., O’Keeffe, B. V., DeFouw, E. R., & Sugai, G. (2022). A Teacher Self-Assessment of Culturally Relevant Practice to Inform Educator Professional Development Decisions in MTSS Contexts. *Assessment for Effective Intervention*, 48(2), 100-112. <https://doi.org/10.1177/15345084221111338>
- Fallon, L. M., Veiga, M., & Sugai, G. (2021). Strengthening MTSS for Behavior (MTSS-B) to Promote Racial Equity. *School Psychology Review*, 52(5), 518-533. <https://doi.org/10.1080/2372966x.2021.1972333>
- Farahani, H., Nepton, A., & Williams, M. T. (2023). Racial Microaggressions. In C. R. Martin, V. R. Preedy, & V. B. Patel (Hrsg.), *Handbook of Anger, Aggression, and Violence* (S. 245-270). Springer International Publishing. https://doi.org/10.1007/978-3-031-31547-3_10
- Faraone, S. V., Bellgrove, M. A., Brikell, I., Cortese, S., Hartman, C. A., Hollis, C., Newcorn, J. H., Philipsen, A., Polanczyk, G. V., Rubia, K., Sibley, M. H., & Buitelaar, J. K. (2024). Attention-deficit/hyperactivity disorder. *Nat Rev Dis Primers*, 10(1), 11. <https://doi.org/10.1038/s41572-024-00495-0>
- Farmer, T. W., Bierman, K. L., Hall, C. M., Brooks, D. S., & Lee, D. L. (2020). Tiered Systems of Adaptive Supports and the Individualization of Intervention: Merging Developmental Cascades and Correlated Constraints Perspectives. *Journal of Emotional and Behavioral Disorders*, 29(1), 3-13. <https://doi.org/10.1177/1063426620957651>
- Farmer, T. W., & Xie, H. (2007). Aggression and school social dynamics: The good, the bad, and the ordinary. *Journal of School Psychology*, 45(5), 461-478. <https://doi.org/10.1016/j.jsp.2007.06.008>
- Fenzl, T., & Mayring, P. (2017). QCAmapp: eine interaktive Webapplikation für Qualitative Inhaltsanalyse. *Zeitschrift für Soziologie der Erziehung und Sozialisation*, 37, 333-339.
- Feuerborn, L. L., & Tyre, A. D. (2012). Establishing Positive Discipline Policies in an Urban Elementary School. *Contemporary School Psychology*, 16, 47-58.

- Feuerborn, L. L., Wallace, C., & Tyre, A. D. (2016). A Qualitative Analysis of Middle and High School Teacher Perceptions of Schoolwide Positive Behavior Supports. *Journal of Positive Behavior Interventions, 18*(4), 219-229. <https://doi.org/10.1177/1098300716632591>
- Flanagan, J. C. (1954). The critical incident technique. *Psychol Bull, 51*(4), 327-358. <https://doi.org/10.1037/h0061470>
- Flannery, K. B., Sugai, G., & Anderson, C. M. (2009). School-Wide Positive Behavior Support in High School. *Journal of Positive Behavior Interventions, 11*(3), 177-185. <https://doi.org/10.1177/1098300708316257>
- Flower, A., McKenna, J. W., Bunuan, R. L., Muething, C. S., & Vega, R. (2014). Effects of the Good Behavior Game on Challenging Behaviors in School Settings. *Review of Educational Research, 84*(4), 546-571. <https://doi.org/10.3102/0034654314536781>
- Fossum, S., Handegard, B. H., Martinussen, M., & Morch, W. T. (2008). Psychosocial interventions for disruptive and aggressive behaviour in children and adolescents: a meta-analysis. *Eur Child Adolesc Psychiatry, 17*(7), 438-451. <https://doi.org/10.1007/s00787-008-0686-8>
- Fox, R. A., Leif, E. S., Moore, D. W., Furlonger, B., Anderson, A., & Sharma, U. (2021). A Systematic Review of the Facilitators and Barriers to the Sustained Implementation of School-Wide Positive Behavioral Interventions and Supports. *EDUCATION AND TREATMENT OF CHILDREN, 45*(1), 105-126. <https://doi.org/10.1007/s43494-021-00056-0>
- Franzen, K., & Kamps, D. (2008). The Utilization and Effects of Positive Behavior Support Strategies on an Urban School Playground. *Journal of Positive Behavior Interventions, 10*(3), 150-161. <https://doi.org/10.1177/1098300708316260>
- Freeman, J., Wilkinson, S., Kowitt, J., Kittelman, A., & Brigid Flannery, K. (2019). Research-supported practices for improving attendance in high schools: a review of the literature. *Educational Research and Evaluation, 24*(8), 481-503. <https://doi.org/10.1080/13803611.2019.1602546>
- Friedrichs, W., & Sanders, O. (2002). *Bildung / Transformation*. <https://doi.org/10.14361/9783839400944>
- Fry, D., Fang, X., Elliott, S., Casey, T., Zheng, X., Li, J., Florian, L., & McCluskey, G. (2018). The relationships between violence in childhood and educational outcomes: A global systematic review and meta-analysis. *Child Abuse Negl, 75*, 6-28. <https://doi.org/10.1016/j.chiabu.2017.06.021>
- Fuchs, L. S., & Fuchs, D. (1986). Effects of systematic formative evaluation: a meta-analysis. *Except Child, 53*(3), 199-208. <https://doi.org/10.1177/001440298605300301>
- Fuchs, L. S., & Fuchs, D. (2007). A Model for Implementing Responsiveness to Intervention. *TEACHING Exceptional Children, 39*(5), 14-20. <https://doi.org/10.1177/004005990703900503>

- Fullan, M. (2015). *The New Meaning of Educational Change*. Teachers College Press.
- Gach, E. J., Ip, K. I., Sameroff, A. J., & Olson, S. L. (2018). Early cumulative risk predicts externalizing behavior at age 10: The mediating role of adverse parenting. *J Fam Psychol*, 32(1), 92-102. <https://doi.org/10.1037/fam0000360>
- Gage, N. A., Beahm, L., Kaplan, R., MacSuga-Gage, A. S., & Lee, A. (2020). Using Positive Behavioral Interventions and Supports to Reduce School Suspensions. *Beyond Behavior*, 29(3), 132-140. <https://doi.org/10.1177/1074295620950611>
- Gage, N. A., Lee, A., Grasley-Boy, N., & Peshak George, H. (2018). The Impact of School-Wide Positive Behavior Interventions and Supports on School Suspensions: A Statewide Quasi-Experimental Analysis. *Journal of Positive Behavior Interventions*, 20(4), 217-226. <https://doi.org/10.1177/1098300718768204>
- Gage, N. A., Sugai, G., Lewis, T. J., & Brzozowy, S. (2013). Academic Achievement and School-Wide Positive Behavior Supports. *Journal of Disability Policy Studies*, 25(4), 199-209. <https://doi.org/10.1177/1044207313505647>
- Gage, N. A., Whitford, D. K., & Katsiyannis, A. (2018). A Review of Schoolwide Positive Behavior Interventions and Supports as a Framework for Reducing Disciplinary Exclusions. *The Journal of Special Education*, 52(3), 142-151. <https://doi.org/10.1177/0022466918767847>
- Gebhardt, M., Scheer, D., & Michael, S. (2022). Sonderpädagogische Diagnostik – Ein Einführung. In M. Gebhardt, Scheer, D. & Schurig, M. (Hrsg.), *Handbuch der sonderpädagogischen Diagnostik. Grundlagen und Konzepte der Statusdiagnostik, Prozessdiagnostik und Förderplanung*. Universitätsbibliothek d. Universität Regensburg. <https://doi.org/10.5283/epub.53149>
- George, H. P., Cox, K. E., Minch, D., & Sandomierski, T. (2018). District Practices Associated With Successful SWPBIS Implementation. *Behavioral Disorders*, 43(3), 393-406. <https://doi.org/10.1177/0198742917753612>
- Gerzel-Short, L. (2018). 'We conquered this together': Tier 2 collaboration with families. *The School Community Journal*, 28(2), 85-112.
- 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. <https://doi.org/10.17886/RKI-GBE-2018-078>
- Good, R. H., & Kaminski, R. A. (2002). *Dynamic Indicators of Basic Early Literacy Skills: administration and scoring guide*. University of Oregon.

- Goodman, R. (1997). The Strengths and Difficulties Questionnaire: a research note. *J Child Psychol Psychiatry*, 38(5), 581-586. <https://doi.org/10.1111/j.1469-7610.1997.tb01545.x>
- Goodman-Scott, E. (2013). Maximizing School Counselors' Efforts by Implementing School-Wide Positive Behavioral Interventions and Supports: A Case Study from the Field. *Professional School Counseling*, 17(1), 111-119.
- Goodman-Scott, E., Betters-Bubon, J., Donohue, P., & Olsen, J. (2023). *The School Counselor's Guide to Multi-Tiered Systems of Support*. <https://doi.org/10.4324/9781003306412>
- Götz, J., Hauenschild, K., Greve, W., & Hellmers, S. (2015). Einstellungen von Lehrerinnen und Lehrern zur inklusiven Grundschule. In D. Blömer, M. Lichtblau, A.-K. Jüttner, K. Koch, M. Krüger, & R. Werning (Hrsg.), *Perspektiven auf inklusive Bildung* (S. 34-39). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-06955-1_4
- Grasley-Boy, N. M., Reichow, B., van Dijk, W., & Gage, N. (2020). A Systematic Review of Tier 1 PBIS Implementation in Alternative Education Settings. *Behavioral Disorders*, 46(4), 199-213. <https://doi.org/10.1177/0198742920915648>
- Grawe, K. (2000). *Psychologische Therapie* (2 Aufl.). Hogrefe.
- Grayson, J. L., & Alvarez, H. K. (2008). School climate factors relating to teacher burnout: A mediator model. *Teaching and Teacher Education*, 24(5), 1349-1363. <https://doi.org/10.1016/j.tate.2007.06.005>
- Greif, S. (2008). *Coaching und ergebnisorientierte Selbstreflexion*. Hogrefe.
- Gresham, F., Watson, T. S., & Skinner, C. H. (2001). Functional behavioral assessment: Principles, procedures, and future directions. *School Psychology Review*, 30(2), 156-172.
- Gresham, F. M. (2016). Response to Intervention and Emotional and Behavioral Disorders. *Assessment for Effective Intervention*, 32(4), 214-222. <https://doi.org/10.1177/15345084070320040301>
- Greulich, L., Otaiba, S. A., Schatschneider, C., Wanzek, J., Ortiz, M., & Wagner, R. K. (2014). Understanding inadequate response to first-grade multi-tier intervention: Nomothetic and ideographic perspectives. *Learning Disability Quarterly*, 37(4), 204-217.
- Grimshaw, J., Eccles, M., Thomas, R., MacLennan, G., Ramsay, C., Fraser, C., & Vale, L. (2006). Toward evidence-based quality improvement. Evidence (and its limitations) of the effectiveness of guideline dissemination and implementation strategies 1966-1998. *J Gen Intern Med*, 21(Suppl 2), S14-20. <https://doi.org/10.1111/j.1525-1497.2006.00357.x>

- Grosche, M., & Volpe, R. J. (2013). Response-to-intervention (RTI) as a model to facilitate inclusion for students with learning and behaviour problems. *European Journal of Special Needs Education, 28*(3), 254-269. <https://doi.org/10.1080/08856257.2013.768452>
- Guest, E. M. (2011). *The impact of positive behavioral interventions and supports in secondary school settings* (G. Tindal, Hrsg.). University of Oregon.
- Guevremont, D. C., & Dumas, M. C. (1994). Peer Relationship Problems and Disruptive Behavior Disorders. *Journal of Emotional and Behavioral Disorders, 2*(3), 164-172.
- Hagen, T., Nitz, J., Brack, F., Hövel, D. C., & Hennemann, T. (2023). Effekte des Good Behavior Game bei Grundschüler_innen mit externalisierenden Verhaltensproblemen. *Lernen und Lernstörungen, 12*(4), 215-228. <https://doi.org/10.1024/2235-0977/a000410>
- Halbheer, U., & Kunz, A. (2011). Kooperation. In U. Halbheer & A. Kunz (Hrsg.), *Kooperation von Lehrpersonen an Gymnasien* (S. 91-108). VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-92720-6_5
- Hanisch, C., Casale, G., Volpe, R. J., Briesch, A. M., Richard, S., Meyer, H., Hövel, D., Hagen, T., Krull, J., & Hennemann, T. (2019). Gestufte Förderung in der Grundschule. *Prävention und Gesundheitsförderung, 14*(3), 237-241. <https://doi.org/10.1007/s11553-018-0700-z>
- Hanisch, C., Eichelberger, I., Richard, S., & Doepfner, M. (2020). Effects of a modular teacher coaching program on child attention problems and disruptive behavior and on teachers' self-efficacy and stress. *School Psychology International, 41*(6), 543-568. <https://doi.org/10.1177/0143034320958743>
- Hanisch, C., Leidig, T., Hennemann, T., & Vögele, U. (2023). Psychische Auffälligkeiten von Kindern und Jugendlichen an Förderschulen mit dem Förderschwerpunkt Emotionale und soziale Entwicklung. *Empirische Sonderpädagogik, 1*, 21-37.
- Hanisch, C., Richard, S., Eichelberger, I., Greimel, L., & Döpfner, M. (2018). *Schulbasiertes Coaching bei Kindern mit expansivem Problemverhalten (SCEP)*. <https://doi.org/10.1026/02813-000>
- Harlacher, J. E., Sakelaris, T. L., & Kattelman, N. M. (2014). Multi-Tiered System of Support. In *Practitioner's Guide to Curriculum-Based Evaluation in Reading* (S. 23-45). Springer New York. https://doi.org/10.1007/978-1-4614-9360-0_3
- Harrison, P. L., & Thomas, A. M. (2014). *Data-based and collaborative decision making*. National Association of School Psychologists.
- Harrison, S. E., Riley-Tillman, T. C., & Chafouleas, S. M. (2014). Direct Behavior Rating. *Canadian Journal of School Psychology, 29*(1), 3-20. <https://doi.org/10.1177/0829573513515424>

- Hastings, R. P., & Brown, T. (2002). Behavioural knowledge, causal beliefs and self-efficacy as predictors of special educators' emotional reactions to challenging behaviours. *J Intellect Disabil Res*, *46*(Pt 2), 144-150. <https://doi.org/10.1046/j.1365-2788.2002.00378.x>
- Hautmann, C., Stein, P., Eichelberger, I., Hanisch, C., Plück, J., Walter, D., & Döpfner, M. (2010). The Severely Impaired Do Profit Most: Differential Effectiveness of a Parent Management Training for Children with Externalizing Behavior Problems in a Natural Setting. *Journal of Child and Family Studies*, *20*(4), 424-435. <https://doi.org/10.1007/s10826-010-9409-0>
- Hawken, L. S., Vincent, C. G., & Schumann, J. (2008). Response to Intervention for Social Behavior. *Journal of Emotional and Behavioral Disorders*, *16*(4), 213-225. <https://doi.org/10.1177/1063426608316018>
- Hechtman, L., Swanson, J. M., Sibley, M. H., Stehli, A., Owens, E. B., Mitchell, J. T., Arnold, L. E., Molina, B. S., Hinshaw, S. P., Jensen, P. S., Abikoff, H. B., Perez Algorta, G., Howard, A. L., Hoza, B., Etcovitch, J., Houssais, S., Lakes, K. D., Nichols, J. Q., & Group, M. T. A. C. (2016). Functional Adult Outcomes 16 Years After Childhood Diagnosis of Attention-Deficit/Hyperactivity Disorder: MTA Results. *J Am Acad Child Adolesc Psychiatry*, *55*(11), 945-952 e942. <https://doi.org/10.1016/j.jaac.2016.07.774>
- Herman, K. C., Hickmon-Rosa, J. e., & Reinke, W. M. (2017). Empirically Derived Profiles of Teacher Stress, Burnout, Self-Efficacy, and Coping and Associated Student Outcomes. *Journal of Positive Behavior Interventions*, *20*(2), 90-100. <https://doi.org/10.1177/1098300717732066>
- Hierck, T., & Weber, C. (2022). *The road to success with MTSS: a ten-step process for schools*. Solution Tree Press,.
- Hillenbrand, C., & Pütz, K. (2008). *KlasseKinderSpiel. Spielerisch Verhaltensregeln lernen*. Edition Körber.
- Hoglund, W. L., Klinge, K. E., & Hosan, N. E. (2015). Classroom risks and resources: Teacher burnout, classroom quality and children's adjustment in high needs elementary schools. *J Sch Psychol*, *53*(5), 337-357. <https://doi.org/10.1016/j.jsp.2015.06.002>
- Hoover, J. J., Soltero-González, L., Wang, C., & Herron, S. (2019). Sustaining a Multitiered System of Supports for English Learners in Rural Community Elementary Schools. *Rural Special Education Quarterly*, *39*(1), 4-16. <https://doi.org/10.1177/8756870519847466>
- Hopman, J. A. B., van Lier, P. A. C., van der Ende, J., Struiksma, C., Wubbels, T., Verhulst, F. C., Maras, A., Breeman, L. D., & Tick, N. T. (2017). Impact of the Good Behavior Game on special education teachers. *Teachers and Teaching*, *24*(4), 350-368. <https://doi.org/10.1080/13540602.2017.1379389>
- Horner, R. H., Kincaid, D., Sugai, G., Lewis, T., Eber, L., Barrett, S., Dickey, C. R., Richter, M., Sullivan, E., Boezio, C., Algozzine, B., Reynolds, H., & Johnson, N. (2013). Scaling Up School-Wide Positive

- Behavioral Interventions and Supports. *Journal of Positive Behavior Interventions*, 16(4), 197-208.
<https://doi.org/10.1177/1098300713503685>
- Horner, R. H., Todd, A. W., Lewis-Palmer, T., Irvin, L. K., Sugai, G., & Boland, J. B. (2016). The School-Wide Evaluation Tool (SET). *Journal of Positive Behavior Interventions*, 6(1), 3-12.
<https://doi.org/10.1177/10983007040060010201>
- Hornsveld, R. H. J., & Kraaimaat, F. W. (2022). Psychological factors predicting violent prison inmates' anger and aggression. *Journal of Forensic Psychology Research and Practice*, 23(5), 472-488.
<https://doi.org/10.1080/24732850.2022.2048768>
- Hövel, D. C., Hennemann, T., & Rietz, C. (2019). Meta-Analyse programmatischer-präventiver Förderung der emotionalen und sozialen Entwicklung in der Primarstufe. In S. Gingelmaier, W. Bleher, M. Hoanzl, & B. Herz (Hrsg.), *Emotionale und Soziale Entwicklung in der Pädagogik der Erziehungshilfe und bei Verhaltensstörungen* (Bd. 1, S. 38-57). Julius Klinkhardt. <https://doi.org/https://doi.org/10.35468/5750-03>
- Hoyle, C. G., Marshall, K. J., & Yell, M. L. (2011). Positive Behavior Supports: Tier 2 Interventions in Middle Schools. *Preventing School Failure*, 55(3), 164-170.
- Hrastinski, S. (2019). What Do We Mean by Blended Learning? *TechTrends*, 63(5), 564-569.
<https://doi.org/10.1007/s11528-019-00375-5>
- Huber, C., & Rietz, C. (2015). Direct Behavior Rating (DBR) als Methode zur Verhaltensverlaufsdagnostik in der Schule: Ein systematisches Review von Methodenstudien. *Empirische Sonderpädagogik*, 2, 75-98.
- Huber, C. G. M. (2013). Inklusive Schulentwicklung durch response-to- intervention (RTI) – Realisierungsmöglichkeiten des RTI-Konzepts im Förderbereich Lesen. In *Gemeinsam Leben* (Bd. 21, S. 79-90). Beltz Juventa.
- Hubert, C., & Grosche, M. (2012). Das response-to-intervention-Modell als Grundlage für einen inklusiven Paradigmenwechsel in der Sonderpädagogik. *Zeitschrift für Heilpädagogik*.
- Huffman, L. G. (2022). *Contextual and Neurobiological Foundations of Externalizing Behavior in Early Adolescence*. University of Georgia.
- Hurrelmann, K., Ditton, H., & Sliwka, A. (2010). *Transmission 03: Herkunft und Chance- Wege zu mehr Bildungsgerechtigkeit an Deutschlands Schulen* (V. S. Deutschland, Hrsg.). ZBW – Leibniz- Informationszentrum Wirtschaft/Leibniz Information Centre for Economics.
- Husky, M. M., Boyd, A., Bitfoi, A., Carta, M. G., Chan-Chee, C., Goelitz, D., Koc, C., Lesinskiene, S., Mihova, Z., Otten, R., Pez, O., Shojaei, T., & Kovess-Masfety, V. (2018). Self-reported mental health in children

- ages 6-12 years across eight European countries. *Eur Child Adolesc Psychiatry*, 27(6), 785-795.
<https://doi.org/10.1007/s00787-017-1073-0>
- Iznardo, M., Rogers, M. A., Volpe, R. J., Labelle, P. R., & Robaey, P. (2020). The Effectiveness of Daily Behavior Report Cards for Children With ADHD: A Meta-Analysis. *J Atten Disord*, 24(12), 1623-1636.
<https://doi.org/10.1177/1087054717734646>
- James, A. G., Noltemeyer, A., Ritchie, R., Palmer, K., & University, M. (2019). Longitudinal disciplinary and achievement outcomes associated with school-wide PBIS implementation level. *Psychology in the Schools*, 56(9), 1512-1521. <https://doi.org/10.1002/pits.22282>
- Jamshidi, L., Declercq, L., Fernandez-Castilla, B., Ferron, J. M., Moeyaert, M., Beretvas, S. N., & Van den Noortgate, W. (2020). Multilevel meta-analysis of multiple regression coefficients from single-case experimental studies. *Behav Res Methods*, 52(5), 2008-2019. <https://doi.org/10.3758/s13428-020-01380-w>
- Janssens, A., Van Den Noortgate, W., Goossens, L., Verschueren, K., Colpin, H., De Laet, S., Claes, S., & Van Leeuwen, K. (2015). Externalizing Problem Behavior in Adolescence: Dopaminergic Genes in Interaction with Peer Acceptance and Rejection. *J Youth Adolesc*, 44(7), 1441-1456.
<https://doi.org/10.1007/s10964-015-0304-2>
- Jenkins, J. R., & Fuchs, L. S. (2012). Curriculum-Based Measurement
 The Paradigm, History, and Legacy. In C. A. Espin, K. L. McMaster, S. Rose, & M. M. Wayman (Hrsg.), *A Measure of Success* (S. 7-24). University of Minnesota Press.
- Jensen, J., Köller, O., & Zimmermann, F. (2023). Does Rude or Kind Behavior Predict Later Academic Achievement? Evidence From Two Samples of Adolescents. *The Journal of Early Adolescence*, 44(3), 334-364. <https://doi.org/10.1177/02724316231178688>
- Jimerson, S. R., Burns, M. K., & VanDerHeyden, A. M. (2016). From Response to Intervention to Multi-Tiered Systems of Support: Advances in the Science and Practice of Assessment and Intervention. In S. R. Jimerson, M. K. Burns, & A. M. VanDerHeyden (Hrsg.), *Handbook of Response to Intervention* (S. 1-6). Springer US. https://doi.org/10.1007/978-1-4899-7568-3_1
- Jornevald, M., Pettersson-Roll, L., & Hau, H. (2023). The Good Behavior Game for students with special educational needs in mainstream education settings: A scoping review. *Psychology in the Schools*, 61(3), 861-886. <https://doi.org/10.1002/pits.23086>
- Joslyn, P. R., Donaldson, J. M., Austin, J. L., & Vollmer, T. R. (2019). The Good Behavior Game: A brief review. *J Appl Behav Anal*, 52(3), 811-815. <https://doi.org/10.1002/jaba.572>

- Kammerl, R., Dertinger, A., & Kramer, M. (2023). Wie verändern sich Kindheit und Grundschule in einer durch Digitalität geprägten Welt? Digitale Bildung als Herausforderung für pädagogische Akteur:innen. In *Grundschule und Digitalität* (S. 54-67). Grundschulverband.
- Kamphaus, R. W., & Reynolds, C. R. (2015). BASC–3 behavioral and emotional screening system. *Psychological Test and Assessment Modeling*.
- Karatas, S., Eckstein, K., Noack, P., Rubini, M., & Crocetti, E. (2023). Positive and negative intergroup contact in school and out-of-school contexts: A longitudinal approach to spillover effects. *J Res Adolesc*, 33(4), 1335-1349. <https://doi.org/10.1111/jora.12881>
- Karhu, A., Paananen, M., Närhi, V., & Savolainen, H. (2020). Implementation of the inclusive CICO Plus intervention for pupils at risk of severe behaviour problems in SWPBS schools. *European Journal of Special Needs Education*, 36(5), 758-772. <https://doi.org/10.1080/08856257.2020.1809801>
- Kauten, R., & Barry, C. T. (2016). Externalizing Behavior. In *Encyclopedia of Personality and Individual Differences* (S. 1-4). https://doi.org/10.1007/978-3-319-28099-8_894-1
- Keil, V., & Price, J. M. (2006). Externalizing behavior disorders in child welfare settings: Definition, prevalence, and implications for assessment and treatment. *Children and youth services review*, 28(7), 761-779. <https://doi.org/10.1016/j.childyouth.2005.08.006>
- Kellermeyer, L., Harnke, B., & Knight, S. (2018). Covidence and Rayyan. *Journal of the Medical Library Association*, 106(4). <https://doi.org/10.5195/jmla.2018.513>
- Kelm, J. L., McIntosh, K., & Cooley, S. (2014). Effects of Implementing School-Wide Positive Behavioural Interventions and Supports on Problem Behaviour and Academic Achievement in a Canadian Elementary School. *Canadian Journal of School Psychology*, 29(3), 195-212.
- Kincaid, D., Childs, K., Blase, K. A., & Wallace, F. (2016). Identifying Barriers and Facilitators in Implementing Schoolwide Positive Behavior Support. *Journal of Positive Behavior Interventions*, 9(3), 174-184. <https://doi.org/10.1177/10983007070090030501>
- Kincaid, D., Knoster, T., Harrower, J. K., Shannon, P., & Bustamante, S. (2016). Measuring the Impact of Positive Behavior Support. *Journal of Positive Behavior Interventions*, 4(2), 109-117. <https://doi.org/10.1177/109830070200400206>
- Kincaid, D., Peshak George, H., & Childs, K. (2016). Review of the Positive Behavior Support Training Curriculum. *Journal of Positive Behavior Interventions*, 8(3), 183-188. <https://doi.org/10.1177/10983007060080030701>

- Kittelman, A., McIntosh, K., & Hoselton, R. (2019). Adoption of PBIS within school districts. *J Sch Psychol*, 76, 159-167. <https://doi.org/10.1016/j.jsp.2019.03.007>
- Klipker, K., Baumgarten, F., Göbel, K., Lampert, T., & Hölling, H. (2018). Psychische Auffälligkeiten bei Kindern und Jugendlichen in Deutschland – Querschnittergebnisse aus KiGGS Welle 2 und Trends. *Journal of Health Monitoring*, 3(3), 37-45. <https://doi.org/10.17886/RKI-GBE-2018-077>
- Klomek, A., Stanley, B., & Sourander, A. (2014). The Bullying Prevention Plan: An Approach to Youth who Bully Others. *Adolescent Psychiatry*, 4(3), 185-193. <https://doi.org/10.2174/221067660403140912162846>
- KMK. (2024). *Statistische Veröffentlichung der Kultursministerkonferenz- Dokumentation Nr. 239*. Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland.
- Knauer, V. (2019). Die Schule als (Re)Produzentin sozialer Ungleichheit? Antimuslimischer Rassismus im deutschen Bildungssystem. In B. Thiessen, C. Dannenbeck, & M. Wolff (Hrsg.), *Sozialer Wandel und Kohäsion* (S. 201-221). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-25765-1_13
- Koenen, A.-K., Vervoort, E., Kelchtermans, G., Verschueren, K., & Spilt, J. L. (2017). Teachers' Daily Negative Emotions in Interactions With Individual Students in Special Education. *Journal of Emotional and Behavioral Disorders*, 27(1), 37-51. <https://doi.org/10.1177/1063426617739579>
- Krasny-Pacini, A., & Evans, J. (2018). Single-case experimental designs to assess intervention effectiveness in rehabilitation: A practical guide. *Ann Phys Rehabil Med*, 61(3), 164-179. <https://doi.org/10.1016/j.rehab.2017.12.002>
- Kratochwill, T. R., Hitchcock, J. H., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M., & Shadish, W. R. (2012). Single-Case Intervention Research Design Standards. *Remedial and Special Education*, 34(1), 26-38. <https://doi.org/10.1177/0741932512452794>
- Kratochwill, T. R., & Levin, J. R. (2010). Enhancing the scientific credibility of single-case intervention research: randomization to the rescue. *Psychol Methods*, 15(2), 124-144. <https://doi.org/10.1037/a0017736>
- Kratochwill, T. R., Levin, J. R., Morin, K. L., & Lindstrom, E. R. (2022). Examining and Enhancing the Methodological Quality of Nonconcurrent Multiple-Baseline Designs. *Perspect Behav Sci*, 45(3), 651-660. <https://doi.org/10.1007/s40614-022-00341-2>
- Krull, J., Urton, K., Kulawiak, P. R., Wilbert, J. r., & Hennemann, T. (2022). Social-relational classroom climate and its link to primary students' behavioral problems. *Empirische Sonderpädagogik*, 14(2), 154-175. <https://doi.org/https://doi.org/10.25656/01:25769>

- Kulkarni, T., Sullivan, A. L., & Kim, J. (2020). Externalizing Behavior Problems and Low Academic Achievement: Does a Causal Relation Exist? *Educational Psychology Review*, 33(3), 915-936. <https://doi.org/10.1007/s10648-020-09582-6>
- Kunzelmann, H. P., Cohen, M. A., Hulten, W. J., Martin, G. L., & Mingo, A. R. (1970). *Precision teaching: An initial training sequence*. WA: Special Child Publications.
- Kuriyan, A. B., Pelham, W. E., Jr., Molina, B. S., Waschbusch, D. A., Gnagy, E. M., Sibley, M. H., Babinski, D. E., Walther, C., Cheong, J., Yu, J., & Kent, K. M. (2013). Young adult educational and vocational outcomes of children diagnosed with ADHD. *J Abnorm Child Psychol*, 41(1), 27-41. <https://doi.org/10.1007/s10802-012-9658-z>
- Lahey, B. B., Krueger, R. F., Rathouz, P. J., Waldman, I. D., & Zald, D. H. (2017). A hierarchical causal taxonomy of psychopathology across the life span. *Psychol Bull*, 143(2), 142-186. <https://doi.org/10.1037/bul0000069>
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1), 159-174. <https://doi.org/10.2307/2529310>
- Lane, K. L., Kalberg, J. R., Bruhn, A. L., E. Mahoney, M., & A. Driscoll, S. (2008). Primary Prevention Programs at the Elementary Level: Issues of Treatment Integrity, Systematic Screening, and Reinforcement. *EDUCATION AND TREATMENT OF CHILDREN*, 31(4), 465-494. <https://doi.org/10.1353/etc.0.0033>
- Langer, S., & Schnebel, S. (2020). Förderkonzepte an Grundschulen – eine Studie zu good-practice Schulen. In *Diversität und soziale Ungleichheit* (S. 149-153). https://doi.org/10.1007/978-3-658-27529-7_22
- Laucht, M., Esser, G., & Schmidt, M. H. (2000). Externalisierende und internalisierende Störungen in der Kindheit: Untersuchungen zur Entwicklungspsychopathologie. *Zeitschrift für Klinische Psychologie und Psychotherapie*, 29(4), 284-292. <https://doi.org/10.1026//0084-5345.29.4.284>
- Lavigne, J. V., Lebailly, S. A., Hopkins, J., Gouze, K. R., & Binns, H. J. (2009). The prevalence of ADHD, ODD, depression, and anxiety in a community sample of 4-year-olds. *J Clin Child Adolesc Psychol*, 38(3), 315-328. <https://doi.org/10.1080/15374410902851382>
- Ledford, J. R., Lambert, J. M., Pustejovsky, J. E., Zimmerman, K. N., Hollins, N., & Barton, E. E. (2022). Single-Case-Design Research in Special Education: Next-Generation Guidelines and Considerations. *Exceptional Children*, 89(4), 379-396. <https://doi.org/10.1177/00144029221137656>
- Lee, A., & Gage, N. A. (2020). Updating and expanding systematic reviews and meta-analyses on the effects of school-wide positive behavior interventions and supports. *Psychology in the Schools*, 57(5), 783-804. <https://doi.org/10.1002/pits.22336>

- Lee, J. M., Kim, J., Hong, J. S., & Marsack-Topolewski, C. N. (2021). From Bully Victimization to Aggressive Behavior: Applying the Problem Behavior Theory, Theory of Stress and Coping, and General Strain Theory to Explore Potential Pathways. *J Interpers Violence, 36*(21-22), 10314-10337. <https://doi.org/10.1177/0886260519884679>
- Lehr, D. (2014). Forschung zur Belastung und Beanspruchung im Lehrerberuf Belastung und Beanspruchung im Lehrerberuf - Gesundheitliche Situation und Evidenz für Risikofaktoren. In *Handbuch der Forschung zum Lehrerberuf* (2 Aufl., S. 947-967). Waxmann.
- Leidig, T., Casale, G., Wilbert, J., Hennemann, T., Volpe, R. J., Briesch, A., & Grosche, M. (2022). Individual, generalized, and moderated effects of the good behavior game on at-risk primary school students: A multilevel multiple baseline study using behavioral progress monitoring. *Frontiers in Education, 7*. <https://doi.org/10.3389/educ.2022.917138>
- Lembke, E. S., Garman, C., Deno, S. L., & Stecker, P. M. (2010). One Elementary School's Implementation of Response to Intervention (RTI). *Reading & Writing Quarterly, 26*(4), 361-373. <https://doi.org/10.1080/10573569.2010.500266>
- Lenz, A. S. (2017). Calculating Effect Size in Single-Case Research. *Measurement and Evaluation in Counseling and Development, 46*(1), 64-73. <https://doi.org/10.1177/0748175612456401>
- Levin, J. R., Ferron, J. M., & Gafurov, B. S. (2018). Comparison of randomization-test procedures for single-case multiple-baseline designs. *Dev Neurorehabil, 21*(5), 290-311. <https://doi.org/10.1080/17518423.2016.1197708>
- Li, P.-H., Mayer, D., & Malmberg, L.-E. (2022). Teacher well-being in the classroom: A micro-longitudinal study. *Teaching and Teacher Education, 115*. <https://doi.org/10.1016/j.tate.2022.103720>
- Liebers, K., Pohlmann-Rother, S., & Lange, S. D. (2023). *Leistungsheterogenität in der Grundschule: Umgang mit Vielfalt im Unterricht*. Kohlhammer Verlag.
- Lipien, L., Kirby, M., & Ferron, J. M. (2023). Single-Case Designs. In *Handbook of Applied Behavior Analysis* (S. 347-362). https://doi.org/10.1007/978-3-031-19964-6_20
- Liu, J. (2004). Childhood externalizing behavior: theory and implications. *J Child Adolesc Psychiatr Nurs, 17*(3), 93-103. <https://doi.org/10.1111/j.1744-6171.2004.tb00003.x>
- Lohrmann, S., Forman, S., Martin, S., & Palmieri, M. (2008). Understanding School Personnel's Resistance to Adopting Schoolwide Positive Behavior Support at a Universal Level of Intervention. *Journal of Positive Behavior Interventions, 10*(4), 256-269. <https://doi.org/10.1177/1098300708318963>

- Losinski, M., Ennis, R. P., Shaw, A., & Gage, N. A. (2021). Supporting Students within an MTSS Framework Using SRSD Fractions: Results of a Regression Discontinuity Design. *Learning Disabilities Research & Practice, 36*(3), 213-223. <https://doi.org/10.1111/ldrp.12253>
- Lower, A., Young, K. R., Christensen, L., Caldarella, P., Williams, L., & Wills, H. (2016). Effects of a Tier 3 Self-Management Intervention Implemented with and without Treatment Integrity. *EDUCATION AND TREATMENT OF CHILDREN, 39*(4), 493-520. <https://doi.org/10.1353/etc.2016.0022>
- Luo, L., Reichow, B., Snyder, P., Harrington, J., & Polignano, J. (2020). Systematic Review and Meta-Analysis of Classroom-Wide Social-Emotional Interventions for Preschool Children. *Topics in Early Childhood Special Education, 42*(1), 4-19. <https://doi.org/10.1177/0271121420935579>
- Lütje-Klose, B., Miller, S., & Ziegler, H. (2014). Professionalisierung für die inklusive Schule als Herausforderung für die LehrerInnenbildung. *Soziale Passagen, 6*(1), 69-84. <https://doi.org/10.1007/s12592-014-0165-7>
- Ma, H. H. (2006). An alternative method for quantitative synthesis of single-subject researches: percentage of data points exceeding the median. *Behav Modif, 30*(5), 598-617. <https://doi.org/10.1177/0145445504272974>
- MacLeod, K. S., Hawken, L. S., O'Neill, R. E., & Bundock, K. (2016). Combining Tier 2 and Tier 3 Supports for Students with Disabilities in General Education Settings. *Journal of Educational Issues, 2*(2), 331-351.
- 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., Blumenthal, Y., Diehl, K., Schöning, A., & Sikora, S. (2014). Das Rügener Inklusionsmodell (RIM). RTI in der Praxis. In M. Hasselhorn, W. Schneider, & U. Trautwein (Hrsg.), *Lernverlaufsdagnostik*. (S. 101-125). Hogrefe.
- 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*(3), 104-118.
- Mahlau, K., Diehl, K., Voss, S., & Hartke, B. (2011). Das Rügener Inklusionsmodell (RIM) - Konzeption einer inklusiven Grundschule. *Zeitschrift für Heilpädagogik*(11), 464-472.
- Manolov, R., Gast, D. L., Perdices, M., & Evans, J. J. (2014). Single-case experimental designs: reflections on conduct and analysis. *Neuropsychol Rehabil, 24*(3-4), 634-660. <https://doi.org/10.1080/09602011.2014.903199>
- Manolov, R., & Moeyaert, M. (2017). Recommendations for Choosing Single-Case Data Analytical Techniques. *Behav Ther, 48*(1), 97-114. <https://doi.org/10.1016/j.beth.2016.04.008>

- Manolov, R., & Solanas, A. (2008). Comparing $N = 1$ effect size indices in presence of autocorrelation. *Behav Modif*, *32*(6), 860-875. <https://doi.org/10.1177/0145445508318866>
- Marlowe, A. (2021). Multi-Tiered System of Supports: A Case Study Examining Effective MTSS Implementation at the Middle School Level. *Doctor of Education Dissertations*, *36*.
- Marraccini, M. E., Lindsay, C. A., Griffin, D., Greene, M. J., Simmons, K. T., & Ingram, K. M. (2023). A Trauma- and Justice, Equity, Diversity, and Inclusion (JEDI)-Informed Approach to Suicide Prevention in School: Black Boys' Lives Matter. *School Psych Rev*, *52*(3), 292-315. <https://doi.org/10.1080/2372966x.2021.2010502>
- Martella, R. C., Marchand-Martella, N. E., Woods, B., Thompson, S., Crockett, C. N., Northrup, E., Benner, G. J., & Ralston, N. C. (2010). Positive behavior support: Analysis of consistency between office discipline referrals and teacher recordings of disruptive classroom behaviors. *Behavioral Development Bulletin*, *16*(1), 25-33. <https://doi.org/10.1037/h0100517>
- Matza, L. S., Paramore, C., & Prasad, M. (2005). A review of the economic burden of ADHD. *Cost Eff Resour Alloc*, *3*, 5. <https://doi.org/10.1186/1478-7547-3-5>
- Mayes, R. D., & Byrd, J. A. (2022). An Antiracist Framework for Evidence-Informed School Counseling Practice. *Professional School Counseling*, *26*(1a). <https://doi.org/10.1177/2156759x221086740>
- Maynard, B. R., Brendel, K. E., Bulanda, J. J., Heyne, D., Thompson, A. M., & Pigott, T. D. (2015). Psychosocial Interventions for School Refusal with Primary and Secondary School Students: A Systematic Review. *Campbell Systematic Reviews*, *11*(1), 1-76. <https://doi.org/10.4073/csr.2015.12>
- Mayring, P., & Fenzl, T. (2019). Qualitative Inhaltsanalyse. In N. Baur & J. Blasius (Hrsg.), *Handbuch Methoden der empirischen Sozialforschung* (S. 633-648). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-21308-4_42
- McHugh, M. L. (2012). Interrater reliability: the kappa statistic. *Biochem Med (Zagreb)*, *22*(3), 276-282.
- McIntosh, K., Campbell, A. L., Carter, D. R., & Zumbo, B. D. (2009). Concurrent Validity of Office Discipline Referrals and Cut Points Used in Schoolwide Positive Behavior Support. *Behavioral Disorders*, *34*(2), 100-113.
- McIntosh, K., & Goodman, S. (2016). *Integrated multi-tiered systems of support: Blending RTI and PBIS*. Guilford Press.
- McIntosh, K., Horner, R. H., & Sugai, G. (2009). Sustainability of Systems-Level Evidence-Based Practices in Schools: Current Knowledge and Future Directions. In W. Sailor, G. Dunlap, G. Sugai, & R. Horner

- (Hrsg.), *Handbook of Positive Behavior Support* (S. 327-352). Springer US. https://doi.org/10.1007/978-0-387-09632-2_14
- McIntosh, K., Kelm, J. L., & Canizal Delabra, A. (2015). In Search of How Principals Change. *Journal of Positive Behavior Interventions*, 18(2), 100-110. <https://doi.org/10.1177/1098300715599960>
- McKenna, M., Castillo, J., Dedrick, R. F., Cheng, K., & Goldstein, H. (2021). Speech-Language Pathologist Involvement in Multi-Tiered System of Supports Questionnaire: Advances in Interprofessional Practice. *Lang Speech Hear Serv Sch*, 52(2), 597-611. https://doi.org/10.1044/2020_LSHSS-20-00084
- McLaughlin, T. W., Denney, M. K., Snyder, P. A., & Welsh, J. L. (2011). Behavior Support Interventions Implemented by Families of Young Children. *Journal of Positive Behavior Interventions*, 14(2), 87-97. <https://doi.org/10.1177/1098300711411305>
- Mecheril, P., & Füllekruss, D. (2023). Bildung als Entprovinzialisierung. In M. S. Baader, T. Freytag, & K. Kempa (Hrsg.), *Politische Bildung in Transformation – Transdisziplinäre Perspektiven* (S. 147-161). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-41027-8_9
- Michie, S., Fixsen, D., Grimshaw, J. M., & Eccles, M. P. (2009). Specifying and reporting complex behaviour change interventions: the need for a scientific method. *Implement Sci*, 4, 40. <https://doi.org/10.1186/1748-5908-4-40>
- Mikami, A. Y., Reuland, M. M., Griggs, M. S., & Jia, M. (2013). Collateral Effects of a Peer Relationship Intervention for Children With Attention Deficit Hyperactivity Disorder on Typically Developing Classmates. *School Psychology Review*, 42(4), 458-476.
- Miller, C. J., Smith, S. N., & Pugatch, M. (2020). Experimental and quasi-experimental designs in implementation research. *Psychiatry Res*, 283, 112452. <https://doi.org/10.1016/j.psychres.2019.06.027>
- Miller, F. G., Crovello, N. J., & Chafouleas, S. M. (2017). Progress Monitoring the Effects of Daily Report Cards Across Elementary and Secondary Settings Using Direct Behavior Rating: Single Item Scales. *Assessment for Effective Intervention*, 43(1), 34-47. <https://doi.org/10.1177/1534508417691019>
- Miller, F. G., & Lee, D. L. (2013). Do Functional Behavioral Assessments Improve Intervention Effectiveness for Students Diagnosed with ADHD? A Single-Subject Meta-Analysis. *Journal of Behavioral Education*, 22(3), 253-282. <https://doi.org/10.1007/s10864-013-9174-4>
- Mitchell, B. S., Stormont, M., & Gage, N. A. (2017). Tier Two Interventions Implemented within the Context of a Tiered Prevention Framework. *Behavioral Disorders*, 36(4), 241-261. <https://doi.org/10.1177/019874291103600404>

- Moeyaert, M., Dehghan-Chaleshtori, M., Xu, X., & Yang, P. (2023). Single-case design meta-analyses in education and psychology: a systematic review of methodology. *Front Res Metr Anal*, 8, 1190362. <https://doi.org/10.3389/frma.2023.1190362>
- Mohammadi, M. R., Salmanian, M., & Keshavarzi, Z. (2021). The Global Prevalence of Conduct Disorder: A Systematic Review and Meta-Analysis. *Iran J Psychiatry*, 16(2), 205-225. <https://doi.org/10.18502/ijps.v16i2.5822>
- Moldenhauer, A. (2023). Zum Wandel von Schule und Unterricht.
- Molela Moukara, F. (2023). Rassismus im Bildungssystem: die schulische Verantwortung im Kampf gegen Rassismus. In B. Aygün, P. Bühler, R. Daraboš, S. Eissa, I. Hagen-Jeske, I. Helmi Hans, M. Arife Kanbur, F. Molela Moukara, & S. Ogiemwonyi (Hrsg.), *RassisMuss MachtKritisch: interdisziplinäre Perspektiven auf Rassismusforschung in Augsburg* (S. 67-88). PubliQation.
- Morse, T. E. (2023). Designing Appropriate Small Group Intensive Instruction within an MTSS for Students with Low Incidence Disabilities. *Journal of the American Academy of Special Education Professionals*, 89-101.
- Müller, M. C. (2021). Internalisierendes und externalisierendes Verhalten von Kindern und ihre Teilhabe an Interaktionen. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 53(3-4), 105-115. <https://doi.org/10.1026/0049-8637/a000244>
- Murawski, W. W., & Hughes, C. E. (2009). Response to Intervention, Collaboration, and Co-Teaching: A Logical Combination for Successful Systemic Change. *Preventing School Failure: Alternative Education for Children and Youth*, 53(4), 267-277. <https://doi.org/10.3200/psfl.53.4.267-277>
- Nelen, M. J. M., Scholte, R. H. J., Blonk, A. M., van der Veld, W. M., Nelen, W. B. L., & Denessen, E. (2021). School-wide positive behavioral interventions and supports in Dutch elementary schools: Exploring effects. *Psychology in the Schools*, 58(6), 992-1006. <https://doi.org/10.1002/pits.22483>
- Nelson, M. A., Caldarella, P., Hansen, B. D., Graham, M. A., Williams, L., & Wills, H. P. (2018). Improving Student Behavior in Art Classrooms: An Exploratory Study of CW-FIT Tier 1. *Journal of Positive Behavior Interventions*, 20(4), 227-238. <https://doi.org/10.1177/1098300718762744>
- Nikstat, A., & Riemann, R. (2023). Differences in Parenting Behavior are Systematic Sources of the Non-shared Environment for Internalizing and Externalizing Problem Behavior. *Behav Genet*, 53(1), 25-39. <https://doi.org/10.1007/s10519-022-10125-8>
- Ninci, J. (2023). Single-Case Data Analysis: A Practitioner Guide for Accurate and Reliable Decisions. *Behav Modif*, 47(6), 1455-1481. <https://doi.org/10.1177/0145445519867054>

- Nitz, J., Brack, F., Hertel, S., Krull, J., Stephan, H., Hennemann, T., & Hanisch, C. (2023). Multi-tiered systems of support with focus on behavioral modification in elementary schools: A systematic review. *Heliyon*, 9(6). <https://doi.org/10.1016/j.heliyon.2023.e17506>
- Nitz, J., Eiben, K., Hanisch, C., Hagen, T., Krull, J., Verbeck, L., Rauterkus, H., & Hennemann, T. (2024). Mehrstufige Förderung in der inklusiven Grundschule – erste qualitative und quantitative Befunde zur Wirksamkeit und Umsetzung des Förderansatzes Multimo. In J. König, C. Hanisch, P. Hanke, T. Hennemann, K. Kaspar, M. Martens, & S. Strauß (Hrsg.), *Teachers and their teaching matters – Auf die Lehrperson und ihren Unterricht kommt es an. 10 Jahre empirische Forschung im IZef der Universität zu Köln*. Waxmann.
- Nitz, J., Hagen, T., Krull, J., Verbeck, L., Eiben, K., Hanisch, C., & Hennemann, T. (2023). Tiers 1 and 2 of a German MTSS: impact of a multiple baseline study on elementary school students with disruptive behavior. *Frontiers in Education*, 8. <https://doi.org/10.3389/educ.2023.1208854>
- Nitz, J., Niederelz, A., Hanisch, C., & Hennemann, T. (2024). Herausforderungen und Gelingensbedingungen erfolgreicher Implementation eines mehrstufigen Förderansatzes an Grundschulen. *Zeitschrift für Heilpädagogik*, 75(1), 4-17.
- Noltemeyer, A., Palmer, K., James, A. G., & Wiechman, S. (2019). School-Wide Positive Behavioral Interventions and Supports (SWPBIS): A Synthesis of Existing Research. *International Journal of School & Educational Psychology*, 7(4), 253-262.
- Norozpour, M., Pourshahbaz, A., Poursharifi, H., Dolatshahi, B., & Habibi, N. (2023). Prevalence of externalising disorder comorbidities in adolescents: a systematic review and meta-analysis protocol. *BMJ open*, 13(7), e065680. <https://doi.org/10.1136/bmjopen-2022-065680>
- O'Connor, P. D., Stuck, G. B., & Wyne, M. D. (2016). Effects of a Short-Term Intervention Resource-Room Program On Task Orientation and Achievement. *The Journal of Special Education*, 13(4), 375-385. <https://doi.org/10.1177/002246697901300405>
- O'Neill, Horner, R., Albin, R., Sprague, J., Storey, K., & Newton, J. (1997). *Functional assessment and program development for problem behavior: A practical handbook* (2 Aufl.). Brooks/Cole.
- Öğülmüş, K., & Vuran, S. (2016). Schoolwide Positive Behavioral Interventions and Support Practices: Review of Studies in the Journal of Positive Behavior Interventions. *Educational Sciences: Theory & Practice*, 16(5), 1693-1710. <https://doi.org/10.12738/estp.2016.5.0264>
- Opp, G. (1998). Gefühls- und Verhaltensstörungen. Begriffliche Problemstellungen und Lösungsversuche. *Zeitschrift für Heilpädagogik*, 49(11), 490-496.

- Oslund, E. L., Elleman, A. M., & Wallace, K. (2021). Factors Related to Data-Based Decision-Making: Examining Experience, Professional Development, and the Mediating Effect of Confidence on Teacher Graph Literacy. *J Learn Disabil*, *54*(4), 243-255. <https://doi.org/10.1177/0022219420972187>
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hrobjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., McGuinness, L. A., Stewart, L. A., Thomas, J., Tricco, A. C., Welch, V. A., Whiting, P., & Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *Syst Rev*, *10*(1), 89. <https://doi.org/10.1186/s13643-021-01626-4>
- Parker, R. I., Hagan-Burke, S., & Vannest, K. (2016). Percentage of All Non-Overlapping Data (PAND). *The Journal of Special Education*, *40*(4), 194-204. <https://doi.org/10.1177/00224669070400040101>
- Parker, R. I., & Vannest, K. (2009). An improved effect size for single-case research: nonoverlap of all pairs. *Behav Ther*, *40*(4), 357-367. <https://doi.org/10.1016/j.beth.2008.10.006>
- Parker, R. I., Vannest, K. J., & Brown, L. (2009). The Improvement Rate Difference for Single-Case Research. *Exceptional Children*, *75*(2), 135-150. <https://doi.org/10.1177/001440290907500201>
- Parker, R. I., Vannest, K. J., & Davis, J. L. (2011). Effect size in single-case research: a review of nine nonoverlap techniques. *Behav Modif*, *35*(4), 303-322. <https://doi.org/10.1177/0145445511399147>
- Parker, R. I., Vannest, K. J., Davis, J. L., & Sauber, S. B. (2011). Combining nonoverlap and trend for single-case research: Tau-U. *Behav Ther*, *42*(2), 284-299. <https://doi.org/10.1016/j.beth.2010.08.006>
- Pas, E. T., Johnson, S. R., Debnam, K. J., Hulleman, C. S., & Bradshaw, C. P. (2019). Examining the Relative Utility of PBIS Implementation Fidelity Scores in Relation to Student Outcomes. *Remedial and Special Education*, *40*(1), 6-15. <https://doi.org/10.1177/0741932518805192>
- Pearce, J., Mann, M. K., Jones, C., van Buschbach, S., Olf, M., & Bisson, J. I. (2012). The most effective way of delivering a train-the-trainers program: a systematic review. *J Contin Educ Health Prof*, *32*(3), 215-226. <https://doi.org/10.1002/chp.21148>
- Pearce, L. R. (2009). Helping Children with Emotional Difficulties: A Response to Intervention Investigation. *Rural Educator*, *30*(2), 34-46.
- Peel, A. J., Oginni, O., Assary, E., Krebs, G., Lockhart, C., McGregor, T., Palaiologou, E., Ronald, A., Danese, A., & Eley, T. C. (2023). A multivariate genetic analysis of anxiety sensitivity, environmental sensitivity and reported life events in adolescents. *J Child Psychol Psychiatry*, *64*(2), 289-298. <https://doi.org/10.1111/jcpp.13725>

- Pennington, B., & McComas, J. J. (2017). Effects of the good behavior game across classroom contexts. *J Appl Behav Anal*, 50(1), 176-180. <https://doi.org/10.1002/jaba.357>
- Perry, K. J., & Price, J. M. (2018). Concurrent child history and contextual predictors of children's internalizing and externalizing behavior problems in foster care. *Child Youth Serv Rev*, 84, 125-136. <https://doi.org/10.1016/j.chilyouth.2017.11.016>
- Petermann, F. (2005). Zur Epidemiologie psychischer Störungen im Kindes- und Jugendalter. *Kindheit und Entwicklung*, 14(1), 48-57. <https://doi.org/10.1026/0942-5403.14.1.48>
- Petermann, F. (2014). Implementationsforschung: Grundbegriffe und Konzepte. *Psychologische Rundschau*, 65(3), 122-128. <https://doi.org/10.1026/0033-3042/a000214>
- Petermann, U., & Petermann, F. (2013). *LSL: Lehrereinschätzliste für Sozial- und Lernverhalten*. Hogrefe.
- Petermann, U., & Petermann, F. (2014). *Schülereinschätzliste für Sozial- und Lernverhalten*. Hogrefe.
- Phillips, J. J., Bruce, M. D., & Bell, M. A. (2023). Setting the stage: Biopsychosocial predictors of early childhood externalizing behaviors. *Dev Psychobiol*, 65(4). <https://doi.org/10.1002/dev.22391>
- Pinkelman, S. E., McIntosh, K., Rasplica, C. K., Berg, T., & Strickland-Cohen, M. K. (2015). Perceived Enablers and Barriers Related to Sustainability of School-Wide Positive Behavioral Interventions and Supports. *Behavioral Disorders*, 40(3), 171-183. <https://doi.org/10.17988/0198-7429-40.3.171>
- Pirani-McGurl, C., Leonard, K., Cohen-Goldberg, S., Soule, K., & Burns, D. (2022). Systems and Processes to Support Data-Based Decision Making Within an MTSS Model. *Perspectives on Language & Literacy*, 48(1), 27-30.
- Plück, J., Wiczorrek, E., Wolff Metternich-Kaizman, T., & Döpfner, M. (2006). *Präventionsprogramm für Expansives Problemverhalten (PEP): Ein Manual für Eltern- und Erziehergruppen*. Hogrefe.
- Polanczyk, G., de Lima, M. S., Horta, B. L., Biederman, J., & Rohde, L. A. (2007). The worldwide prevalence of ADHD: a systematic review and metaregression analysis. *Am J Psychiatry*, 164(6), 942-948. <https://doi.org/10.1176/ajp.2007.164.6.942>
- Polanczyk, G. V., Salum, G. A., Sugaya, L. S., Caye, A., & Rohde, L. A. (2015). Annual research review: A meta-analysis of the worldwide prevalence of mental disorders in children and adolescents. *J Child Psychol Psychiatry*, 56(3), 345-365. <https://doi.org/10.1111/jcpp.12381>
- Polanczyk, G. V., Willcutt, E. G., Salum, G. A., Kieling, C., & Rohde, L. A. (2014). ADHD prevalence estimates across three decades: an updated systematic review and meta-regression analysis. *Int J Epidemiol*, 43(2), 434-442. <https://doi.org/10.1093/ije/dyt261>

- Polanin, J. R., Espelage, D. L., Grotzinger, J. K., Ingram, K., Michaelson, L., Spinney, E., Valido, A., Sheikh, A. E., Torgal, C., & Robinson, L. (2022). A Systematic Review and Meta-analysis of Interventions to Decrease Cyberbullying Perpetration and Victimization. *Prev Sci*, 23(3), 439-454. <https://doi.org/10.1007/s11121-021-01259-y>
- Pool Maag, S. (2022). Multiprofessionelle Zusammenarbeit an inklusiven Schulen. Nicht die Zuständigkeit, sondern das Ergebnis zählt. *Schweizerische Zeitschrift für Heilpädagogik*, 28, 15-21. <https://doi.org/10.5281/zenodo.6542713>
- Posit-Team. (2022). *RStudio: Integrated Development Environment for R*. In PBC, Posit Software. <http://www.posit.co/>.
- Powers, K., & Mandal, A. (2011). Tier III assessments, data-based decision-making, and interventions. *Contemporary School Psychology*, 15, 21-33.
- Prenzel, A. (2013). *Inklusive Bildung in der Primarstufe. Eine wissenschaftliche Expertise des Grundschulverbandes*. Grundschulverband e.V. <https://doi.org/10.25656/01:18827>
- Pullen, P. C., van Dijk, W., Gonsalves, V. E., Lane, H. B., & Ashworth, K. E. (2018). RTI and MTSS: Response to Intervention and Multi-Tiered Systems of Support How Do They Differ and How Are They the Same, If at All? In M. J. K. Paige C. Pullen (Hrsg.), *Handbook of Response to Intervention and Multi-Tiered Systems of Support*. Routledge.
- Pustejovsky, J. E., Swan, D. M., & English, K. W. (2023). An Examination of Measurement Procedures and Characteristics of Baseline Outcome Data in Single-Case Research. *Behav Modif*, 47(6), 1423-1454. <https://doi.org/10.1177/0145445519864264>
- Pyle, K., & Fabiano, G. A. (2017). Daily Report Card Intervention and Attention Deficit Hyperactivity Disorder: A Meta-Analysis of Single-Case Studies. *Exceptional Children*, 83(4), 378-395. <https://doi.org/10.1177/0014402917706370>
- Quinn, M. M., Gable, R. A., Fox, J., Rutheford, R. B., Rutherford, R. B., Van Acker, R., & Conroy, M. (2001). Putting Quality Functional Assessment into Practice in Schools: A Research Agenda on Behalf of E/BD Students. *EDUCATION AND TREATMENT OF CHILDREN*, 24(3), 261-275.
- Reinelt, T., Samdan, G., Kiel, N., & Petermann, F. (2019). Frühkindliche Prädiktoren externalisierender Verhaltensauffälligkeiten. *Kindheit und Entwicklung*, 28(1), 19-32. <https://doi.org/10.1026/0942-5403/a000268>
- Reiss, F., Napp, A. K., Erhart, M., Devine, J., Dadaczynski, K., Kaman, A., & Ravens-Sieberer, U. (2023). [Prevention perspective: mental health of schoolchildren in Germany]. *Bundesgesundheitsblatt*

- Gesundheitsforschung Gesundheitsschutz*, 66(4), 391-401. <https://doi.org/10.1007/s00103-023-03674-8>
(Perspektive Prävention: Psychische Gesundheit von Schülerinnen und Schülern in Deutschland.)
- Renner, K.-H., Heydasch, T., & Ströhlein, G. (2012). *Forschungsmethoden der Psychologie: Von der Fragestellung zur Präsentation*.
- Reynolds, M. C. (1962). A framework for considering some issues in special education. *Exceptional Children*, 28, 367–370.
- Riden, B. S., Taylor, J. C., Lee, D. L., & Scheeler, M. C. (2018). A Synthesis of the Daily Behavior Report Card Literature from 2007 to 2017. *The Journal of Special Education Apprenticeship*, 7(1).
- Rivenbark, J. G., Odgers, C. L., Caspi, A., Harrington, H., Hogan, S., Houts, R. M., Poulton, R., & Moffitt, T. E. (2018). The high societal costs of childhood conduct problems: evidence from administrative records up to age 38 in a longitudinal birth cohort. *J Child Psychol Psychiatry*, 59(6), 703-710. <https://doi.org/10.1111/jcpp.12850>
- Robinson, L. E., Clements, G., Drescher, A., El Sheikh, A., Milarsky, T. K., Hanebutt, R., Graves, K., Delgado, A. V., Espelage, D. L., & Rose, C. A. (2023). Developing a Multi-Tiered System of Support-Based Plan for Bullying Prevention Among Students with Disabilities: Perspectives from General and Special Education Teachers During Professional Development. *School Ment Health*, 1-13. <https://doi.org/10.1007/s12310-023-09589-8>
- Rodriguez, B. J., Loman, S. L., & Borgmeier, C. (2016). Tier 2 Interventions in Positive Behavior Support: A Survey of School Implementation. *Preventing School Failure*, 60(2), 94-105.
- Röhner, C., Decker-Ernst, Y., Salem, S., & Hettich, N. (2020). Kinder nach Flucht und Migration – Herausforderungen an die Integrationskraft des Schul- und Aufnahmesystems. In *Diversität und soziale Ungleichheit* (S. 46-57). https://doi.org/10.1007/978-3-658-27529-7_6
- Rojas-Andrade, R., & Bahamondes, L. L. (2018). Is Implementation Fidelity Important? A Systematic Review on School-Based Mental Health Programs. *Contemporary School Psychology*, 23(4), 339-350. <https://doi.org/10.1007/s40688-018-0175-0>
- Rubow, C. C., Vollmer, T. R., & Joslyn, P. R. (2018). Effects of the Good Behavior Game on student and teacher behavior in an alternative school. *J Appl Behav Anal*, 51(2), 382-392. <https://doi.org/10.1002/jaba.455>
- Ruiz-Olivares, R., Pino, M. J., & Herruzo, J. (2010). Reduction of disruptive behaviors using an intervention based on the Good Behavior Game and the Say-Do-Report Correspondence. *Psychology in the Schools*, 47(10), 1046-1058. <https://doi.org/10.1002/pits.20523>

- Sadjadi, M., Blanchard, L., Brulle, R., & Bonell, C. (2022). Barriers and facilitators to the implementation of Health-Promoting School programmes targeting bullying and violence: a systematic review. *Health Educ Res*, 36(5), 581-599. <https://doi.org/10.1093/her/cyab029>
- Sanetti, L. M. H., & Collier-Meek, M. A. (2015). Data-Driven Delivery of Implementation Supports in a Multi-Tiered Framework: A Pilot Study. *Psychology in the Schools*, 52(8), 815-828. <https://doi.org/10.1002/pits.21861>
- Savage, C., Lewis, J., & Colless, N. (2011). Essentials for implementation: Six years of School Wide Positive Behaviour Support in New Zealand. *New Zealand Journal of Psychology*, 40(1), 29-37.
- Schaffer, G. E. (2023). *Multi-Tiered Systems of Support: A practical guide to preventative practice*. SAGE Publications.
- Schemme, D., Novak, H., & Garcia-Wülfing, I. (2017). *Transfer von Bildungsinnovationen – Beiträge aus der Forschung*.
- Scherzinger, M., & Wettstein, A. (2023). Externalisierende Verhaltensschwierigkeiten als Ursache von Lernschwierigkeiten. *Lern-und Verhaltensschwierigkeiten in der Schule: Erscheinungsformen-Entwicklungsmodelle-Implikationen für die Praxis*, 155.
- Scheuermann, B. K., & Nelson, C. M. (2019). Sustaining PBIS in Secure Care for Juveniles. *EDUCATION AND TREATMENT OF CHILDREN*, 42(4), 537-556. <https://doi.org/10.1353/etc.2019.0025>
- Schubarth, W., & Seidel, A. (2013). Gewalt an Schulen in Deutschland: Diskurse, Befunde, Prävention. In M. A. Marchwacka (Hrsg.), *Gesundheitsförderung im Setting Schule* (S. 259-273). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-00528-3_15
- Scott, T. M., Lingo, A. S., & Hirn, R. G. (2013). *The Academic and Behavior Response to Intervention School Assessment (ASA)*. Center for Instructional and Behavioral Interventions in Schools, University of Louisville.
- Scotti, J. R., Evans, I. M., Meyer, L. H., & Walker, P. (1991). A meta-analysis of intervention research with problem behavior: treatment validity and standards of practice. *Am J Ment Retard*, 96(3), 233-256.
- Scruggs, T. E., & Mastropieri, M. A. (2012). PND at 25. *Remedial and Special Education*, 34(1), 9-19. <https://doi.org/10.1177/0741932512440730>
- Scruggs, T. E., Mastropieri, M. A., & Casto, G. (2016). The Quantitative Synthesis of Single-Subject Research. *Remedial and Special Education*, 8(2), 24-33. <https://doi.org/10.1177/074193258700800206>

- Sewell, A. (2019). An adaption of the Good Behaviour Game to promote social skill development at the whole-class level. *Educational Psychology in Practice*, 36(1), 93-109. <https://doi.org/10.1080/02667363.2019.1695583>
- Sewell, A. (2020). An Adaption of the Good Behaviour Game to Promote Social Skill Development at the Whole-Class Level. *Educational Psychology in Practice*, 36(1), 93-109.
- Shadish, W. R., & Sullivan, K. J. (2011). Characteristics of single-case designs used to assess intervention effects in 2008. *Behav Res Methods*, 43(4), 971-980. <https://doi.org/10.3758/s13428-011-0111-y>
- Shepley, C., & Grisham-Brown, J. (2019). Multi-tiered systems of support for preschool-aged children: A review and meta-analysis. *Early Childhood Research Quarterly*, 47, 296-308. <https://doi.org/10.1016/j.ecresq.2019.01.004>
- Shepley, C., Grisham-Brown, J., & Lane, J. D. (2020). Multitiered Systems of Support in Preschool Settings: A Review and Meta-Analysis of Single-Case Research. *Topics in Early Childhood Special Education*, 41(4), 307-320. <https://doi.org/10.1177/0271121419899720>
- Sherrod, M. D., Getch, Y. Q., & Ziomek-Daigle, J. (2009). The Impact of Positive Behavior Support to Decrease Discipline Referrals with Elementary Students. *Professional School Counseling*, 12(6), 421-427.
- Shinn, M. R., & Garman, G. (2006). *AIMSweb*. Edformation, Inc.
- Simonsen, B., MacSuga-Gage, A. S., Briere, D. E., Freeman, J., Myers, D., Scott, T. M., & Sugai, G. (2013). Multitiered Support Framework for Teachers' Classroom-Management Practices. *Journal of Positive Behavior Interventions*, 16(3), 179-190. <https://doi.org/10.1177/1098300713484062>
- Simonsen, B., Robbie, K., Meyer, K., Freeman, J., Everett, S., & Feinberg A. . (2021). Multi-Tiered System of Supports (MTSS) in the Classroom. *Center on PBIS, University of Oregon*.
- Skaalvik, E. M., & Skaalvik, S. (2010). Teacher self-efficacy and teacher burnout: A study of relations. *Teaching and Teacher Education*, 26(4), 1059-1069. <https://doi.org/10.1016/j.tate.2009.11.001>
- Skorsetz, N., Bonanati, M., & Kucharz, D. (2020). *Diversität und soziale Ungleichheit*. <https://doi.org/10.1007/978-3-658-27529-7>
- Smith, J. D. (2012). Single-case experimental designs: a systematic review of published research and current standards. *Psychol Methods*, 17(4), 510-550. <https://doi.org/10.1037/a0029312>
- Smith, J. D., Borckardt, J. J., & Nash, M. R. (2012). Inferential precision in single-case time-series data streams: how well does the em procedure perform when missing observations occur in autocorrelated data? *Behav Ther*, 43(3), 679-685. <https://doi.org/10.1016/j.beth.2011.10.001>

- 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. *Sch Psychol Q*, 33(1), 83-93. <https://doi.org/10.1037/spq0000209>
- Snijders, T., & Bosker, R. (2011). *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling* (2 Aufl.). Sage.
- Solesbee, C., & Davies, S. C. (2021). Data-Based Decision-Making within an MTSS Framework for Students with Concussion/TBI. *Communique*, 50(4), 18-20.
- Solomon, B. G., Klein, S. A., Hintze, J. M., Cressey, J. M., & Peller, S. L. (2012). A meta-analysis of school-wide positive behavior support: An exploratory study using single-case synthesis. *Psychology in the Schools*, 49(2), 105-121. <https://doi.org/10.1002/pits.20625>
- Sørli, M.-A., & Ogden, T. (2015). School-Wide Positive Behavior Support–Norway: impacts on problem behavior and classroom climate. *International Journal of School & Educational Psychology*, 3(3), 202-217.
- Sørli, M. A., & Ogden, T. (2007). Immediate Impacts of PALS: A school-wide multi-level programme targeting behaviour problems in elementary school. *Scandinavian Journal of Educational Research*, 51(5), 471-492. <https://doi.org/10.1080/00313830701576581>
- Spilles, M., & Hagen, T. (2019). Schulpraktische Erhebungs- und Auswertungsmöglichkeiten von Einzelfalldaten. *Potsdamer Zentrum für empirische Inklusionsforschung (ZEIF)*(3).
- Spilles, M., Hagen, T., & Hennemann, T. (2019). Playing the Good Behavior Game during a Peer-Tutoring Intervention: Effects on Behavior and Reading Fluency of Tutors and Tutees with Behavioral Problems. *Insights into Learning Disabilities*, 16(1), 59-77.
- Stanley, K. (2007). Design of randomized controlled trials. *Circulation*, 115(9), 1164-1169. <https://doi.org/10.1161/CIRCULATIONAHA.105.594945>
- Stephan, S. H., Sugai, G., Lever, N., & Connors, E. (2015). Strategies for integrating mental health into schools via a multitiered system of support. *Child Adolesc Psychiatr Clin N Am*, 24(2), 211-231. <https://doi.org/10.1016/j.chc.2014.12.002>
- Sterling, S., & Huckle, J. (1996). *Education for Sustainability*. Routledge. <https://doi.org/10.4324/9781315070650>
- Stoiber, K. C., & Gettinger, M. (2016). Multi-Tiered Systems of Support and Evidence-Based Practices. In S. R. Jimerson, M. K. Burns, & A. M. VanDerHeyden (Hrsg.), *Handbook of Response to Intervention* (S. 121-141). Springer US. https://doi.org/10.1007/978-1-4899-7568-3_9

- Stoltz, S., Londen, M. v., Deković, M., Castro, B. O. d., & Prinzie, P. (2012). Effectiveness of individually delivered indicated school-based interventions on externalizing behavior. *International Journal of Behavioral Development*, 36(5), 381-388. <https://doi.org/10.1177/0165025412450525>
- Sturmeijer, P. (2022). Psychological and Sociological Theories of Violence and Aggression. In P. Sturmeijer (Hrsg.), *Violence and Aggression* (S. 215-232). Springer International Publishing. https://doi.org/10.1007/978-3-031-04386-4_9
- Sugai, G., & Horner, R. R. (2019). A Promising Approach for Expanding and Sustaining School-Wide Positive Behavior Support. *School Psychology Review*, 35(2), 245-259. <https://doi.org/10.1080/02796015.2006.12087989>
- Suri, H., & Clarke, D. (2009). Advancements in Research Synthesis Methods: From a Methodologically Inclusive Perspective. *Review of Educational Research*, 79(1), 395-430. <https://doi.org/10.3102/0034654308326349>
- Sutherland, M., Lesner, T., Kosty, D., Lussier, C., Smolkowski, K., Turtura, J., Doabler, C. T., & Clarke, B. (2023). Examining Interactions Across Instructional Tiers: Do Features of Tier 1 Predict Student Responsiveness to Tier 2 Mathematics Intervention? *J Learn Disabil*, 56(4), 243-256. <https://doi.org/10.1177/00222194221102644>
- Taillieu, T. L., & Brownridge, D. A. (2015). The Impact of Aggressive Parental Discipline Experienced in Childhood on Externalizing Problem Behavior in Early Adulthood. *Journal of child & adolescent trauma*, 8(4), 253-264. <https://doi.org/10.1007/s40653-015-0063-y>
- te Poel, K., Gollub, P., Siedenbiedel, C., Greiten, S., & Veber, M. (2023). *Heterogenität und Inklusion in den Schulpraktischen Studien* (Bd. 8). Waxmann.
- Tehrani, H. D., Yamini, S., & Vazsonyi, A. T. (2023). The effectiveness of parenting program components on disruptive and delinquent behaviors during early and middle childhood: a component network meta-analysis. *Journal of Experimental Criminology*. <https://doi.org/10.1007/s11292-023-09562-0>
- Timothy, J. L., Barbara, S. M., Robert, T., & Lori, N. (2014). School-Wide Positive Behavior Support. In *Handbook of Classroom Management*. Routledge. <https://doi.org/10.4324/9780203074114.ch3>
- Tobin, T. J., & Sugai, G. (2005). Preventing Problem Behaviors: Primary, Secondary, and Tertiary Level Prevention Interventions for Young Children. *Journal of Early and Intensive Behavior Intervention*, 2(3), 125-144.
- Tschannen-Moran, M., Hoy, A. W., & Hoy, W. K. (2016). Teacher Efficacy: Its Meaning and Measure. *Review of Educational Research*, 68(2), 202-248. <https://doi.org/10.3102/00346543068002202>

- Tully, L. A., & Hunt, C. (2015). Brief Parenting Interventions for Children at Risk of Externalizing Behavior Problems: A Systematic Review. *Journal of Child and Family Studies*, 25(3), 705-719. <https://doi.org/10.1007/s10826-015-0284-6>
- Tyre, A., Feuerborn, L., Beaudoin, K., & Bruce, J. (2019). Middle School Teachers' Concerns for Implementing the Principles of SWPBIS. *Journal of Positive Behavior Interventions*, 22(2), 93-104. <https://doi.org/10.1177/1098300719867858>
- Utley, C. A. (2012). Measuring Cultural Responsiveness in the Classroom Component of a School-Wide Model of Positive Behavior Support at the Elementary Level. *Multicultural Learning and Teaching*, 7(2).
- Vainikainen, M.-P., Hienonen, N., & Hotulainen, R. (2017). Class size as a means of three-tiered support in Finnish primary schools. *Learning and Individual Differences*, 56, 96-104. <https://doi.org/10.1016/j.lindif.2017.05.004>
- Vannest, K. J., Davis, J. L., Davis, C. R., Mason, B. A., & Burke, M. D. (2019). Effective Intervention for Behavior With a Daily Behavior Report Card: A Meta-Analysis. *School Psychology Review*, 39(4), 654-672. <https://doi.org/10.1080/02796015.2010.12087748>
- Vergunst, F., Commisso, M., Geoffroy, M. C., Temcheff, C., Poirier, M., Park, J., Vitaro, F., Tremblay, R., Cote, S., & Orri, M. (2023). Association of Childhood Externalizing, Internalizing, and Comorbid Symptoms With Long-term Economic and Social Outcomes. *JAMA Netw Open*, 6(1). <https://doi.org/10.1001/jamanetworkopen.2022.49568>
- Vierbuchen, M.-C., & Bartels, F. (2020). Die inklusive Schule. In H. Ricking, T. Bolz, B. Rieß, & M. Wittrock (Hrsg.), *Prävention und Intervention bei Verhaltensstörungen: Gestufte Hilfen in der schulischen Inklusion* (S. 48-58). Kohlhammer Verlag.
- Vincent, C. G., Swain-Bradway, J., Tobin, T. J., & May, S. (2011). Disciplinary Referrals for Culturally and Linguistically Diverse Students with and without Disabilities: Patterns Resulting from School-Wide Positive Behavior Support. *Exceptionality*, 19(3), 175-190. <https://doi.org/10.1080/09362835.2011.579936>
- Volpe, R. J., Fabiano, G. A., & Pelham, W. E., Jr. (2013). *Daily Behavior Report Cards: An Evidence-Based System of Assessment and Intervention*. Guilford Publications.
- von Wirth, E., Breuer, D., Mandler, J., Schurmann, S., & Dopfner, M. (2022). Prediction of Educational Attainment and Occupational Functioning in Young Adults With a Childhood Diagnosis of ADHD: Results from the Cologne Adaptive Multimodal Treatment (CAMT) Study. *J Atten Disord*, 26(7), 1018-1032. <https://doi.org/10.1177/10870547211045740>

- von Wirth, E., & Döpfner, M. (2023). *ADHS in der Lerntherapie: Ein verhaltenstherapeutischer Praxisleitfaden*. Hogrefe.
- Voorhees, M. D., Walker, V. L., Snell, M. E., & Smith, C. G. (2013). A Demonstration of Individualized Positive Behavior Support Interventions by Head Start Staff to Address Children's Challenging Behavior. *Research and Practice for Persons with Severe Disabilities*, 38(3), 173-185. <https://doi.org/10.1177/154079691303800304>
- Voß, S., Blumenthal, Y., Ehrich, K., & Kathrin, M. (2020). Multimodale Diagnostik als Ausgangspunkt für spezifische Förderung. Eine Darlegung am Beispiel der Rechtschreibung. *Zeitschrift für Heilpädagogik*, 71, 88-89.
- Voß, S., Blumenthal, Y., Sikora, S., Mahlau, K., Diehl, K., & Hartke, B. (2014). Rügener Inklusionsmodell (RIM) – Effekte eines Beschulungsansatzes nach dem Response to Intervention-Ansatz auf die Rechen und Leseleistungen von Grundschulkindern. *Empirische Sonderpädagogik*, 2, 114-132.
- Voß, S., Marten, K., Diehl, K., Mahlau, K., Sikora, S., Blumenthal, Y., & Hartke, B. (2016). Evaluationsergebnisse des Projekts "Rügener Inklusionsmodell (RIM) - Präventive und Integrative Schule auf Rügen (PISaR)" nach vier Schuljahren. *Zeitschrift für Heilpädagogik*(3), 133-149.
- Walker, B., Cheney, D., Stage, S., Blum, C., & Horner, R. H. (2016). Schoolwide Screening and Positive Behavior Supports. *Journal of Positive Behavior Interventions*, 7(4), 194-204. <https://doi.org/10.1177/10983007050070040101>
- Walker, V. L., Conradi, L. A., Strickland-Cohen, M. K., & Johnson, H. N. (2023). School-wide positive behavioral interventions and supports and students with extensive support needs: a scoping review. *Int J Dev Disabil*, 69(1), 13-28. <https://doi.org/10.1080/20473869.2022.2116232>
- Wanzek, J., Vaughn, S., Scammacca, N., Gatlin, B., Walker, M. A., & Capin, P. (2016). Meta-Analyses of the Effects of Tier 2 Type Reading Interventions in Grades K-3. *Educ Psychol Rev*, 28(3), 551-576. <https://doi.org/10.1007/s10648-015-9321-7>
- Wartenberg, G., Aldrup, K., Grund, S., & Klusmann, U. (2023). Satisfied and High Performing? A Meta-Analysis and Systematic Review of the Correlates of Teachers' Job Satisfaction. *Educational Psychology Review*, 35(4). <https://doi.org/10.1007/s10648-023-09831-4>
- Waschbusch, D. A., Breaux, R. P., & Babinski, D. E. (2018). School-Based Interventions for Aggression and Defiance in Youth: A Framework for Evidence-Based Practice. *School Mental Health*, 11(1), 92-105. <https://doi.org/10.1007/s12310-018-9269-0>
- Weber, C. (2017). *Elterliche Erziehung und externalisierende Verhaltensprobleme von Kindern*. Springer.

- Weiland, C. A., Murakami, E. T., Aguilera, E., & Richards, M. G. (2014). Advocates in Odd Places: Social Justice for Behaviorally Challenged, Minority Students in a Large Urban School District. *Education, Citizenship and Social Justice*, 9(2), 114-127.
- Weist, M. D., Splett, J. W., Halliday, C. A., Gage, N. A., Seaman, M. A., Perkins, K. A., Perales, K., Miller, E., Collins, D., & DiStefano, C. (2022). A randomized controlled trial on the interconnected systems framework for school mental health and PBIS: Focus on proximal variables and school discipline. *J Sch Psychol*, 94, 49-65. <https://doi.org/10.1016/j.jsp.2022.08.002>
- Wexler, D. (2018). School-based multi-tiered systems of support (MTSS): An introduction to MTSS for neuropsychologists. *Appl Neuropsychol Child*, 7(4), 306-316. <https://doi.org/10.1080/21622965.2017.1331848>
- White, O. R. (1986). Precision teaching--precision learning. *Except Child*, 52(6), 522-534. <https://doi.org/10.1177/001440298605200605>
- WHO. (2016). *International statistical classification of diseases and related health problems* (10 Aufl., Bd. 2).
- Wilbert, J., Börnert-Ringleb, M., & Lüke, T. (2022). Statistical Power of Piecewise Regression Analyses of Single-Case Experimental Studies Addressing Behavior Problems. *Frontiers in Education*, 7, 1-13. <https://doi.org/10.3389/feduc.2022.917944>
- Wilbert, J., & Lueke, T. (2022). Scan: Single-case data analyses for single and multiple baseline designs.
- Willcutt, E. G. (2012). The prevalence of DSM-IV attention-deficit/hyperactivity disorder: a meta-analytic review. *Neurotherapeutics*, 9(3), 490-499. <https://doi.org/10.1007/s13311-012-0135-8>
- Wilmers, A., Achenbach, M., & Keller, C. (2021). *Bildung im digitalen Wandel*. Waxmann.
- Wilson, A., Beribisky, N., Desrocher, M., Dlamini, N., Williams, T., & Westmacott, R. (2024). Predictors of externalizing behavior outcomes following pediatric stroke. *Child Neuropsychol*, 30(2), 241-263. <https://doi.org/10.1080/09297049.2023.2185216>
- Wirth, S., Friesen, M., Philipp, K., & Streit, C. (2023). Wie kann die diagnostische Kompetenz von Lehrkräften mit Text-Bild-Vignetten gefördert werden? Konzeption und Erprobung einer Lerngelegenheit für angehende Mathematiklehrkräfte in der Primarstufe. *Herausforderung Lehrer*innenbildung - Zeitschrift zur Konzeption, Gestaltung und Diskussion*, 6(1). <https://doi.org/10.11576/hlz-6245>
- Wiskow, K. M., Urban-Wilson, A., Ishaya, U., DaSilva, A., Nieto, P., Silva, E., & Lopez, J. (2021). A comparison of variations of the good behavior game on disruptive and social behaviors in elementary school classrooms. *Behavior Analysis: Research and Practice*, 21(2), 102-117. <https://doi.org/10.1037/bar0000208>

- Wohnhas, V., Goldan, J., & Lütje-Klose, B. (2024). Beteiligung an Aufgaben als eine Facette der Arbeitszufriedenheit von Lehrkräften in inklusiven Schulen. *QfI - Qualifizierung für Inklusion. Online-Zeitschrift zur Forschung über Aus-, Fort- und Weiterbildung pädagogischer Fachkräfte*, 5(3). <https://doi.org/10.21248/qfi.143>
- Wong, T. Y., Zhang, H., White, T., Xu, L., & Qiu, A. (2021). Common functional brain networks between attention deficit and disruptive behaviors in youth. *Neuroimage*, 245. <https://doi.org/10.1016/j.neuroimage.2021.118732>
- Woolcock, M. (2013). Using case studies to explore the external validity of 'complex' development interventions. *Evaluation*, 19(3), 229-248. <https://doi.org/10.1177/1356389013495210>
- Wu, Y. C., Chen, P. Y., Tsai, S. P., Tsai, S. F., Chou, Y. C., & Chiu, C. Y. (2019). The effects of the class-wide function-related intervention teams on behaviors of an elementary student with autism spectrum disorder in an inclusive classroom in Taiwan. *Int J Dev Disabil*, 65(5), 368-377. <https://doi.org/10.1080/20473869.2019.1647031>
- Yeung, A. S., Mooney, M., Barker, K., & Dobia, B. (2009). Does School-Wide Positive Behaviour System Improve Learning in Primary Schools? Some Preliminary Findings. *New Horizons in Education*, 57(1), 17-32.
- Yoon, D., Yoon, S., Pei, F., & Ploss, A. (2021). The roles of child maltreatment types and peer relationships on behavior problems in early adolescence. *Child Abuse Negl*, 112. <https://doi.org/10.1016/j.chiabu.2020.104921>
- Yucesoy-Ozkan, S., Rakap, S., & Gulboy, E. (2019). Evaluation of treatment effect estimates in single-case experimental research: comparison of twelve overlap methods and visual analysis. *British Journal of Special Education*, 47(1), 67-87. <https://doi.org/10.1111/1467-8578.12294>
- Zarakoviti, E., Shafran, R., Papadimitriou, D., & Bennett, S. D. (2021). The Efficacy of Parent Training Interventions for Disruptive Behavior Disorders in Treating Untargeted Comorbid Internalizing Symptoms in Children and Adolescents: A Systematic Review. *Clin Child Fam Psychol Rev*, 24(3), 542-552. <https://doi.org/10.1007/s10567-021-00349-1>
- Zhai, S., Liang, Y., Lu, C., & He, J. (2024). Influence of parenting styles on children's development of externalizing behaviors: The role of resting respiratory sinus arrhythmia. *International Journal of Behavioral Development*, 48(2), 156-165. <https://doi.org/10.1177/01650254231222434>
- Zhang, Y., Xu, B., Kim, H. H., Muetzel, R., Delaney, S. W., & Tiemeier, H. (2022). Differences in cortical morphology and child internalizing or externalizing problems: Accounting for the co-occurrence. *JCPP Adv*, 2(4). <https://doi.org/10.1002/jcv2.12114>

Anteile der Eigenleistung

Beitrag 1:

- Beitrag JN: Ideengenerierung, Manuskriptverfassung, grafische Umsetzung, Kodierung, Abstract-Screening, Volltext-Screening, Auswertung, Durchführung der Suchen, Einreichung des Manuskripts samt Überarbeitung und Wiedereinreichung
- Beitrag Co-Autor*innen: Kodierung, Abstract-Screening, Volltext-Screening, Auswertung, Korrekturlesen

Beitrag 2:

- Beitrag JN: Ideengenerierung, Literaturrecherche, Manuskriptverfassung, Studienplanung, Stichprobenakquise, Entwicklung der Datenerhebungsmethode, grafische Umsetzung, Auswertung, Durchführung der Datenerhebung, Verfassen der Auswertungs-Skript-Dateien, Einreichung des Manuskripts samt Überarbeitung und Wiedereinreichung
- Beitrag Co-Autor*innen: Planung der Datenerhebung, Stichprobenakquise, Auswertung, Korrekturlesen, Entwicklung der Datenerhebungsmethode

Beitrag 3:

- Beitrag JN: Ideengenerierung, Literaturrecherche, Manuskriptverfassung, Studienplanung, Stichprobenakquise, Planung der Datenerhebung, Entwicklung der Datenerhebungsmethode, grafische Umsetzung, Auswertung, Durchführung der Datenerhebung, Zusammenführen der Daten, Verfassen der Auswertungs-Skript-Dateien, Einreichung des Manuskripts samt Überarbeitung und Wiedereinreichung
- Beitrag Co-Autor*innen: Planung der Datenerhebung, Stichprobenakquise, Auswertung, Korrekturlesen, Entwicklung der Datenerhebungsmethode, Durchführung der Interviews, Durchführung der Datenauswertung

Beitrag 4:

- Beitrag JN: Literaturrecherche, Manuskriptverfassung, grafische Umsetzung, Auswertung, Aufbereiten der Daten, Verfassen der Auswertungs-Skript-Dateien
- Beitrag Co-Autor*innen: Ideengenerierung, Literaturrecherche, Manuskriptverfassen, Studienplanung, Stichprobenakquise, Entwicklung der Datenerhebungsmethode, grafische Umsetzung, Auswertung, Durchführung der Datenerhebung, Zusammenführen der Daten, Verfassen der Auswertungs-Skript-Dateien, Einreichung des Manuskripts samt Überarbeitung und Wiedereinreichung

Beitrag 5:

- Beitrag JN: Ideengenerierung, Literaturrecherche, Manuskriptverfassung, Studienplanung, Stichprobenakquise, Planung der Datenerhebung, Entwicklung der Datenerhebungsmethode, grafische Umsetzung, Auswertung, Durchführung der Datenerhebung, Verfassen der Auswertungs-Skript-Dateien, Einreichung des Manuskripts samt Überarbeitung und Wiedereinreichung
- Beitrag Co-Autor*innen: Planung der Datenerhebung, Korrekturlesen

Beitrag 6:

- Beitrag JN: Ideengenerierung, Literaturrecherche, Manuskriptverfassung, Studienplanung, Stichprobenakquise, Planung der Datenerhebung, Entwicklung der Datenerhebungsmethode, Auswertung, Durchführung der Datenerhebung (Interviews), Kodierung, Verfassen der Auswertungs-Skript-Dateien, Einreichung des Manuskripts samt Überarbeitung und Wiedereinreichung
- Beitrag Co-Autor*innen: Datenerhebung, Studienplanung, Korrekturlesen