CARL VON OSSIETZKY universität OLDENBURG
FAKULTÄT I
BILDUNGS- UND SOZIALWISSENSCHAFTEN

Die Förderung mathematischer Basiskompetenzen zu Beginn der Sekundarstufe I

Evaluation und Implementation eines Trainingsprogramms zur Förderung der mathematischen Basiskompetenzen im inklusiven Setting zu Beginn der Sekundarstufe I

Dissertation

zur Erlangung des Grades einer Doktorin der Philosophie (Dr. phil.)

Vorgelegt von:	Carolina Käter geboren am 24.06.1990 in Wilhelmshaven
Referent:	Prof. Dr. Clemens Hillenbrand
Korreferentin:	Prof. Dr. Conny Melzer
Tag der Abgabe:	15.02.2017
Tag der Disputation:	20.06.2017

gefördert durch
Cusanuswerk
Bischöfliche Studienförderung

EIN DANKESCHÖN

Das Verfassen einer Doktoarbeit kann nur gelingen, wenn es Menschen gibt, die den langen Weg mit dir gemeinsam gehen, Durststrecken überwinden und vor allem auch die kleinen und größeren erreichten Ziele mit dir feiern. Aus diesem Grund möchte ich gerne „Danke“ sagen.

Lieber Clemens, ich danke dir für die vielen gemeinsamen Jahre, die wir gemeinsam in deinem Team an der Universität Oldenburg verbracht haben. Ich danke dir dabei vor allem für die inspirierende Unterstützung in der gesamten Projektphase und im Verfassen der Dissertation, für die vielen Freiräume und für das wertschätzende Feedback.

Liebe Conny, unsere Wege haben sich 2011 in Oldenburg getroffen. Seitdem konnte und durfte ich schon viel von dir lernen. Du bist für mich in vielen Fragen eine wertvolle Ansprechpartnerin und ich danke dir sehr für die tolle Unterstützung in der gesamten Projektphase und in allen Belangen der Dissertation.

Weiterhin Danke sagen möchte ich meinen Projektpartnerinnen und -partnern Julia, Tobias und Carolin. Zu viert haben wir es mit vereinten Kräften geschafft, die hier vorgestellte Fördermaßnahme in zahlreichen Treffen und Gesprächen zu entwickeln und auch Realität werden zu lassen. Auch den mitwirkenden Studierenden möchte ich dabei für die Unterstützung und die fleißige Mitarbeit danken.

Ebenso gilt ein Dank unseren Hilfskräften Celina und Mareike, die uns in allen Belangen des Projekts tatkräftig zur Seite standen.

Für die vergangenen Jahre möchte ich auch einen Dank an meine wunderbaren Kolleginnen und Kollegen des Teams der Pädagogik und Didaktik bei Beeinträchtigungen des Lernens am Institut für Sonder- und Rehabilitationspädagogik an der Universität Oldenburg aussprechen. Ich habe mich bei euch immer aufgehoben und unterstützt gefühlt. Ich habe die tolle gemeinsame Zeit mit euch vor allem während unserer Arbeitswochen auf Spiekeroog sehr genossen! Danke an Alissa, Annika, Flora, Imke, Julia P., Julia V., Marie, Matthias, Ralf und Tobias für die fachliche und zwischenmenschliche Unterstützung!

Ein ganz besonderer Dank gilt meinen Eltern Franz und Hildegard. Eure bedingunglose Unterstützung, euer offenes Ohr, euer Zuspruch und Vertrauen haben mir nicht nur in den Jahren der Promotion besonders den Rücken gestärkt. Ihr habt mich zu dem Menschen gemacht, der ich heute bin und dafür danke ich euch von Herzen! Besonders für den Ruhepol in anstrengenden Zeiten möchte ich mich bei meiner Schwester Leana und meinem Bruder Tobias bedanken. Auch den beiden Familien Börries und Drerup danke ich dafür, dass sie immer an mich und mein Vorhaben geglaubt haben.

Meinen lieben Mädels aus meiner Heimat, meinen Freundinnen und Freunden in Oldenburg und mittlerweile überall auf der Welt danke ich für die Geduld und den oftmaligen Verzicht, der mit dem Schreiben meiner Dissertation leider einherging. Jedes Treffen, jeder Anruf und jede Nachricht haben mich stets ein Stückchen weitergebracht. Dafür danke ich vor allem euch, liebe Kerstin, Malin, Mona und Sarah.

Der größte Dank verbunden mit großer Liebe gilt abschließend dir, Rainer. Ich danke dir für die vielen gemeinsamen Jahre, die du mittlerweile an meiner Seite gehst, für die Gelduld und den beständigen Zuspruch – vor allem in den herausfordernden Zeiten.

Eure Carolina

ZUSAMMENFASSUNG

Die PISA-Ergebnisse aus dem Jahr 2015 zeigen, dass sich 17,2 % der 15-jährigen Schülerinnen und Schüler in der Sekundarstufe I unter oder auf der Kompetenzstufe I (von VI) befindet (Hammer et al., 2016, S. 233). Ihr Mathematikverständnis entspricht allenfalls dem Niveau der vierten Jahrgangsstufe. Das bedeutet, dass ca. „jeder sechste Jugendliche in Deutschland die Mindestanforderungen für ein anschlussfähiges mathematisches Verständnis nicht erreicht und erhebliche Probleme haben dürfte, einen Ausbildungsplatz zu finden beziehungsweise eine Ausbildung erfolgreich abzuschließen sowie anspruchsvollere mathematische Anforderungen im Alltag zu bewältigen“ (Prenzel, Sälzer, Klieme & Köller, 2012, S. 75). Diese Erkenntnisse sind alarmierend. Auch wenn die PISA-Studie die Diskussion um mathematische Bildung aktuell weiter in das Zentrum der bildungspolitischen Diskussion rückt, handelt es sich bei ihren Ergebnissen keinesfalls um neuartige Befunde. So weisen hierzulande u. a. schon Mittelberg (2004), Moser Opitz (2005), Humbach (2008) und die Autorinnen und Autoren der großangelegten MARKUS-Studie (Balzer & Jäger, 2001; Helmke et al., 2000, 2002) sowie der QuaSum-Studie (Lehmann, Peek, et al., 2001) auf die geringen mathematischen Basiskompetenzen vieler Schülerinnen und Schüler in der Sekundarstufe I hin. In Fragen der Förderung mathematischer Basiskompetenzen gerade in inklusiven Lerngruppen wird derzeit die Forderung nach Evidenzbasierung lauter. *Evidenzbasiert* bedeutet hier, dass in Unterrichts- und Fördersettings Maßnahmen und Methoden zum Einsatz kommen, deren Wirksamkeit durch nachvollziehbare wissenschaftliche Forschung bestätigt ist und deren Einsatz sich im betrachteten Einzelfall und aus Sicht der beteiligten Akteure (z. B. der Lehrkraft) eignet (Hillenbrand, 2015, S. 314). Die vorliegende Forschungsarbeit versucht, den offensichtlichen Bedarf an Förderung mathematischer Basiskompetenzen in der Sekundarstufe I systematisch zu bedienen und greift dabei auf das Konzept der Evidenzbasierung zurück. Vorgestellt und evaluiert wird eine vierwöchige Maßnahme zur Förderung mathematischer Basiskompetenzen zu Beginn der Sekundarstufe I in inklusiven Settings. Sie zeichnet sich durch die Implementation evidenzbasierter Fördermethoden wie der Direkten Instruktion, peergestützten bzw. kooperativen Lernformen und der Lernverlaufsdiagnostik aus. Die Förderung soll künftig die Anschlussfähigkeit der Schülerinnen und Schüler an den mathematischen Kompetenzerwerb in der Sekundarstufe I gewährleisten. Die Ergebnisse zeigen, dass sich die Schülerinnen und Schüler der Experimentalgruppe (n = 260) mit einem geringen Vorwissen, geringen kognitiven Fähigkeiten, einer geringen Lesekompetenz und einem risikobehafteten Sozialverhalten durch die Maßnahme im Vergleich zur Kontrollgruppe (n = 376) signifikant in ihren mathematischen Basiskompetenzen verbessern. Langfristige Effekte können für die Schülerinnen und Schüler mit geringen Lesekompetenzen nachgewiesen werden. Die Lernfreude der Schülerinnen und Schüler erweist sich dabei als Prädiktor für die Wirksamkeit der Förderung. In Fragen der Implementationsqualität zeigen sich die Interventionsleiterinnen und -leiter und die Schülerinnen und Schüler insgesamt zufrieden mit der Maßnahme. Diese und überraschende weitere Befunde werden im Kontext bestehender Forschung diskutiert. Eine methodenkritische Analyse eröffnet einen Ausblick auf maßnahmenbezogene Optimierungen wie zum Beispiel einer stärkeren Differenzierung insbesondere für die sehr schwachen Schülerinnen und Schüler.

Inhaltsverzeichnis

1 EINLEITUNG 1

2 ENTWICKLUNG UND ERWERB MATHEMATISCHER KOMPETENZEN 5

2.1 BEGRIFFLICHKEITEN 5

2.2 ENTWICKLUNG FRÜHER MATHEMATISCHER VORLÄUFERFERTIGKEITEN 10

2.2.1 Das Stufenmodell nach Fritz und Ricken (2008) 11

2.2.2 Das ZGV-Modell nach Krajewski (2013) 13

2.2.3 Bedeutung früher mathematischer Vorläuferfertigkeiten für den schulischen Kompetenzerwerb 15

2.3 ERWERB MATHEMATISCHER BASISKOMPETENZEN IN DER GRUNDSCHULZEIT 18

2.3.1 Pädagogisch-Psychologische Perspektive 19

2.3.1.1 Das Kompetenzstufenmodell nach Reiss (2004) 19

2.3.1.2 Bildungsstandards für den Primarbereich 21

2.3.1.3 Internationale Perspektive 22

2.3.2 Kognitiv-Neurowissenschaftliche Perspektive 25

2.3.3 Bedeutung mathematischer Basiskompetenzen für den weiteren Kompetenzerwerb 30

2.4 ERWERB MATHEMATISCHER KOMPETENZEN IN DER SEKUNDARSTUFE I 32

2.4.1 Der Prozess der Transition in die Sekundarstufe I 33

2.4.2 Erwerb mathematischer Kompetenzen in der Sekundarstufe I im Längsschnitt 36

2.4.2.1 Nationale Studien 36

2.4.2.2 Internationale Studien 43

2.4.3 Erwerb mathematischer Kompetenzen in der Sekundarstufe I im Querschnitt 50

2.5 ZUSAMMENFASSUNG 57

3 SCHWIERIGKEITEN IN DER ENTWICKLUNG UND IM ERWERB MATHEMATISCHER KOMPETENZEN 60

3.1 BEGRIFFLICHKEITEN 60

3.1.1 Schwierigkeiten im Lernen 60

3.1.2 Schwierigkeiten im Rechnen 64

3.1.2.1 Rechenschwierigkeiten 64

3.1.2.2 Rechenstörung 64

3.1.3 Begrifflichkeiten international 65

3.2 PRÄVALENZ UND KOMORBIDITÄTEN 66

3.3 ÄTIOLOGIE 69

3.3.1 Lern-/Rechenschwierigkeiten 69

3.3.2 Rechenstörung 72

3.4 DIAGNOSTIK 74

3.4.1 Hinweise und Symptome in Vorschule und Schule ... 75
3.4.2 Anamnese ... 78
3.4.3 Analyse der Mathematikleistungen ... 78
3.4.3.1 Formelle Verfahren ... 78
3.4.3.2 Informelle Verfahren ... 84

3.5 Grundsätze der Förderung ... 87

3.6 Zusammenfassung ... 89

4 Evidenzbasierte Förderung mathematischer Basiskompetenzen ... 91

4.1 Das Konzept Evidenzbasierung ... 91

4.2 Rahmenkonzept Responsives Handlungsmodell (RTI) ... 100
4.2.1 Entstehung und Entwicklung ... 102
4.2.2 Grundprinzipien ... 102
4.2.2.1 Mehrebenenprävention ... 103
4.2.2.2 Evidenzbasierte Praxis ... 103
4.2.2.3 Datenbasierte Entscheidung ... 104
4.2.2.4 Lernverlaufsdiagnostik ... 106
4.2.3 Wirksamkeit ... 108
4.2.4 Kritik ... 110

4.3 Evidenzbasierte Förderung mathematischer Basiskompetenzen in der Sek. I ... 111
4.3.1 Hinweise zur Auswertung des Forschungsstandes ... 111
4.3.2 Deutschsprachige standardisierte Förderprogramme ... 113
4.3.2.1 Identifikation (potentiell) evidenzbasierter standardisierter Programme ... 114
4.3.2.2 Analyse (potentiell) evidenzbasierter standardisierter Programme ... 118
4.3.2.3 Zusammenfassung und Forschungsdesiderata ... 123
4.3.3 Fördermethoden für den Bereich Mathematik ... 124
4.3.3.1 Identifikation (potentiell) evidenzbasierter Fördermethoden ... 125
4.3.3.2 Analyse (potentiell) evidenzbasierter Fördermethoden ... 128
4.3.3.3 Zusammenfassung und Forschungsdesiderata ... 149

4.4 Implementation von Fördermaßnahmen ... 150
4.4.1 Begriffliches und Konzeption ... 151
4.4.2 Bedeutsamkeit und Desiderata ... 153
4.4.3 Kriterien von Implementation ... 153
4.4.4 Zusammenfassung ... 155

5 Zusammenfassung der theoretischen Grundlagen ... 156

6 Grundlagen der Maßnahme zur Förderung mathematischer Basiskompetenzen zu Beginn der Sekundarstufe I ... 159

6.1 Ziele der Förderung ... 159

6.2 Entwicklungsprozess ... 160

6.3 Darstellung der Fördermaßnahme ... 161

6.3.1 Didaktische Prinzipien ... 162
6.3.2 Methodische Prinzipien ... 166

7 FORSCHUNGSZIEL, -FRAGEN UND HYPOTHESEN ... 176

7.1 FORSCHUNGSZIEL UND -FRAGEN ... 176

7.2 HYPOTHESEN ... 176

7.2.1 Hypothesenblock I: Wirksamkeit auf universeller Ebene ... 177
7.2.2 Hypothesenblock II: Wirksamkeit bei Schülerinnen und Schüler mit ungünstigen Lernvoraussetzungen ... 177
7.2.2.1 Geringes Vorwissen ... 179
7.2.2.2 Geringe kognitive Fähigkeiten ... 180
7.2.2.3 Geringe Lesekompetenz ... 181
7.2.2.4 Risikobehaftetes Sozialverhalten ... 182
7.2.2.5 Risikobehaftetes Lernverhalten ... 183
7.2.2.6 Sonderpädagogischer Unterstützungsbedarf im Bereich Lernen ... 184
7.2.3 Hypothesenblock III: Einfluss der Lernfreude und der Leistungsheterogenität der eigenen Klasse auf die Wirksamkeit ... 185
7.2.4 Hypothesenblock IV: Implementationsqualität der Maßnahme ... 186
7.2.4.1 Beurteilung der Schülerinnen und Schüler ... 186
7.2.4.2 Beurteilung der Interventionsleiterinnen und -leiter ... 187

8 METHODIK UND EVALUATION ... 188

8.1 FORSCHUNGSDESIGN ... 188

8.2 ERHEBUNGSINSTRUMENTE ... 192

8.2.1 Schülererhebung ... 192
8.2.1.1 Heidelberger Rechentest (HRT 1–4) ... 192
8.2.1.2 Salzburger Lese-Screening (SLS 1–4) ... 193
8.2.1.3 Culture Fair Intelligence Test (CFT 20-R) ... 194
8.2.1.4 Fragebogen zur Erfassung emotionaler und sozialer Schulerfahrungen (FEESS 3–4) ... 195
8.2.1.5 Implementationsfragebogen ... 196
8.2.2 InterventionsleiterInnen- und Lehrererhebung ... 198
8.2.2.1 Lehrereinschätzliste für Sozial- und Lernverhalten (LSL) ... 198
8.2.2.2 Implementationsfragebogen ... 199

8.3 STICHPROBENREKRUTIERUNG ... 201

8.4 UNTERSUCHUNGSDURCHFÜHRUNG ... 202

8.5 FORSCHUNGSETHIK ... 204

8.6 AUSWERTUNGSSTRATEGIEN ... 205

8.6.1 Überblick über die Auswertungsstrategien der Hypothesen ... 206
8.6.2 Beschreibung der Strategien und Verfahren ... 206
8.6.2.1 χ^2-Test ... 206
8.6.2.2 t-Test ... 207
8.6.2.3 Varianzanalyse mit Messwiederholung ... 208
8.6.2.4 Effektstärke ... 209
8.6.2.5 α-Fehler-Adjustierung ... 210
8.6.2.6 Poweranalyse ... 211

8.6.2.7 (Multiple lineare) Regression 211

8.7 Operationalisierung der Wirksamkeit 213

8.8 Datenpflege und -aufbereitung 214

9 Ergebnisse 216

9.1 Beschreibung der Stichprobe 216

9.2 Darstellung des Drop-Outs 224

9.3 Hypothesenblock I 225

9.4 Hypothesenblock II 228

9.4.1 Wirksamkeit bei Schülerinnen und Schülern mit einem geringen Vorwissen 229

9.4.2 Wirksamkeit bei Schülerinnen und Schülern mit geringen kognitiven Fähigkeiten 232

9.4.3 Wirksamkeit bei Schülerinnen und Schülern mit geringen Lesekompetenzen 235

9.4.4 Wirksamkeit bei Schülerinnen und Schülern mit einem risikobehafteten Sozialverhalten 239

9.4.5 Wirksamkeit bei Schülerinnen und Schülern mit einem risikobehafteten Lernverhalten 242

9.4.6 Wirksamkeit bei Schülerinnen und Schülern mit einem sonderpädagogischen Unterstützungsbedarf im Bereich Lernen 245

9.4.7 Ergebnisse der Hypothesenkomplexe I und II im Überblick 249

9.5 Hypothesenblock III 255

9.6 Hypothesenblock IV 257

9.6.1 Implementationsqualität aus Sicht der Schülerinnen und Schüler 257

9.6.2 Implementationsqualität aus Sicht der Interventionsleiterinnen und-leiter 260

9.6.3 Ergebnisse des Hypothesenkomplexes IV im Überblick 264

10 Interpretation und Diskussion 266

10.1 Interpretation und Diskussion der Ergebnisse 266

10.1.1 Hypothesenblock I und II 266

10.1.2 Hypothesenblock III 281

10.1.3 Hypothesenblock IV 282

10.2 Methodenkritische Reflexion 286

10.2.1 Entwicklung und Durchführung der Fördermaßnahme 286

10.2.2 Forschungsmethodik 290

11 Abschließende Bewertung und Ausblick 295

12 Literaturverzeichnis 299

13 Anhang 327

Abbildungsverzeichnis

Abbildung 1 Mathematische Kompetenzen im Entwicklungsverlauf 10

Abbildung 2 Unterscheidung zwischen primären und sekundären mathematischen Kompetenzen (Geary, 2000, S. 12) 11

Abbildung 3 Entwicklungsmodell der Zahl-Größen-Verknüpfung nach Krajewski (2013, S. 201) 14

Abbildung 4 Zusammenhang zwischen numerischer Kompetenz im Vorschulalter und mathematischer Intelligenz mit 23 Jahren (Stern, 2008, S. 196) 17

Abbildung 5 Überblick über die Bereiche der Bildungsstandards für den Primarbereich (Kultusministerkonferenz, 2004, S. 7) 21

Abbildung 6 Principles and Standards for School Mathematics (NCTM, 2000) 23

Abbildung 7 Ergebnisse des Early Numeracy Research Projects (Clarke, 2001) 24

Abbildung 8 Triple-Code-Modell nach Dehaene (1992) 26

Abbildung 9 Entwicklungsschritte der Rechenfähigkeit (Jacobs & Petermann, 2007) 27

Abbildung 10 Arbeitsgedächtnismodell nach Baddeley (Baddeley, 2010a, 2010b) (grau: Arbeitsgedächtniskomponenten) 28

Abbildung 11 Entwicklungsaufgaben in der Transition in die Sekundarstufe I (Griebel & Niesel, 2011, S. 208 ff) 34

Abbildung 12 Überblick über nationale Längsschnittstudien zum Erwerb mathematischer Kompetenzen in der Sekundarstufe I 37

Abbildung 13 Überblick über internationale Längsschnittstudien zum Erwerb mathematischer Kompetenzen in der Sekundarstufe I 44

Abbildung 14 Erwerb mathematischer Kompetenzen in der ORACLE-Replikationsstudie (Galton et al., 2002, S. 147) 47

Abbildung 15 Ergebnisse des Leverhulme Numeracy Research-Programms im Bereich „Methods of computation“ (Addition) (Brown et al., 2003) 48

Abbildung 16 Arithmetische Kompetenzen von 7- und 8-Klässlern (Mittelberg, 2004, S. 68) 54

Abbildung 17 Vergleich der Fachleistungen im Bereich Mathematik nach Beschulung (Hoffmann, 2009, S. 36) 55

Abbildung 18 Dreidimensionale Terminologie von Lernbeeinträchtigungen (Schröder, 2005, S. 88 f.; nach Bach, 1977) 62

Abbildung 19 Handlungsmodell bei Lernstörungen (Matthes, 2006, S. 26) 63

Abbildung 20 Geschlechterunterschiede im Risiko des Auftretens von LRS und Dyskalkulie sowie des gemeinsamen Auftretens (Hasselhorn & Schuchardt, 2006, S. 213) 68

Abbildung 21 Erklärungsmodelle von Lernschwierigkeiten (Heimlich, 2009, S. 35) ... 69

Abbildung 22 Mögliche Ursachen einer Dyskalkulie (Jacobs & Petermann, 2007, S. 16) 72

Abbildung 23 Standardisierte Testverfahren im Bereich Mathematik für die Sekundarstufe I 80

Abbildung 24 Differenzialdiagnostik Rechenstörung vs. Rechenschwierigkeit (Jacobs & Petermann, 2012, S. 96; modifiziert durch d. Verf.) 85

Abbildung 25 Aspekte der Evidenzbasierung (grau: nach Blumenthal & Mahlau, 2015, S. 409; weiß: nach Whitehurst, 2002) 93
Abbildung 26 Schritte in der Auswahl evidenzbasierter Verfahren (Freeman & Sugai, 2013, S. 10) 99
Abbildung 27 Respose to Intervention-Modell (Huber & Grosche, 2012, S. 314 erweitert nach Brown-Chidsey & Steege, 2010, S. 5; d. Verf.) 101
Abbildung 28 Formative Diagnostik der Leistungsentwicklung eines Kindes im Bereich Mathematik (Lembke & Stecker, 2007, S. 2) 105
Abbildung 29 Vorgehen und Gliederung der Kapitel 4.3.2 und 4.3.3 zur Auswertung des Forschungsstandes 112
Abbildung 30 Ergebnisse der Calcularis-Pilotstudie (Käser et al., 2013, S. 9)..... 122
Abbildung 31 Ergebnisse des Follow-Through-Projekts in den Domänen Lesen, Rechnen, Rechtschreiben und Sprache (W. C. Becker & Engelmann, 1996, S. 36 ff; Eigene Darstellung) 131
Abbildung 32 Vergleich der Direkten Instruktion und dem sog. Drill and Practice (Tournaki, 2003)..... 132
Abbildung 33 Ergebnisse der Untersuchung (Calhoon & Fuchs, 2003, S. 241)..... 137
Abbildung 34 Entwicklung nach der Intervention (Fuchs, Fuchs, Yazdian, et al., 2002, S. 577) 139
Abbildung 35 Wirksamkeit Kooperatives Lernen vs. Klassenunterricht in Klasse 3 .. 142
Abbildung 36 Wirksamkeit Kooperatives Lernen vs. Klassenunterricht in Klasse 4 .. 143
Abbildung 37 Akteure von Implementation (Reinmann-Rothmeier & Mandl, 1998, S. 308) 152
Abbildung 38 Einordnung der Maßnahme in die Klassifikation von Präventions- und Interventionsmaßnahmen (Durlak, 1997a, S. 3) 162
Abbildung 39 Vorwissensaktivierung durch den Advance-Organizer..... 169
Abbildung 40 Elemente der Arbeitsblätter a) Orientierungshilfe, b) Beispielaufgaben, c) Instruktion und Selbstkontrolle 172
Abbildung 41 Auszug aus einer peergestützten Übung 173
Abbildung 42 Die Rollenkarten der kooperativen Lerngruppen..... 173
Abbildung 43 Der Stadionplan mit Verhaltensfeedback und dem Lernverlauf..... 174
Abbildung 44 Schema des Forschungsdesigns 189
Abbildung 45 Entscheidungsbaum für t-Tests (Rasch et al., 2006b, S. 109)..... 211
Abbildung 46 Statistische Fundierung der kurzfristigen Wirksamkeit 213
Abbildung 47 Statistische Fundierung der langfristigen Wirksamkeit 214
Abbildung 48 Prä-Post-Entwicklung der Gesamtgruppe 226
Abbildung 49 Post-Follow-Up-I-Follow-Up-II-Entwicklung der Gesamtgruppe 227
Abbildung 50 Prä-Post-Entwicklung der Schülerinnen und Schüler mit einem geringen Vorwissen..... 229
Abbildung 51 Post-Follow-Up-I-Follow-Up-II-Entwicklung der Schülerinnen und Schüler mit einem geringen Vorwissen 231
Abbildung 52 Prä-Post-Entwicklung der Schülerinnen und Schüler mit geringen kognitiven Fähigkeiten..... 233

Abbildung 53 Post-Follow-Up-I-Follow-Up-II-Entwicklung der Schülerinnen und Schüler mit geringen kognitiven Fähigkeiten 234
Abbildung 54 Prä-Post-Entwicklung der Schülerinnen und Schüler mit geringen Lesekompetenzen 236
Abbildung 55 Post-Follow-Up-I-Follow-Up-II-Entwicklung der Schülerinnen und Schüler mit geringen Lesekompetenzen 237
Abbildung 56 Prä-Post-Entwicklung der Schülerinnen und Schüler mit einem risikobehafteten Sozialverhalten 239
Abbildung 57 Post-Follow-Up-I-Follow-Up-II-Entwicklung der Schülerinnen und Schüler mit einem risikobehafteten Sozialverhalten 241
Abbildung 58 Prä-Post-Entwicklung der Schülerinnen und Schüler mit einem risikobehafteten Lernverhalten 243
Abbildung 59 Post-Follow-Up-I-Follow-Up-II-Entwicklung der Schülerinnen und Schüler mit einem risikobehafteten Lernverhalten 244
Abbildung 60 Prä-Post-Entwicklung der Schülerinnen und Schüler mit einem sonderpädagogischen Unterstützungsbedarf im Bereich Lernen 246
Abbildung 61 Post-Follow-Up-I-Follow-Up-II-Entwicklung der Schülerinnen und Schüler mit einem sonderpädagogischen Unterstützungsbedarf im Bereich Lernen 247
Abbildung 62 Korrelation zwischen dem Prädiktor Lernfreude und dem Leistungszuwachs im HRT von der Prä- zur Posterhebung inklusive der Regressionsgeraden 257
Abbildung 63 Die durchschnittliche Zustimmung zu den Implementationsskalen aus Sicht der Schülerinnen und Schüler 258
Abbildung 64 Die durchschnittliche Zustimmung zu den Implementationsskalen aus Sicht der Interventionsleiterinnen und -leiter 260
Abbildung 65 Vergleich der Effektstärken des Interaktionseffekts in den Hypothesenkomplexen 1–7 281

Tabellenverzeichnis

Tabelle 1 Übersicht über Definitionen des Begriffs „Mathematische Basiskompetenzen“ 8

Tabelle 2 Kompetenzstufenmodell nach Fritz & Ricken (2008) mit Alters- bzw. Klassenangaben (Fritz, Ricken, Balzer, Leutner & Willmes, 2012) 12

Tabelle 3 Kompetenzstufenmodell (Reiss, 2004; Reiss & Winkelmann, 2008) 19

Tabelle 4 Inhaltsbezogene Kompetenzen der Bildungsstandards Mathematik für den Primarbereich (Kultusministerkonferenz, 2004, S. 9 ff.) 22

Tabelle 5 Detailangaben zu nationalen Längsschnittstudien zum Erwerb mathematischer Kompetenzen in der Sekundarstufe I 38

Tabelle 6 Korrelation zwischen den Mathematikleistung in den ersten sechs Schuljahren mit denen in der elften Klasse (Stern, 2013) 42

Tabelle 7 Detailangaben zu internationalen Längsschnittstudien zum Erwerb mathematischer Kompetenzen in der Sekundarstufe I 45

Tabelle 8 Detailangaben zu nationalen Querschnittstudien zum Erwerb mathematischer Kompetenzen in der Sekundarstufe I 51

Tabelle 9 Arten von Lernstörungen (Lauth et al., 2014, S. 18) 61

Tabelle 10 Prävalenz von Rechenstörungen (angelehnt an Jacobs & Petermann, 2007; erweiterte Zusammenstellung d. Verf.) 66

Tabelle 11 Anzeichen einer Rechenschwäche in der Vorschulzeit (vgl. Tröster, 2008, S. 353) 76

Tabelle 12 Kompetenzstufen nach Reiss (2004) bzw. nach Reiss und Winkelmann (2008) mit Hürden und Symptomen bei Rechenschwierigkeiten/-störungen 77

Tabelle 13 Detailangaben zu deutschsprachigen standardisierten Testverfahren im Bereich Mathematik für die Sekundarstufe I 81

Tabelle 14 Operationsübergreifende Fehler in (halb-)schriftlichen und mündlichen arithmetischen Rechenverfahren (Winter, 2010, S. 43) 86

Tabelle 15 Aspekte der Förderung bei Rechenschwierigkeiten (Kategorisierung nach Lorenz, 2014, S. 50) 88

Tabelle 16 Klassifikation evidenzbasierter Praxis (Council for Exceptional Children, 2014, S. 8 ff; Übers. d. Verf.) 94

Tabelle 17 Internet-Plattformen und Webseiten zur Auswahl evidenzbasierter Methoden und Maßnahmen 97

Tabelle 18 „Growth Rates“ im Bereich Mathematik (Hosp et al., 2007, S. 47; S. 79; S. 109) 106

Tabelle 19 Häufigkeit curriculumbasierter Messungen im RTI-Modell (Käter, Käter, Martenstein & Hillenbrand, 2016, S. 169) 107

Tabelle 20 Potenziell evidenzbasierte deutschsprachige standardisierte Förderprogramme im Bereich Mathematik für die Sekundarstufe I 116

Tabelle 21 Informationsquellen zu (potenziell) evidenzbasierten Fördermethoden im Bereich Mathematik 126

Tabelle 22 Vergleich der Effektstärken (d) der Methoden aus den recherchierten Metaanalysen 127

Tabelle 23 Schrittweise Umsetzung der Direkten Instruktion (nach Lebens & Lauth, 2014, S. 424) 129
Tabelle 24 Effektstärken der Treatments (d) (Fuchs, Fuchs, Hamlett, et al., 2002, S. 100) 138
Tabelle 25 Zuwachs je nach Anforderungsniveau (Shapiro et al., 2005, S. 23) 147
Tabelle 26 Implementationsvariablen (Petermann, 2015b, S. 123 ff.) 154
Tabelle 27 Ziele der Maßnahme 160
Tabelle 28 Zeitlicher Ablauf der Konzeption und Entwicklung der Maßnahme 161
Tabelle 29 Überblick über die Inhalte der Förderstunden 166
Tabelle 30 Überblick über die eingesetzten (evidenzbasierten) Methoden in den Unterrichtsphasen 167
Tabelle 31 Überblick über den Einsatz der CBM 175
Tabelle 32 Theoriebasierte Bestimmung der Schülerinnen und Schüler mit ungünstigen Lernvoraussetzungen 178
Tabelle 33 Detaillierter Überblick über das Forschungsdesign 190
Tabelle 34 Operationalisierung der Variablen der Hypothesen 190
Tabelle 35 Überblick über die Untertests „Rechenoperationen“ des HRT 1–4 193
Tabelle 36 Einordnung der LQ-Werte (Mayringer & Wimmer, 2003, S. 22) 194
Tabelle 37 IQ-Wertebereiche und Intelligenzniveaus 195
Tabelle 38 Konstruktion des Implementationsfragebogens für die Schülerinnen und Schüler 197
Tabelle 39 Überblick über die Skalen der Lehrereinschätzliste für Sozial- und Lernverhalten (Petermann & Petermann, 2013, S. 7) 198
Tabelle 40 Konstruktion des Implementationsfragebogens für die Interventionsleiterinnen und -leiter 200
Tabelle 41 Zeitlicher Ablauf der Pilotstudie 202
Tabelle 42 Zeitlicher Ablauf der Hauptstudie 204
Tabelle 43 Auswertungsstrategien zu den Hypothesen 206
Tabelle 44 Parametrische und zugehörige non-parametrischen Verfahren 207
Tabelle 45 Klassifikation von Cohen’s d und $\eta 2$ (Lenhard & Lenhard, 2016b) 210
Tabelle 46 Effektstärkenbasierte sukzessive α-Abwärtsadjustierung 210
Tabelle 47 Überblick über imputierte Werte für einzelne Subtests im HRT 1–4 (t_i–t_i) 215
Tabelle 48 Soziodemografische Beschreibung der Stichprobe inklusive fehlender Werte und χ^2-Tests 218
Tabelle 49 Testung der Untersuchungsgruppen auf Vortestunterschiede 221
Tabelle 50 Aufschlüsselung fehlender Schülerinnen und Schüler (Drop-Outs) zu den vier Messzeitpunkten Prä, Post, Follow-Up-I und Follow-Up-II 224
Tabelle 51 Zentrale Kennwerte zur Hypothese 1a – kurzfristige Wirksamkeit in der Gesamtgruppe 226
Tabelle 52 Zentrale Kennwerte zur Hypothese 1b – langfristige Wirksamkeit in der Gesamtgruppe 228

Tabelle 53 Zentrale Kennwerte zur Hypothese 2a – kurzfristige Wirksamkeit bei einem geringen Vorwissen 230
Tabelle 54 Zentrale Kennwerte zur Hypothese 2b – langfristige Wirksamkeit in der Gesamtgruppe 232
Tabelle 55 Zentrale Kennwerte zur Hypothese 3a – kurzfristige Wirksamkeit bei geringen kognitiven Fähigkeiten 233
Tabelle 56 Zentrale Kennwerte zur Hypothese 3b – langfristige Wirksamkeit bei geringen kognitiven Fähigkeiten 235
Tabelle 57 Zentrale Kennwerte zur Hypothese 4a – kurzfristige Wirksamkeit bei geringen Lesekompetenzen 237
Tabelle 58 Zentrale Kennwerte zur Hypothese 4b – langfristige Wirksamkeit bei geringen Lesekompetenzen 238
Tabelle 59 Zentrale Kennwerte zur Hypothese 5a – kurzfristige Wirksamkeit bei einem risikobehafteten Sozialverhalten 240
Tabelle 60 Zentrale Kennwerte zur Hypothese 5b – langfristige Wirksamkeit bei einem risikobehafteten Sozialverhalten 242
Tabelle 61 Zentrale Kennwerte zur Hypothese 6a – kurzfristige Wirksamkeit bei einem risikobehafteten Lernverhalten 243
Tabelle 62 Zentrale Kennwerte zur Hypothese 6b – langfristige Wirksamkeit bei einem risikobehafteten Lernverhalten 245
Tabelle 63 Zentrale Kennwerte zur Hypothese 7a – kurzfristige Wirksamkeit bei einem Unterstützungsbedarf im Bereich Lernen 246
Tabelle 64 Zentrale Kennwerte zur Hypothese 7b – langfristige Wirksamkeit bei einem Unterstützungsbedarf im Bereich Lernen 248
Tabelle 65 Ergebnisse der Hypothesenkomplexe I und II im tabellarischen Überblick 250
Tabelle 66 Analyse der Teststärke im Hypothesenblock I und II 253
Tabelle 67 Effektstärkenbasierte sukzessive α-Abwärtsadjustierung im Hypothesenblock I und II 253
Tabelle 68 Ergänzende Ergebnisse: Entwicklung in den sechs Skalen des HRT-Subtests „Rechenoperationen“ (Rohwerte, deskriptiv) 254
Tabelle 69 Modellübersicht I 256
Tabelle 70 Modellübersicht II 256
Tabelle 71 Ergebnisse der Skala „Akzeptanz“ des Implementationsfragebogens der Schülerinnen und Schüler auf Subskalenebene (deskriptiv) 259
Tabelle 72 Ergebnisse der Skala „Angemessenheit“ des Implementationsfragebogens der Schülerinnen und Schüler auf Subskalenebene (deskriptiv) 260
Tabelle 73 Ergebnisse der Skala „Akzeptanz“ des Implementationsfragebogens der Interventionsleiterinnen und -leiter (deskriptiv) 261
Tabelle 74 Ergebnisse der Skala „Angemessenheit“ des Implementationsfragebogens der Interventionsleiterinnen und -leiter (deskriptiv) 262
Tabelle 75 Ergebnisse der Skala „Machbarkeit“ des Implementationsfragebogens der Interventionsleiterinnen und -leiter (deskriptiv) 263
Tabelle 76 Ergebnisse des Hypothesenkomplexes IV im tabellarischen Überblick 265

Abkürzungsverzeichnis

ABKÜRZUNGEN IN DER THEORETISCHEN GRUNDLEGUNG

BELLA	Berliner Erhebung der Lernausgangslagen arbeitsrelevanter Basiskompetenzen von Schülerinnen und Schülern mit einem Förderbedarf
BIJU	Bildungsverläufe und psychosoziale Entwicklung im Jugendalter
CBM	Curriculumbasiertes Messen
IDEA	Individuals with Disabilities Education Improvement Act
IGS	Integrierte Gesamtschule
IQ	Intelligenzquotient
KESS	Kompetenzen und Einstellungen von Schülerinnen und Schülern (Studie)
LAU	Aspekte der Lernausgangslage und der Lernentwicklung (Studie)
LNRP	Levelhume Numeracy Research Programme
LOGIK	Longitudinalstudie zur Genese individueller Kompetenzen
LOSO	Longitudinal Research Secondary Education
MARKUS	Mathematik-Gesamterhebung Rheinland-Pfalz: Kompetenzen, Unterrichtsmerkmale, Schulkontext
OBS	Oberschule
ORACLE	Observational and Classroom-Learning Evaluation-Study
ORACLE-R	Observational and Classroom-Learning Evaluation-Study-Replication
PALMA	Projekt zur Analyse der Leistungsentwicklung in Mathematik
PISA	Programme for International Student Assessment
QuaSUM	Qualitätsuntersuchung an Schulen zum Unterricht in Mathematik
RTI	Response-to-Intervention-Modell / Responsives Handlungsmodell
SCHOLASTIK	Schulorganisierte Lernangebote und Sozialisation von Talenten, Interessen und Kompetenzen
Sek I	Sekundarstufe I
SEELS	Special Education Elementary Longitudinal Study
NELS:88	National Education Longitudinal Study of 1988
TIMSS	Trends in International Mathematics and Science Study (TIMS-Study)
ZGV-Modell	Zahl-Größen-Verknüpfungs-Modell

Abkürzungen in der empirischen Untersuchung

α	Signifikanzniveau
β	Regressionskoeffizient
AV	Abhängige Variable
d	Effektstärke (ES)
d_{korr}	Korrigierte Effektstärke
EG	Experimentalgruppe
g	Gruppe
gkF	Geringe kognitive Fähigkeiten
gLk	Geringe Lesekompetenz
gV	Geringes Vorwissen
K-S	Kolmogorov-Smirnov-Test
KG	Kontrollgruppe
N/n	Stichprobe (N = Gesamtstichprobe; n = Teilstichprobe)
η^2	Eta-Quadrat
p_t	Signifikanzwert des t-Tests
p_u	Signifikanzwert des Mann-Whitney-U-Tests
$p_{Mauchly}$	Signifikanzwert des Mauchly-W-Tests
r	Korrelationskoeffizient
R^2	Bestimmtheitsmaß der Regressionsanalyse
rL	Risikobehaftetes Lernverhalten
rS	Risikobehaftetes Sozialverhalten
SD	Standardabweichung
SPU	Sonderpädagogischer Unterstützungsbedarf (Lernen)
t	Messzeitpunkt
UV	Unabhängige Variable
χ^2	Chi-Quadrat

1 Einleitung

„Wem mathematische Grundkenntnisse fehlen, der muss viel glauben – den Steuerbescheid oder das beim Barzahlen erhaltene Wechselgeld."
(Prof. Dr. Albrecht Beutelspacher im Interview mit DIE ZEIT am 13. März 2016)

Die Studie „Bürgerkompetenz Rechnen" unter Schirmherrschaft Johanna Wankas zeigt im Jahr 2013, dass nur jede bzw. jeder zweite Deutsche zwischen 18 und 65 Jahren ausrechnen kann, wie viele Eimer Farbe sie oder er kaufen muss, um einen fensterlosen Kellerraum mit gegebenen Abmessungen zu streichen – und dies sogar trotz Nutzung eines Taschenrechners (Stiftung Rechnen, 2013). Viele Erwachsene scheitern im Alltag schon beim Lösen solcher eher einfachen mathematischen Aufgaben. Sie sind in alltäglichen mathematischen Herausforderungen auf Grund fehlender Kompetenzen kaum handlungsfähig (Basendowski, 2013). Es fehlt ihnen an „mathematical literacy", die in der PISA-Studie als Voraussetzung für eine mündige gesellschaftliche Teilhabe gesehen wird (OECD, 2013, S. 17). In Fragen der Ursachen für diese geringe Handlungsfähigkeit kommt Norbert Lossau, Journalist des o. g. Interviews, pointiert zum Schluss, dass vor allem „[die] Schule entscheidet, mit wie viel mathematischer Bildung ein Mensch durchs Leben geht".

Ein nicht zu missachtender Teil der Schülerinnen und Schüler zeigt in der Schule massive Schwierigkeiten im Erwerb dieser für den Alltag benötigten grundlegenden mathematischen Kompetenzen. Ein Blick auf die mathematische Bildung in Schulen ist ernüchternd. Die Defizite erkannten viele Politikerinnen und Politiker sowie (Bildungs-)Wissenschaftlerinnen und -Wissenschaftler spätestens durch den „Pisa-Schock" im Jahr 2000. Trotz des unmittelbar verabschiedeten 7-Punkte-Handlungsplans (Kultusministerkonferenz, 2002) weisen aber auch in den Folgejahren zahlreiche nationale Befunde auf geringe mathematische Kompetenzen vieler Schülerinnen und Schüler in der Sekundarstufe I hin. Zu erwähnen sind exemplarisch die groß angelegte MARKUS-Studie (Balzer & Jäger, 2001; Helmke et al., 2000, 2002), die QuaSum-Studie (Lehmann, Peek, et al., 2001) sowie die Untersuchungen von Mittelberg (2004), Moser Opitz (2005) und Humbach (2008).

Die PISA-Studien rücken die Diskussion um gute mathematische Bildung und schulischen Kompetenzerwerb in regelmäßigen Abständen immer wieder in das Zentrum der bildungspolitischen Diskussion. Doch zeigen die aktuellen PISA-Ergebnisse aus dem Jahr 2015, dass sich immer noch 17,2 % der 15-jährigen Schülerinnen und Schüler unter oder auf der Kompetenzstufe I (von VI) befindet (Hammer et al., 2016, S. 233). Ihr Mathematikverständnis entspricht allenfalls dem Niveau der vierten Jahrgangsstufe. Das bedeutet, dass ca. „jeder sechste Jugendliche in Deutschland die Mindestanforderungen für ein anschlussfähiges mathematisches Verständnis nicht erreicht und erhebliche Probleme haben dürfte, einen Ausbildungsplatz zu finden beziehungsweise eine Ausbildung erfolgreich abzuschließen sowie anspruchsvollere mathematische Anforderungen im Alltag zu bewältigen" (Prenzel et al., 2012, S. 75). Im Vergleich zur Pisa-Studie aus dem Jahr 2012 ist die Zahl der Schülerinnen und Schüler auf der ersten Kompetenzstufe zwar um 0,5 % gesunken, doch kann dies eher als „kosmetischer Effekt" betrachtet werden (Lossau, 2016, S. 19). Die Befunde sind alarmierend und lassen bisher keine zuversichtlichen Prognosen zu.

Die Entwicklung und der Erwerb mathematischer Kompetenzen verläuft dabei kumulativ (Humbach, 2008). In der Grundschulmathematik, die die Grundlage für den weiteren schulischen Kompetenzerwerb darstellt, sind das Verständnis und die Anwendungen arithmetischer Rechenoperationen erklärtes Ziel. Arithmetische Kompetenzen in Form der Grundrechenarten und des Stellenwertsystems werden im Rahmen dieser Forschungsarbeit unter dem Begriff der „mathematischen Basiskompetenzen" gefasst. Mathematische Schwierigkeiten in der Sekundarstufe I können eine Folge von Rückständen im Primarbereich sein. Sind die mathematischen Voraussetzungen im Übergang nicht gegeben, sind Probleme in der Sekundarstufe I weitgehend vorprogrammiert. Rückstände führen dann zu Anschlussschwierigkeiten im Mathematikunterricht der Sekundarstufe I, der gerade für schwache Schülerinnen und Schüler durch das ungewohnte Fachlehrerprinzip und eine divergierende Unterrichtsmethodik und -organisation erschwerend gekennzeichnet ist (Van Ophuysen & Wendt, 2009).

Studien zeigen, dass in der Sekundarstufe I gerade im Lösen arithmetischer Aufgaben oftmals noch erhebliche Defizite bestehen. Mittelberg (2004, S. 68) zeigt zum Beispiel, dass die Aufgabe „15.000 – 15 =" von 19,1 % der von ihm untersuchten Hauptschülerinnen und -schüler in Klasse sieben und acht *nicht* richtig gelöst wird. Moser Opitz (2005, S. 123) bestätigt diese Befunde. In ihrer Studie lösen die Aufgabe „12 : 4 =" lediglich 19,1 % der rechenschwachen Fünftklässlerinnen und -klässler, jedoch auch nur 51,1 % der durchschnittlich mathematisch begabten Schülerinnen und Schüler. Im achten Schuljahr beträgt die Lösungsquote bei dieser Aufgabe 31,8 % (rechenschwach) bzw. 63,8 % (durchschnittlich mathematisch begabt). Dabei gehören diese Aufgaben zum Anforderungsbereich der Primarstufe.

Neben Schwierigkeiten im Lösen arithmetischer Operationen ist auch das Wissen um die dahinterliegenden Konzepte oftmals nicht vollständig ausgebildet. Das bedeutet, dass die Schülerinnen und Schüler einer Sachsituation keine angemessene Rechenoperation zuordnen können. Als stellvertretend steht dafür eine Studie von Ehlert und Kollegen (2013). Die Sachaufgabe „Der Tierpfleger Paul verteilt 96 Fische an die Pinguine. Jeder Pinguin bekommt 12 Fische. Wie viele Pinguine hat er gefüttert?" lösen 10,5 % der untersuchten Hauptschülerinnen und -schüler mit der Lösung „96 + 12 =". Weitere 36,84 % rechnen „96 – 12 =" und 5,26 % der Schülerinnen und Schüler gar „12 : 96 =". „Dieses Beispiel zeigt, dass weniger fehlendes Textverstehen Grund für die fehlerhafte Bearbeitung der Aufgabe ist als [auch; Anm. d. Verf.] das fehlende Konzeptwissen um die Operation der Division" (Ehlert et al., 2013, S. 256). Dabei gehen diese mathematischen Schwierigkeiten nur bei ca. 2–6 % der Schülerinnen und Schüler mit einer gravierenden Rechenstörung im Sinne einer Dyskalkulie einher. Bei einem Großteil der Schülerinnen und Schüler handelt es sich vielmehr um allgemeine Rechenschwierigkeiten.

In Fragen mathematischer Bildung erweisen sich die Haupt- und Gesamtschulen als größte Herausforderung, da die Schülerinnen und Schüler in diesen Schulformen die größten Schwierigkeiten und Defizite im mathematischen Kompetenzerwerb zeigen (Bos & Gröhlich, 2010; Lehmann, Husfeldt & Peek, 2001; Van Ophuysen & Wendt, 2009; Pekrun et al., 2006). Es ist anzunehmen, dass gerade in diesen Schulformen das bereits vorhandene Spektrum an mathematischen Kompetenzen mit dem weiteren Ausbau des inklusiven Schulsystems steigen wird (Schels, Janz & Hintermair, 2015). Dies geht dann wiederum einher mit größeren Anforderungen der Schülerinnen und Schüler (Special Needs) und einer höheren Belastung der Lehrkräfte der allgemeinen Schule (Schels et al., 2015). Den bisherigen Befunden zufolge ist der Handlungsbedarf immens.

In Fragen der Förderung mathematischer Basiskompetenzen gerade in inklusiven Lerngruppen wird zugleich die Forderung nach Evidenzbasierung lauter. *Evidenzbasierung* bedeutet hier, dass in Unterrichts- und Fördersettings Maßnahmen und Methoden zum Einsatz kommen, deren Wirksamkeit durch nachvollziehbare wissenschaftliche Forschung bestätigt ist und deren Einsatz sich im betrachteten Einzelfall aus Sicht der beteiligten Akteure (z. B. der Lehrkraft) auch eignet (Hillenbrand, 2015, S. 314). Das pädagogisch-psychologische Konzept der Evidenzbasierung stammt ursprünglich aus den Vereinigten Staaten und genießt in Deutschland aktuell große Aufmerksamkeit. Ungefähr seit dem Jahr 2010 gehört der Ansatz „in der Wissenschaft, aber auch in der Kultusbürokratie und in Praktikerkreisen [...] zu den meist diskutierten" (Kuhl & Hecht, 2014, S. 409).

In den Vereinigten Staaten gibt es bereits umfangreiche Internetplattformen, auf denen Lehrkräfte Informationen zur evidenzbasierten Förderung bei diversen spezifischen Problemlagen einholen können. Da das Konzept in Deutschland noch Einzug hält, steht ein vergleichbarer, ökonomischer und nutzerorientierter Zugang zu entsprechenden Informationen für Sekundarstufen-I-Lehrkräfte noch aus. Dennoch gibt es Möglichkeiten zur Identifizierung evidenzbasierter Fördermethoden. Als geeignet erweist sich zum Beispiel der Handlungsleitfaden von Freeman und Sugai (2013), der von Hillenbrand (2015) ins Deutsche übersetzt worden ist. Er besteht aus sechs Schritten und führt u. a. über eine genaue Beschreibung des Problems und der Bestimmung grundsätzlicher Handlungsoptionen hin zu einer Auswahl geeigneter und passender Förderverfahren und -methoden.

Die vorliegende Forschungsarbeit versucht, den – aus wissenschaftlicher und schulpraktischer Sicht – offensichtlichen Bedarf an Förderung mathematischer Basiskompetenzen in der Sekundarstufe I systematisch zu bedienen und greift dabei auf das Konzept der Evidenzbasierung zurück. Vorgestellt wird eine vierwöchige Maßnahme zur Förderung mathematischer Basiskompetenzen zu Beginn der Sekundarstufe I in inklusiven Settings, die sich durch die Implementation evidenzbasierter Fördermethoden kennzeichnet. Die Maßnahme verortet sich dabei auf der universellen Ebene im „Responsiven Handlungsmodell", das ein wissenschaftlich evaluiertes Rahmenmodell für inklusive Bildung darstellt (Brown-Chidsey & Steege, 2010; Ennemoser, 2014; Fuchs & Fuchs, 2006; Huber, Grosche & Schütterle, 2013).

Forschungsziel dieser Arbeit ist es dabei, die kurz- und langfristige Wirksamkeit der Maßnahme zu evaluieren, mögliche Prädiktoren der Wirksamkeit zu identifizieren und die Implementationsqualität zu analysieren. Damit wird ein Beitrag zu Möglichkeiten evidenzbasierter Förderung mathematischer Basiskompetenzen in der Sekundarstufe I geleistet. Die Intervention soll künftig die Anschlussfähigkeit der Schülerinnen und Schüler nach der Transition an den mathematischen Kompetenzerwerb in der Sekundarstufe I gewährleisten.

Die Forschungsarbeit verfolgt folgenden Aufbau:

Im zweiten Kapitel werden in einem eher kurzen Abriss die Entwicklung mathematischer Vorläuferfertigkeiten und der Erwerb mathematischer Basiskompetenzen in der Grundschulzeit beschrieben. Schwerpunkt des Kapitels bildet dabei der Erwerb mathematischer Kompetenzen in der Sekundarstufe I. Es wird dabei der Prozess der Transition beschrieben und umfassend sowohl nationale und internationale Querschnitt- und Längsschnittuntersuchungen zum Erwerb mathematischer Kompetenzen vorgestellt.

Das darauffolgende Kapitel 3 beschäftigt sich mit Schwierigkeiten und Störungen in der Entwicklung und im Erwerb mathematischer Kompetenzen. Neben Begrifflichkeiten, Prävalenzangaben und Komorbiditäten werden die Ätiologie, Diagnostik und Grundsätze der Förderung dargestellt. Mit Blick auf die Ausrichtung der Forschungsarbeit auf inklusive Settings thematisiert das Kapitel nicht nur Schülerinnen und Schüler mit massiven Rechenstörungen (Dyskalkulie), sondern wirft durchgängig einen Blick auf diejenigen mit allgemeinen Rechenschwierigkeiten sowie auf Schülerinnen und Schüler in diversen Risikolagen für schulisches Lernen.

Den Schwerpunkt der theoretischen Grundlagen bildet das Kapitel 4 zur evidenzbasierten Förderung mathematischer Basiskompetenzen. Neben der Darstellung grundlegender Informationen zum Konzept der Evidenzbasierung und ihrer Verankerung im „Responsiven Handlungsmodell“ wird in einer Auswertung des aktuellen Forschungsstandes dargelegt, welche Methoden und standardisierten Förderprogramme sich zur evidenzbasierten Förderung mathematischer Basiskompetenzen in der Sekundarstufe I grundsätzlich anbieten bzw. aktuell schon vorhanden sind. Das Kapitel 4.4 thematisiert abschließend die Bedeutsamkeit der strukturierten und präzisen Implementation von Förderungkonzepten oder -programmen.

Nach der Zusammenfassung der theoretischen Grundlagen in Kapitel 5 wird die hier entwickelte Maßnahme zur Förderung mathematischer Kompetenzen zu Beginn der Sekundarstufe I vorgestellt (Kapitel 6). Neben den Zielen der Fördermaßnahme und ihrem Entwicklungsprozess werden ihre didaktischen und methodischen Prinzipien dargelegt.

Der empirische Abschnitt dieser Forschungsarbeit beginnt in Kapitel 7.1 mit der Formulierung des Forschungsziels und drei untergeordneten Forschungsfragen. Es leiten sich davon 15 Hypothesenkomplexe ab, die sich mit der kurz- und langfristigen Wirksamkeit, möglichen Prädiktoren und der Implementationsqualität beschäftigen.

Im Anschluss erfolgt die Darstellung des Forschungsdesigns, der eingesetzten Messinstrumente, der rekrutierten Stichprobe und der Untersuchungsdurchführung. Zudem werden Aspekte zur Forschungsethik, die Auswertungsstrategien sowie Hinweise zur Datenpflege und -aufbereitung dargelegt. Die Ergebnisse werden anschließend hypothesengeleitet ausgewertet und im Kontext aktueller Forschung diskutiert. Eine sich anschließende Methodenkritik betrachtet sowohl die Entwicklung und Durchführung der Fördermaßnahme als auch den forschungsmethodischen Zugang. Eine abschließende Bewertung und ein Ausblick runden die Arbeit im Kapitel 11 ab.

2 Entwicklung und Erwerb mathematischer Kompetenzen

2.1 Begrifflichkeiten

Im sogenannten „Zeitalter nach Pisa“ (Müller-Ruckwitt, 2008, S. 13) ist der Kompetenzbegriff in das Zentrum bildungspolitischer Aufmerksamkeit gerückt. Begriffe wie Schlüsselkompetenz, Kernkompetenz, Basiskompetenz, Alltagskompetenz, Methodenkompetenz, personale Kompetenz, soziale Kompetenz, Fachkompetenz, Kompetenzkonzept, Kompetenzmodell oder Medienkompetenz zeigen einen Ausschnitt des mittlerweile weit verbreiteten und häufig verwendeten Kompetenzbegriffs. So waren laut Deutscher Bibliothek in Frankfurt 70 % der Bücher, die den Begriff „Kompetenz“ im Titel tragen, alleine in den sieben Jahren von 1997 bis 2004 erschienen (Huber, 2004, S. 15). Sogleich erstarkt die Kritik an einer inflationären und unreflektierten Verwendung des Kompetenzbegriffs. Krautz (2009) schreibt von einer ökonomisierten Bildung, in der Kompetenzformulierungen dazu dienen, Schülerinnen und Schüler für das Zeitalter der Globalisierung vorzubereiten und dienstleistungsfähig zu machen. Der wesentliche Auftrag von Bildung, dass Schülerinnen und Schüler die Welt als Vermittlungsgegenstand reduziert, altersgemäß und bewältigbar kennenlernen, würde dabei oftmals vergessen werden. Die darüber hinaus an vielen Stellen geäußerte Kritik am Kompetenzbegriff richtet sich u. a. an dessen undefinierten und unscharfen Verwendung (Huber, 2004, S. 22; Müller-Ruckwitt, 2008, S. 31).

Diese Diskussion fordert, den im Rahmen dieser Arbeit verwendeten Terminus der mathematischen Basiskompetenz zu definieren und von anderen, oftmals parallel verwendeten Begriffen, abzugrenzen. Die Annäherung erfolgt in einem Dreischritt:

1. Was ist eine Kompetenz?
2. Was sind mathematische Kompetenzen?
3. Was sind mathematische Basiskompetenzen?

KOMPETENZ

Der Begriff „Kompetenz“ (von lat. *competentia*) war ursprünglich ein Synonym für „Recht auf Einkünfte“ und galt erst im 19. Jahrhundert als Abstraktum des Adjektivs „kompetent,“ was „zuständig“ bedeutete (Kluge, 2012). Heute wird der Begriff „Kompetenz“ mit „Sachverstand und Fähigkeiten“, im rechtssprachlichen Kontext noch mit „Zuständigkeit“ gleichgesetzt (DUDEN, 2013). Wer kompetent ist, ist etymologisch sachverständig und befähigt.

Im akademischen und bildungssprachlichen Kontext geht es bei Kompetenzen, besonders gegenüber früheren Definitionen, stets um Handlungsfähigkeit – um die Anwendungsfähigkeit von Wissen und Können (Sälzer, Reiss, Schiepe-Tiska, Prenzel & Heinze, 2013, S. 51). Schülerinnen und Schüler sollen nicht lediglich „totes Wissen“ (Meyer & Klapper, 2006, S. 90) kontextlos anreichern. Eine mittlerweile als „Referenzzitat“ (Klieme, 2004, S. 12) häufig rezipierte Definition des Begriffs Kompetenz stammt von Weinert (2001, S. 27).

Kompetenzen sind demnach

> „[...] die bei Individuen verfügbaren oder durch sie erlernbaren kognitiven Fähigkeiten und Fertigkeiten, um bestimmte Probleme zu lösen, sowie die damit verbundenen motivationalen, volitionalen und sozialen Bereitschaften und Fähigkeiten, um die Problemlösungen in variablen Situationen erfolgreich und verantwortungsvoll nutzen zu können." (Weinert, 2001, S. 27)

Die Definition nach Weinert (2001) zeigt, dass Kompetenzen erst dann zur Problemlösung nützen, wenn auch die Motivation und Volition dafür entsprechend vorhanden sind. Diese Definition ist eine übergeordnete, auf schulische Domänen bezogene, eher fachunspezifische Begriffsbestimmung. Was heißt es nun, mathematisch kompetent zu sein? Die Mathematikdidaktik präzisiert dafür eine spezifische Definition.

Mathematische Kompetenzen

In der Literatur finden sich diverse Beschreibungen und Definitionen des Begriffs der mathematischen Kompetenz. Allgemeine Anerkennung kommt dabei der Definition des PISA-Konsortiums zu, das „mathematische Kompetenz" mit dem englischsprachigen Begriff „mathematical literacy" gleichsetzt. „Literacy" bedeutet ins Deutsche übersetzt „Bildung" oder „Kenntnis" – wird jedoch oftmals auch mit „Lese- und Schreibfähigkeit" in Verbindung gebracht.

> „Mathematical Literacy: An individual's capacity to formulate, employ, and interpret mathematics in a variety of contexts. It includes reasoning mathematically and using mathematical concepts, procedures, facts, and tools to describe, explain, and predict phenomena. It assists individuals to recognise the role that mathematics plays in the world and to make the well-founded judgments and decisions needed by constructive, engaged and reflective citizens." (OECD, 2013, S. 17)

Trotz der in Deutschland konzeptionellen Redigierung des Begriffs „Mathematische Grundbildung", wie PISA 2000 ihn noch verwendete, zum nun aktuellen Begriff „Mathematische Kompetenz", besteht Kritik an der von PISA aufgestellten Definition. Neben der definitorischen Unschärfe des Begriffs (Müller-Ruckwitt, 2008, S. 31) richtet sich die Kritik an die OECD, die sich zumutet, gesellschaftlich normativ festzulegen, welche Fähigkeit jemand bedarf, um Anforderungen des Lebens bewältigen zu können (Mugerauer, 2012, S. 48 ff.; Müller-Ruckwitt, 2008, S. 43).

Die angeführte Definition der mathematischen Kompetenz nach der OECD (2013) betont auch für diesen speziellen Gegenstandsbereich die Anwendungsorientierung mit dem Ziel, Wissen flexibel zur Lösung innermathematischer und alltäglicher Probleme im persönlichen, beruflichen und gesellschaftlichen Kontext anwenden zu können (Sälzer et al., 2013, S. 50).

Aktuell liegen in Deutschland verschiedene Konzepte zur Darstellung und Strukturierung von mathematischen Kompetenzen in der schulischen Bildung vor. In der Darstellung dieser verschiedenen Kompetenzkonzepte und -modelle muss jedoch unterschieden werden zwischen

a) der *Struktur* eines Kompetenzbereichs,
b) der *Graduierung* einzelner Kompetenzen in Niveaus bzw. Stufen und
c) der Darstellung *des Erwerbs und* der *Entwicklung von Kompetenzen* im Zeitverlauf (Pant, Böhme & Köller, 2012, S. 50).

Unter Punkt a) kann zum Beispiel die Strukturierung der allgemeinen mathematischen Kompetenzen in den Bildungsstandards im Fach Mathematik für die Primarstufe in die fünf Kompetenzbereiche Problemlösen, Argumentieren, Modellieren, Kommunizieren und Darstellen von Mathematik verstanden werden (vgl. Kultusministerkonferenz, 2004). Punkt b) meint zum Beispiel die weitere Differenzierung des inhaltsbezogenen mathematischen Kompetenzbereichs „Zahlen und Operationen" in den Bildungsstandards im Fach Mathematik für die Primarstufe in fünf Niveaustufen (vgl. Reiss, Roppelt, Haag, Pant, & Köller, 2012, S. 76) und Punkt c) z. B. die Darstellung der Entwicklung mathematischer Vorläuferfertigkeiten (z. B. Fritz & Ricken, 2008).

Es liegen derzeit Untersuchungen vor, in denen mathematische Kompetenzen für eine bestimmte Alters- oder Klassenstufe modelliert werden: Das Programme for International Student Assessment (PISA) im Jahr 2012 bildet im Zuge seiner Erhebung mathematischer Kompetenzen ausschließlich 15-jähriger Schülerinnen und Schüler fünf Kompetenzstufen für eben diese Zielgruppe ab (Sälzer et al., 2013, S. 61). Auch im Rahmen der Bildungsstandards für den mittleren Schulabschluss (Kultusministerkonferenz, 2003) werden Kompetenzstufen ausschließlich für die zehnte Jahrgangsstufe ausgewiesen. Im Rahmen der TIMS-Studie werden mathematische Kompetenzstufen für die Zielgruppe der Viertklässlerinnen und Viertklässler entwickelt (Selter, Walther, Wessels & Wendt, 2012, S. 102). Neben den alters- und klassenstufenspezifischen Kompetenzmodellen (Fall a) gibt es Kompetenzstufenmodelle für ein spezielles mathematisches Themenfeld (Fall b), wie zum Beispiel das dreistufige Kompetenzmodell von Reiss und Kollegen (2002) zum Themenbereich „Beweisen und Begründen" für die gymnasialen Jahrgangsstufen sieben und acht. Diese Modelle ermöglichen es jedoch nicht, Erwerbsprozesse über mehrere Schuljahre abzubilden.

Mathematisches Lernen verläuft im Erwerbskontext kumulativ (Humbach, 2008). Inhalte bauen aufeinander auf, Wissen wird verknüpft und die Anwendungskompetenz ausgebaut. Aus diesem Grund gibt es den Begriff der mathematischen *Basis*kompetenz, der versucht, einen Bereich des Erwerbs abzugrenzen und seine Relevanz im weiteren Kompetenzerwerb herauszustellen.

Mathematische Basiskompetenzen

Gerade für die weiterführende Mathematik sind bestimmte Grundvorstellungen „notwendige Voraussetzung" (Ehlert et al., 2013, S. 239). Diese Voraussetzungen finden sich unter diversen Sammelbegriffen wie „Basiskompetenz" (zum Beispiel Reiss & Winkelmann, 2008; Schneider, Küspert, & Krajewski, 2013), „Basisstoff" (Deutscher, Prediger & Selter, 2013; Freesemann, 2014), „Grundfertigkeiten und -vorstellungen" (Leiß & Blum, 2006) oder auch „Basiskonzepte" (Ehlert et al., 2013) wieder. Die Begrifflichkeiten werden dabei oftmals synonym und parallel nebeneinander verwendet, obschon sie in ihrer inhaltlichen Ausgestaltung in Teilen divergieren. Mit der Verwendung des Begriffs wird grundsätzlich versucht, eine Phase bzw. eine Stufe im Erwerb mathematischer Kompetenzen zeitlich und inhaltlich zu beschreiben und von vorhergehenden sowie darauf aufbauenden Entwicklungs- bzw. Erwerbsschritten abzugrenzen.

Die Tabelle 1 zeigt in chronologischer Reihenfolge, wie unterschiedlich und vielfältig der Begriff „mathematische Basiskompetenzen" in der deutschsprachigen Literatur verwendet wird.

TABELLE 1 ÜBERSICHT ÜBER DEFINITIONEN DES BEGRIFFS „MATHEMATISCHE BASISKOMPETENZEN“

Autor/en	Begriff	Definition
Leiß & Blum (2006)	Basis der Kompetenzen	• elementares Grundwissen im flexiblen und flüssigen Umgang mit Zahlen, Größen und geometrischen Objekten • inhaltliche Vorstellung von mathematischen Begriffen und Verfahren
Reiss & Winkelmann (2008)	Mathematische Basiskompetenzen	• technische Grundlagen
Sinner & Kuhl (2010)	Mathematischen Basiskompetenzen	• numerische Basisfertigkeiten • Anzahlkonzept • Kardinalzahlverständnis • Verständnis von Anzahlrelationen
Drüke-Noe & Kollegen (2011)	Mathematische Basiskompetenzen	• mathematische Kompetenzen, über die alle Schülerinnen und Schüler aller Bildungsgänge am Ende der allgemeinen Schulpflicht verfügen
De Vries (2013)	Mathematische Basiskompetenzen	• pränumerische Kompetenzen • Vorläuferfertigkeiten
Ehlert, Fritz, Arndt & Leutner (2013)	Arithmetische Basiskompetenzen	• Beherrschung der Grundrechenarten und des Stellenwertsystems • Modellieren von Textaufgaben
Ennemoser & Krajewski (2013)	Mathematische Basiskompetenzen	• grundlegende Kompetenzen in der Zahl-Größen-Verknüpfung und in der Anwendung des Konventions- und Regelwissens
Freesemann (2014)	Basisstoff	• Verständnis des Dezimalsystems und der Grundoperationen • Zählen in Schritten und der Umgang mit Sachaufgaben
Lambert (2015)	Mathematische Basiskompetenzen	• Zählen • Zahlenwissen • Transformation von Zahlen • Schätzen • Zahlenmuster

Sinner und Kuhl (2010) sind der Auffassung, dass die Entwicklung der „mathematischen Basiskompetenzen“ bereits ab der Geburt beginnt, beim Schuleintritt jedoch noch nicht abgeschlossen sein müsse. Die Autoren fassen unter „mathematischen Basiskompetenzen“ Entwicklungsbereiche, die in anderen Darstellungen jedoch eher unter den Begriff der mathematischen Vorläuferfertigkeiten fallen (Gaupp, Zoelch & Schumann-Hengsteler, 2004; Schneider et al., 2013). Eine ähnliche inhaltliche Ausrichtung verfolgt auch de Vries (2013). Dabei wird der Begriff der Basiskompetenzen nicht nur auf frühe Entwicklungsphasen begrenzt, sondern auch für den Sekundarstufenbereich als tragfähig angenommen (Ennemoser & Krajewski, 2013, S. 226). Dahingegen fassen Drüke-Noe und Kollegen (2011, S. 8) mathematische Basiskompetenzen sehr weit als diejenigen Kompetenzen, die am Ende der allgemeinen Schulpflicht vor dem Übergang in die berufliche Ausbildung gefestigt und dauerhaft vorliegen müssen.

Leiß und Blum (2006, S. 34) sind der Auffassung, dass zwei Gesichtspunkte die „Basis“ ausmachen, auf der weiterführende Kompetenzen fußen: Zum einen ein elementares Grundwissen

bezüglich des flexiblen und flüssigen Umgangs mit Zahlen und Größen und mit geometrischen Figuren (Leitfrage: „Wie geht das?“), zum anderen die inhaltliche Vorstellung von mathematischen Begriffen und Verfahren (Leitfrage: „Was bedeutet das?“). Reiss und Winkelmann (2008) bezeichnen „Technische Grundlagen“ als mathematische Basiskompetenzen. Diese sind in der ersten Stufe ihres schulischen Kompetenzstufenmodells verortet (s. Kapitel 2.3.1.1). Eine ähnliche, jedoch etwas differenziertere inhaltliche Auffassung vertritt Freesemann (2014), die unter dem Begriff „Basisstoff“ das Verständnis des Dezimalsystems und der Grundoperationen, das Zählen in Schritten und den Umgang mit Sachaufgaben fasst. Ehlert und Kollegen (2013) nehmen die Beherrschung der Grundrechenarten, des Stellenwertsystems und das Modellieren von Textaufgaben hinzu. Ennemoser und Krajewski (2013, S. 226) fassen mathematische Basiskompetenzen als das Verständnis des mathematischen Sprach- und Symbolsystems auf und unterscheiden zwei Arten basaler Kompetenzen: grundlegende Kompetenzen in der Zahl-Größen-Verknüpfung und in der Anwendung des Konventions- und Regelwissens. Dass sich diese mathematischen Basiskompetenzen nach der Grundschulzeit *nicht* weiterentwickeln, konnte in jüngster Zeit widerlegt werden (Schneider et al., 2013, S. 116).

Die angeführten Definitionen verdeutlichen die weitgefächerten Dimensionen des Begriffs der mathematischen Basiskompetenz. Gemeinsam ist allen Ausführungen, dass es sich stets um grundlegende Inhalte handelt, ohne dessen Beherrschung der weitere Kompetenzerwerb deutlich erschwert ist oder gar unmöglich wird.

Zusammenfassend, und dabei nicht allen o. g. aufgeführten Vorstellungen folgend, sollen für diese Arbeit folgende Bereiche unter dem Begriff „Mathematische Basiskompetenz“ verstanden werden:

> Mathematische Basiskompetenzen umfassen das Operationsverständnis und die kontextbezogene Anwendung arithmetischer Operationen und des Stellenwertsystems (Ehlert et al., 2013; Ennemoser & Krajewski, 2013; Freesemann, 2014; Leiß & Blum, 2006; Reiss & Winkelmann, 2008).

Die Abbildung 1 veranschaulicht zusammenfassend das dieser Arbeit zugrundeliegende Verständnis der stufenweisen Entwicklung bzw. des Erwerbs mathematischer Kompetenzen. Es dient einer allgemeinen Orientierung und legt weder eine starre Abfolge noch zeitliche Abschnitte fest, weswegen die Linien nicht durchgezogen sind. Die Entwicklung mathematischer Vorläuferfertigkeiten (Mengen- und Zahlenwissen; Krajewski, Grüßing & Peter-Kopp, 2009) beginnt bereits ab der Geburt. Darauf aufbauend erfolgt zum Eintritt in die Schule der Erwerb mathematischer Basiskompetenzen. Diese Basiskompetenzen stellen sowohl die Grundlage für den weiteren mathematischen Kompetenzerwerb, als auch für das benötigte mathematische Verständnis im Alltag dar. Am Ende der Schulzeit sollten gute (schul-)mathematische Kompetenzen ausgebildet sein mit dem langfristigen Ziel, mündig, eigenständig und partizipativ am gesellschaftlichen Leben teilzuhaben und mathematisch handlungsfähig zu sein (Basendowski, 2013).

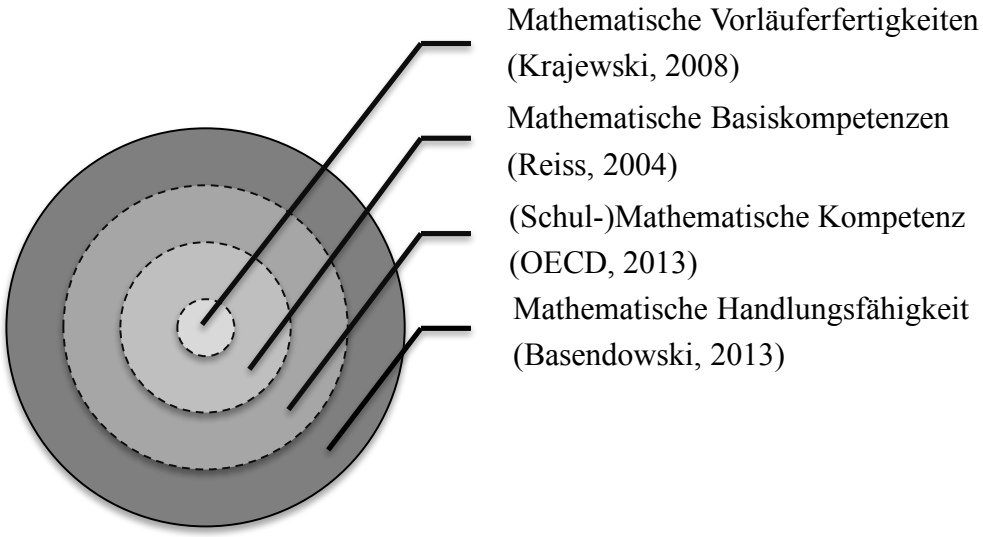


ABBILDUNG 1 MATHEMATISCHE KOMPETENZEN IM ENTWICKLUNGSVERLAUF

Nach der Erarbeitung der Arbeitsdefinition schließt sich die Frage an, wie sich mathematische Vorläuferfertigkeiten und (Basis-)Kompetenzen im Laufe der Zeit entwickeln. Denn klar ist, dass die Entwicklung mathematischer Kompetenzen nicht erst mit Schuleintritt beginnt.

Zu unterscheiden sind in diesem Kontext die beiden Begriffe „Entwicklung“ und „Erwerb“. Der Begriff „Entwicklung“ umfasst eher die dem Menschen innewohnenden Prozesse der „Herausbildung“ bzw. „Entfaltung“ (DUDEN, 2016a). Darunter ist zum Beispiel die Ausbildung der mathematischen Vorläuferfertigkeiten ab der Geburt zu verstehen. Der Begriff des „Erwerbs“ fokussiert verstärkt einen gesteuerten Prozess der Aneigung und des Lernens (DUDEN, 2016b). Dieser Prozess des Erwerbs mathematischer Kompetenzen findet zum Beispiel – durch Bildungspläne gesteuert – im schulischen Kontext statt (siehe auch die Begriffswahl bei Jacobs & Petermann, 2007). Deswegen wird für den vorschulischen Abschnitt der Begriff „Entwicklung“ und für den schulischen Kontext der Begriff „Erwerb“ gewählt.

Einen Überblick darüber, wie mathematische Kompetenzentwicklung von der Vorschule bis über den Kompetenzerwerb in der Primar- und Sekundarstufe I verläuft, geben die folgenden Kapitel 2.2 (Vorläuferfertigkeiten), 2.3 (Erwerb in der Grundschule) und 2.4 (Erwerb in der Sekundarstufe I).

2.2 Entwicklung früher mathematischer Vorläuferfertigkeiten

Schülerinnen und Schüler verfügen zu Schulbeginn bereits über erste mathematische Kompetenzen, die sehr unterschiedlich ausgeprägt vorliegen können. In der ersten Klasse beginnt nicht die „Stunde Null der Mathematik“ (Schneider et al., 2013, S. 100). Diese vorschulischen Kompetenzen, „welche als Startbedingungen zum Schulanfang gute oder schlechte Voraussetzungen für das Erlernen der Mathematik darstellen“ (Krajewski, 2008, S. 13), werden oftmals Vorläuferfertigkeiten genannt (Schneider et al., 2013, S. 67).

Geary (2000, S. 12 f.) bezeichnet Vorläuferfertigkeiten als „biologically primary quantitative abilities“. Er unterscheidet sie von den „biologically secondary abilities“, die kultur- und gesellschaftsabhängig in der Schule vermittelt werden (siehe Abbildung 2).

ABBILDUNG 2 UNTERSCHEIDUNG ZWISCHEN PRIMÄREN UND SEKUNDÄREN MATHEMATISCHEN KOMPETENZEN (GEARY, 2000, S. 12)

Der frühe Erwerb mathematischer Kompetenzen, der bereits unmittelbar nach der Geburt einsetzt (Geary, 2000; Jacobs & Petermann, 2007), stellt ein wichtiges Fundament für den Erwerb mathematischer Kompetenzen in der Schulzeit dar. So nennen Weißhaupt und Peucker (2009, S. 52) verschiedene Aspekte vorschulischer Kompetenzen wie zum Beispiel das Mengenwissen oder den Anzahlaspekt auch „arithmetisches Vorwissen" und betonen dabei, dass mathematische Entwicklung als Aufbau komplexer Verbindungen verstanden werden muss.

Die Darstellung der Entwicklung früher mathematischer Kompetenzen ist im Rahmen dieser Arbeit für das Verstehen der später aufgezeigten Modelle schulischen Kompetenzerwerbs grundlegend. Es liegen für die Entwicklung erster mathematischer Kompetenzen schon frühe Modelle aus den 1970er und 1980er Jahren vor (zum Beispiel Fuson, 1988; Gelman & Gallistel, 1978; Resnick, 1983), die jedoch nur einzelne Aspekte mathematischer Kompetenzentwicklung abbilden und inhaltlich begrenzt sind (Weißhaupt & Peucker, 2009). Jüngere Modelle bündeln diese einzelnen Aspekte und integrieren sie (sog. Skills-Integration-Modelle). Es werden im Folgenden zwei jüngere, empirisch überprüfte Skills-Integration-Modelle nach Fritz und Ricken (2008) bzw. Krajewski (2013) vorgestellt. Anschließend wird die Bedeutsamkeit der frühen Vorläuferfertigkeiten für den schulischen Erwerb aufgezeigt.

2.2.1 Das Stufenmodell nach Fritz und Ricken (2008)

Fritz und Ricken (2008, 2009) generieren ein fünfstufiges Modell zur Entwicklung früher mathematischer Kompetenzen für das Alter von 4;6 bis 6;11 Jahren. Mit diesem Modell werden wesentliche Meilensteile der frühen Kompetenzentwicklung abgebildet. In jüngeren Ausführungen ist die Stufe 1 in zwei Stufen ausdifferenziert, sodass im Folgenden auch eine Entwicklungsstufe 0 dargestellt wird, die in älteren Ausführungen der Autoren nicht zu finden ist. Das Entwicklungsmodell implementiert und verknüpft die Theorie Fusons (1988) zum Zählschema und die Theorie Resnicks (1983) zum Mengenkonzept, die beide aus den 1980er Jahren stammen. Die Entwicklung der komplexen frühen mathematischen Kompetenzen dauert bis in das 3. Schuljahr hinein und verläuft nicht linear. Siegler (1987) nennt dies einen Prozess „sich überlappender Wellen", in dem erste Rechenstrategien nicht sofort abgelegt werden, sondern parallel nebeneinander angewandt werden.

Tabelle 2 gibt einen Überblick über die Zuordnung von Altersstufen und Entwicklungschritten nach Fritz und Ricken. Die Altersangaben sind als ungefähre Richtgrößen zu verstehen.

TABELLE 2 KOMPETENZSTUFENMODELL NACH FRITZ & RICKEN (2008) MIT ALTERS- BZW. KLASSENANGABEN (FRITZ, RICKEN, BALZER, LEUTNER & WILLMES, 2012)

Stufe	Alter bzw. Klassenstufe	Entwicklungsschrittt
0	3–4	Zählzahl
1	4 ½–5 ½	Ordinaler Zahlenstrahl
2	Anfang Klasse 1	Kardinalität und Zerlegbarkeit
3	Mitte/Ende Klasse 1	Enthaltensein und Klasseninklusion
4	Anfang/Mitte Klasse 2	Relationaler Zahlbegriff
5	Ende Klasse 2	Einheiten in Zahlen erkennen

Basierend auf dem Entwicklungsmodell entwickeln Fritz und Ricken in Forschungsgruppen das KALKULIE Diagnose- und Trainingsprogramm für rechenschwache Kinder für die Primarstufe (Gerlach, Fritz, Ricken & Schmidt, 2007) sowie einen Test zur Erfassung mathematischer Kompetenzen im Vorschulalter (MARKO-D; Ricken, Fritz-Stratmann, & Balzer, 2013) inklusive einem an die Diagnostik angepassten Förderprogramm für das Vor- und Grundschulalter (MARKO-T; Gerlach, Fritz, & Leutner, 2013).

ENTWICKLUNGSSTUFE 0

Auf dieser ersten Entwicklungsstufe sind Kinder in der Lage, Mengen (ohne Benennung konkreter Anzahlen) nach *mehr* oder *weniger* zu unterscheiden. Objekte können in diesem Entwicklungsschritt der Größe nach geordnet werden. Die Kinder zeigen eine erste Kenntnis über die Zahlwörter, sodass sie diese geordnet in der Form „einszweidreivierfünf“ aufsagen können (unbreakable chain level).

ENTWICKLUNGSSTUFE I

Im nächsten Entwicklungsschritt gelingt es Kindern, Zahlwörter einzelnen Objekten in einer Reihe zuzuordnen. So können sie zum Beispiel beurteilen, ob zwei unterschiedlich hohe Türme aus Bauklötzen gleich oder unterschiedlich hoch sind, indem sie einen Stein des ersten Turms mit einem Stein des zweiten Turms abgleichen (Eins-zu-Eins-Zuordnung). Eine Sammlung von Objekten kann also „ausgezählt“ werden. Zahlenkenntnis und Mengenkenntnis stehen jedoch immer noch relativ unverbunden nebeneinander. Zum Beispiel wird nicht verstanden, dass fünf ausgezählte Klötze im Sinne des Kardinalzahlprinzips für die konkrete Anzahl an Klötzen steht.

ENTWICKLUNGSSTUFE II

Die Entwicklungsstufe II ist dadurch gekennzeichnet, dass Kinder merken, dass Zahlen, die aufeinander folgen, mit *mehr werden* und *größer werden* zusammenhängen. Dies passiert in der Vorstellung des sog. mentalen Zahlenstrahls. Ein erstes Verständnis von Vorgänger/Nachfolger wird aufgebaut. Erste arithmetische Vorwärts-Operationen („L. hat drei Äpfel, F. gibt ihm noch zwei dazu. Wie viele hat er nun?“) können – oftmals mit Hinzunahme der Finger – zählend gelöst werden. Die Startzahl beim Lösen solcher ersten Aufgaben ist stets die 1. Das Grundwissen besteht aus der Annahme, dass Aufgaben zum Vermehren (Addition) mit einem Voranschreiten auf dem mentalen Zahlenstrahl zusammenhängen.

ENTWICKLUNGSSTUFE III

In einem nächsten Schritt findet die zentrale Entwicklung und Vorstellung des Kardinalzahlprinzips statt. Kinder erkennen, dass zum Beispiel die Zahl 4 für eine bestimmte Menge an Objekten (zum Beispiel Klötzen) steht. Kindern gelingt es, nun auch ab der Zahl 4 weiter zu zählen, ohne bei der 1 anfangen zu müssen (breakable chain level). Dies ist ein gesteigertes Verständnis der Stufe II, auf der ein Voranschreiten auf dem Zahlenstrahl lediglich unspezifisch mit *mehr werden* oder *größer werden* verbunden ist. Ein Summand wird als Stellvertreter einer Menge aufgefasst, zu dem eine zweite Menge hinzugezählt werden kann. „Im Prozess des Rechenlernens stellt diese Integration [von Zählwissen und Mengenwissen; d. Verf.] ein erstes bedeutsames Nadelöhr dar, eine Hürde, die nicht alle Kinder problemlos meistern" (Fritz & Ricken, 2008, S. 36). Es handelt sich hier bereits um erste Ansätze des Teile-Ganzes-Konzepts.

ENTWICKLUNGSSTUFE IV

Kinder wissen in dieser vierten Stufe nicht nur, dass eine Zahl für eine bestimmte Menge an Objekten steht, sondern erkennen zudem, dass die Zahl die vorangegangenen Zahlen enthält. Dies ist insofern eine Erweiterung, da Kinder nun die Aufgabe „Gib mir fünf Klötze, drei davon sollen rot sein" lösen können. Kinder, die dieses Prinzip nicht verinnerlicht haben, wählen wahrscheinlich fünf blaue und zusätzlich drei rote Klötze aus (Fritz & Ricken, 2009, S. 383). Das Verständnis, dass eine Gesamtmenge zwei Teilmengen enthalten kann, ermöglicht es Kindern, Teilmengen und auch Ausgangsmengen sowohl bei Aufgaben der Addition als auch der Subtraktion zu bestimmen. Verstanden wird, dass der Abstand zwischen aufeinanderfolgenden Zahlen stets gleich ist und dass eine Zahl (zum Beispiel die 2) auch den Abstand *zwischen* zwei Zahlen darstellen kann (zum Beispiel zwischen 5 und 7 oder 11 und 13) (relationaler Zahlbegriff). In der Entwicklungsstufe IV werden mit dem Teile-Ganzes-Konzept bedeutsame Voraussetzungen für das Verständnis weiterführender Rechenoperationen gelegt. Die Stufe IV vollzieht sich im Regelfall in der ersten und zweiten Klasse.

ENTWICKLUNGSSTUFE V

Die fünfte Entwicklungsstufe ist durch eine Vertiefung des relationalen Zahlbegriffs mit einer gleichzeitigen Verknüpfung des Teile-Ganzes-Konzepts gekennzeichnet. Es ist den Kindern nun möglich, Additions- und Subtraktionsaufgaben zu lösen, die komplementär als triadische Teil-Teil-Ganzes (T-T-G)-Aufgaben gesehen werden. Die Kinder sind in der Lage, Differenzmengen und Ausgangsmengen im Zahlenraum bis 20 zu bestimmen. Sachaufgaben, die sprachlich eine Addition implizieren, können durch Subtraktionsrechnungen gelöst werden (zum Beispiel bei Fragen zur Ausgangsmenge). Genutzt wird das T-T-G-Verständnis bei Rechenaufgaben mit einem Zehnerübergang sowie bei Aufgaben zur relationalen Zahlbeziehung (Welche Zahl ist um 3 größer als 4?). Typische Aufgaben sind hier Zahlenmauern.

2.2.2 Das ZGV-Modell nach Krajewski (2013)

Ein ebenso junges Modell zur Erklärung der Entwicklung früher mathematischer Kompetenzen stammt von Krajewski und hat bereits stetig begriffliche Präzisierungen und Modifikationen erfahren (Krajewski, 2008; Krajewski et al., 2009). Das Zahl-Größen-Verknüpfung-Modell (kurz: ZGV-Modell) unterscheidet sich an verschiedenen Stellen von dem Modell von Fritz und

Ricken (2009) und auch von den Annahmen Fusons (1988), Resnicks (1983) sowie Gelman und Gallistels (1978). Die wissenschaftliche und praktische Bedeutsamkeit des Modells Krajewskis liegt in den Befunden, die die Autorin im Rahmen einer umfangreichen Längsschnittuntersuchung gewinnt und das Modell empirisch bestätigen (Krajewski, 2008). Die Autorin begleitet dafür 148 Kinder über die Zeit des letzten Kindergartenjahrs bis in die zweite Klasse hinein und kann damit Aussagen über spezifische und unspezifische Prädiktoren mathematischer Kompetenzentwicklung treffen.

Auf Grundlage des ZGV-Modells sind zwei bisher unveröffentlichte diagnostische Verfahren zur Erfassung mathematischer Basiskompetenzen für das Vor- und Grundschulalter (MBK-0 und MBK-1; Krajewski & Ennemoser, 2013, S. 57) sowie das Programm Mengen-Zählen-Zahlen zur Förderung mathematischer Kompetenzen im Vor- und Grundschulalter (Krajewski, Nieding & Schneider, 2013) entwickelt werden.

Das ZGV-Modell gliedert sich in drei Ebenen: Basisfertigkeiten (Ebene 1), Einfaches Zahlverständnis (Ebene 2) und Tiefes Zahlverständnis (Ebene 3).

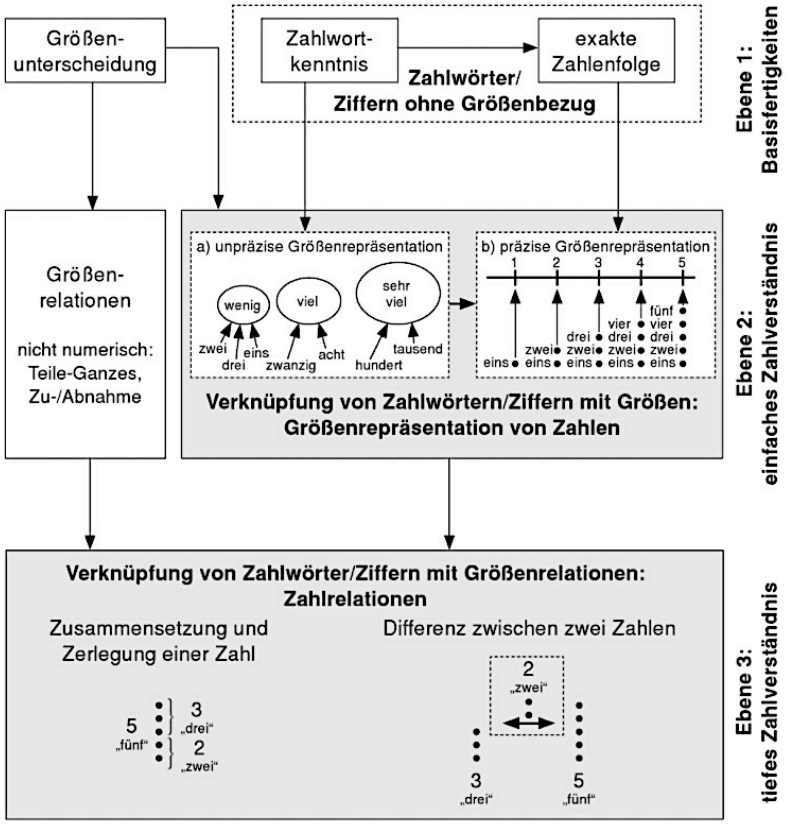


ABBILDUNG 3 ENTWICKLUNGSMODELL DER ZAHL-GRÖßEN-VERKNÜPFUNG NACH KRAJEWSKI (2013, S. 201)

EBENE 1: BASISFERTIGKEITEN

Auf der Ebene 1 erwerben Kinder zum einen die Fähigkeit, Mengen hinsichtlich ihres Volumens oder ihrer Ausdehnung voneinander zu unterscheiden. Zudem erwerben sie eine Zahlwortkenntnis und können die Zahlwortreihe korrekt aufsagen (vorwärts wie rückwärts). Ein konkreter Zahl-Größen-Bezug besteht noch nicht. Wie in der Abbildung 3 zu sehen, entwickeln sich die Größenunterscheidung und die Zahlwortkenntnis zu Beginn unabhängig voneinander.

Krajewski führt das Aufsagen der Zahlwortreihe auf minimalistische Kompetenzen zurück und verortet auch das Rückwärtszählen oder das nicht bei 1 beginnende Zählen auf der Ebene 1. Hier unterscheidet sich ihr Modell von dem Fritz und Rickens und von der Idee Fusons (Schneider et al., 2013, S. 34).

EBENE 2: EINFACHES ZAHLVERSTÄNDNIS

Auf der Ebene 2 findet eine Verknüpfung von Zahlwörtern mit Größen statt. Auf der Ebene 2a kategorisieren Kinder, teilweise auch unabhängig von der Zählkompetenz, bestimmte Zahlen mit großen Mengen oder kleinen Mengen. Sie erkennen, dass die Anzahl „zwei" eher eine kleine Menge ist und die Anzahl „tausend" eher eine große. Im Unterschied zum Modell nach Fritz und Ricken wird bei Krajewski in Ebene 2a nicht angenommen, dass eine konkrete 1-zu-1-Zuordnung zwischen Zahl und Objekt stattfindet. Diese Zuordnung ist hier unpräzise. Der Schritt 2a ist Voraussetzung für die Ebene 2b, in der die exakte Zahlenfolge mit aufsteigenden Mengen und Größen verknüpft wird. Durch das Verstehen des Kardinalverständnisses ist es den Kindern möglich, benachbarte Zahlen nach ihrer quantitativen Größe zu unterscheiden (Krajewski & Ennemoser, 2013, S. 44). Der differenzierte Zweischritt in Ebene 2 kann erklären, weshalb einige Kinder in der Lage sind, weit auseinander liegende Zahlen in der Größe zu unterscheiden, jedoch nicht Nachbarzahlen (Schneider et al., 2013, S. 35).

Parallel zu den Ebenen 2a und 2b entwickelt sich, bevor der Übergang zur Ebene 3 erfolgt, das Verständnis für Teil-Ganzes-Beziehungen und die Veränderung von Größen durch Zu- oder Abnahme.

EBENE 3: TIEFES ZAHLVERSTÄNDNIS

Auf der Ebene 3 erkennen Kinder, im Regelfall im Alter von sechs Jahren, dass eine Zahl aus mehreren „Zahlen" besteht und dass somit Zahlen auch zerlegt werden können. Dieser Schritt stellt eine Erweiterung des unpräzisen Teil-Ganzes-Konzepts (eine große Menge wird in kleine Mengen aufgeteilt) aus Stufe 2 dar, da hier nun die einzelnen Teile mit konkreten Zahlen benannt werden können (5 wird zu 2 und 3). Zudem verstehen die Kinder, dass Zahlen auch Relationen anzeigen können, also zum Beispiel die Differenz zwischen zwei Zahlen (zum Beispiel 2 zeigt die Differenz zwischen 3 und 5 an). Das Verständnis der Zahlbeziehung ist für das weitere Rechnen in der Schule unabdingbar und stellt häufig noch eine Hürde für Schülerinnen und Schüler dar (Schneider et al., 2013, S. 31).

2.2.3 Bedeutung früher mathematischer Vorläuferfertigkeiten für den schulischen Kompetenzerwerb

Die Modelle von Fritz und Ricken (2008, 2009) sowie das von Krajewski (2013) bilden empirisch abgesichert ab, wie sich frühe mathematische Kompetenzen entwickeln. Diese Entwicklungsmodelle ermöglichen eine Einordung und Einschätzung von Entwicklungsständen von Kindern und jungen Schulkindern. Die mathematischen Vorläuferfertigkeiten haben grundsätzlich eine hohe Bedeutung für den weiteren mathematischen Kompetenzerwerb in der Schule. Auf bildungspolitischer Ebene wird dies dadurch bestätigt, dass die frühe mathematische Bildung im „Orientierungsplan für Bildung und Erziehung für den Elementarbereich"

(Niedersächsisches Kultusministerium, 2005) als Förderbereich expliziert und fest verankert ist.

Die Bedeutung früher mathematischer Kompetenzen für den weiteren Erwerb basiert inzwischen auf einer umfassenden wissenschaftlichen Erkenntnislage (Schneider et al., 2013, S. 53). Krajewski (2008) zeigt, dass das verbale Zählen eine der frühsten Entwicklungsschritte im mathematischen Lernen darstellt (Stufe 0 bei Fritz und Ricken, 2008; Ebene 1 bei Krajewski, 2013) und ein spezifischer Prädiktor für den Kompetenzerwerb in der Grundschule ist. „Zählen [dient] als Voraussetzung für die Herausbildung von Rechenoperationen wie Addition und Subtraktion und damit als Voraussetzung für die arithmetische Kompetenz" (Schneider et al., 2013, S. 23). Das sog. Zahlenvorwissen im Vorschulalter (Zählfertigkeiten, arabisches Zahlenwissen, erste Rechenfertigkeiten) kann zu ungefähr 26 % die Unterschiede in den Mathematikleistungen im ersten Schuljahr erklären (Krajewski, 2008, S. 172) und gilt nach der Intelligenz (fluide Intelligenz; gemessen durch den CFT 1) als zweitwichtigster Prädiktor für mathematische Kompetenz. Interessant ist in dem Zusammenhang, dass der Einfluss der Intelligenz im Laufe der ersten beiden Schuljahre sinkt, die Bedeutsamkeit des Zahlenvorwissens jedoch konstant hoch bleibt (Krajewski, 2008, S. 183 f.). Hinsichtlich der spezifischen Prädiktoren halten Krajewski und Schneider (2006) zusammenfassend fest, dass die Kompetenzen auf den ersten beiden Ebenen des ZGV-Modells (Ebene 1: Basisfertigkeiten, Ebene 2: Einfaches Zahlverständnis) als diejenigen Vorläuferfertigkeiten zu betrachten sind, die einen Großteil der Unterschiede der mathematischen Leistungen bis zum Ende der Grundschulzeit aufklären können. Fritz und Ricken (2008, S. 37) schließen sich dem an und konstatieren hinsichtlich des Teil-Teil-Ganzes-Konzepts:

> „Die Anwendung dieses Wissens – nämlich dass Mengen in Teile zerlegt und ohne Schaden zu nehmen wieder zusammengesetzt werden können – auf konkrete Quantitäten gilt als bedeutsamster Schritt in der Entwicklung des mathematischen Verständnisses (Gerster, 2003; Resnick, 1983). Er ist deshalb besonders bedeutsam, weil ein sicheres Verständnis der Grundrechenarten und vieler weiterführender Rechenoperationen auf diesem Konzept basieren."

Dornheim (2008) repliziert diese Ergebnisse und nennt das spezifische Zahlenvorwissen (Beherrschung der Zahlwortreihe, Anzahl- und Kardinalzahlkonzept, flexibles Zählen) den Hauptprädiktor für die Rechenleistungen in den ersten beiden Schulbesuchsjahren. Auch Stern (1997) statuiert für Schülerinnen und Schüler, die in der Vorschulzeit ein besseres numerisches Verständnis (Zahleninvarianz, Mengenschätzen) aufweisen, dass eben diese den Vorsprung im Laufe der Grundschulzeit sogar weiter ausbauen können.

Neben den spezifischen Prädiktoren spielen auch „unspezifische Prädiktoren" eine bedeutsame Rolle im weiteren Kompetenzerwerb. Sie unterscheiden sich von den spezifischen Faktoren dahingehend, dass sie nicht nur Aussagekraft für das mathematische Lernen haben, sondern auch die Entwicklung zum Beispiel des Schriftspracherwerbs beeinflussen. Als unspezifische Prädiktoren werden u. a. die Intelligenz, der familiäre Anregegungsgehalt, das Geschlecht, die phonologische Bewusstheit (Schneider et al., 2013, S. 56–65) und die Schnelligkeit beim Zugriff auf Zahlworte im Langzeitgedächtnis (Krajewski & Schneider, 2006) identifiziert.

Auch das familiäre Anregungspotenzial spielt eine bedeutsame Rolle. Der Einfluss dieses Prädiktors nimmt im Laufe der Grundschulzeit stark zu. Am Ende der vierten Klasse kann dieser Faktor 18 % der Unterschiede in den Mathematikleistungen aufklären (Krajewski & Schneider, 2006; Schneider et al., 2013). Erwähnt sei dazu eine Untersuchung von Niklas und Schneider

(2012) zum „Home Numeracy Environments“. Die Autoren stellen fest, dass der häusliche mathematische Anregungsgehalt durch die Eltern einen über die Dauer der Vorschulzeit hinweg signifikanten Einfluss auf die mathematische Kompetenzentwicklung der Kinder hat und die Unterschiede in den Mathematikleistungen am Ende der ersten Klasse zu 19 % aufklären kann.

Innerhalb der LOGIK-Studie (Weber & Stefanek, 1998) werden über einen Zeitraum von 14 Jahren u. a. mathematische Kompetenzen bei Kleinkindern bis in das Jugend- bzw. Erwachsenenalter erhoben (siehe Kapitel 2.4.2.1). Stern arbeitet dabei in der Arbeitsgruppe zu dem Bereich der mathematischen Kompetenzen und kann Zusammenhänge zwischen frühen und späten mathematischen Kompetenzen ermitteln. So kann die Autorin eine Korrelation von $r = .43$ zwischen dem numerischen Verständnis im Alter von fünf Jahren und der mathematischen Intelligenz, gemessen durch den Untertest „Mathematisches Denken“ in einem Intelligenztest, im Alter von 23 Jahren ermitteln (vgl. Abbildung 4). Nach Hehn (2014) handelt es sich dabei um einen positiven und nahezu deutlichen linearen Zusammenhang.

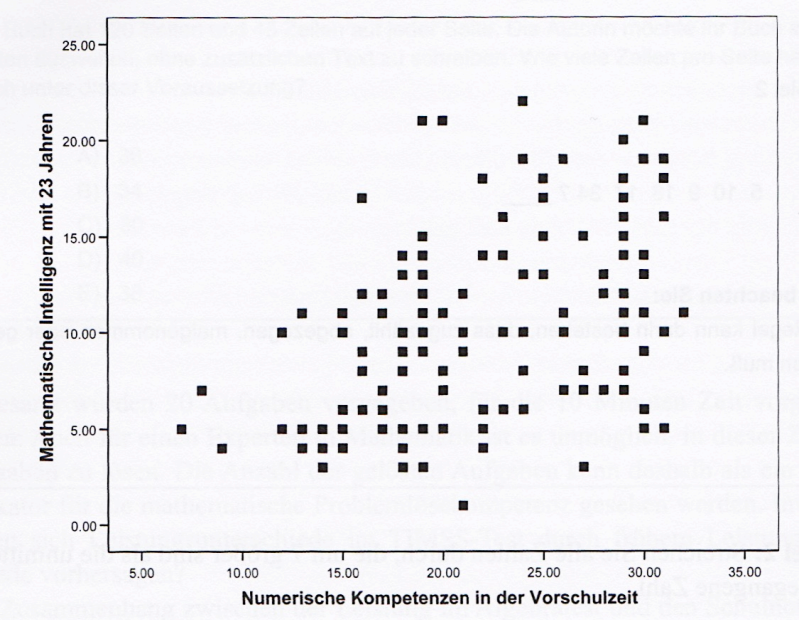


ABBILDUNG 4 ZUSAMMENHANG ZWISCHEN NUMERISCHER KOMPETENZ IM VORSCHULALTER UND MATHEMATISCHER INTELLIGENZ MIT 23 JAHREN (STERN, 2008, S. 196)

Auch in der internationalen Forschung kommt der Bedeutung früher mathematischer Kompetenzentwicklung für die weitere Schullaufbahn große Aufmerksamkeit zu (Cross, Woods & Schweingruber, 2010; Denton & West, 2002; Education Commission of the States, 2013; Jordan, 2010). Eine finnische Forschergruppe um Lerkkanen und Kollegen (2005) fand heraus, dass frühe mathematische Kompetenzen spätere Leseleistungen besser vorhersagen als die Lesekompetenzen selbst. Allerdings betrachtet die Forschergruppe lediglich die Entwicklung bis zu Beginn der Grundschule in Klasse eins und zwei.

Eine weitere zentrale Arbeit zur Bedeutung vorschulischer Kompetenzen leisten Duncan und Kollegen (2007). In ihrer Metaanalyse ziehen sie sechs Längsschnittstudien hinzu und untersuchen, welche Faktoren spätere schulische Kompetenzen vorhersagen können. Alle sechs Studien starten mit den Kompetenzerhebungen im Vorschulalter. Drei Studien begleiten die Kinder bis zur dritten Klasse, eine Studie bis zur vierten Klasse, eine Studie bis zur fünften Klasse und eine weitere bis in das frühe Erwachsenenalter. Es stellt sich heraus, dass frühe mathematische

Kompetenzen wie das „Zahlenwissen“ (knowledge of numbers, number sense) oder „Größenvergleich/-unterscheidung“ (ordinality) mit einer durchschnittlichen kleinen bis mittelgroßen Effektstärke (nach Cohen, 1988) von $d = .34$ die stärkste Vorhersagekraft für spätere Kompetenzen hat. Erkannt wird auch hier, dass frühe mathematische Kompetenzen die spätere Lesekompetenz besser vorhersagen als die frühe Lesekompetenz selbst (vgl. Lerkkanen u. a., 2005).

Die Ausführungen zeigen, dass die Entwicklung früher mathematischer Kompetenzen auf die weitere Entwicklung einen großen Einfluss hat. Für junge Heranwachsende ist der Mathematikunterricht dennoch der zentrale Lernkontext, in dem sie die Welt der Mathematik begreifen und verstehen (Fritz & Ricken, 2008, S. 78; Schneider et al., 2013, S. 209).

Im folgenden Kapitel wird eben dieser Erwerb mathematischer Basiskompetenzen in der Grundschulzeit thematisiert. Auch hier wird zu Beginn, wie in Kapitel 2.2, der allgemeine mathematische Kompetenzerwerb dargelegt und anschließend die Bedeutung für den weiteren Erwerb in der Sekundarstufe I aufgeschlüsselt.

2.3 Erwerb mathematischer Basiskompetenzen in der Grundschulzeit

Die bisher dargestellte Entwicklung *früher* mathematischer Kompetenzen basiert inszwischen auf einer umfangreichen Forschungs- und Erkenntnislage. Weniger differenzierte Ergebnisse gibt es bezüglich der Betrachtung des *schulischen* Kompetenzerwerbs, der in Deutschland erst seit den 1990er Jahren größere Aufmerksamkeit erhält (Walther, Geiser, Langeheine & Lobemeier, 2003). Zu den wenigen nationalen Forschungsergebnissen kommt erschwerend hinzu, dass ein internationaler Vergleich aufgrund der vielen unterschiedlichen Bildungssysteme erschwert und nur mit Einschränkungen möglich ist (Schmitt, 2001).

Einige nationale Studien belegen, dass das Leistungsspektrum von Schülerinnen und Schülern bereits *zu Beginn* der Grundschule breit gefächert ist (Hasemann, 2001; Hasemann & Gasteiger, 2014). Diese Erkenntnis wird international weitgehend repliziert (Denton & West, 2002). Gerade bei den divergierenden Lernausgangslagen der Schülerinnen und Schüler beim Schuleintritt, die sich im Zuge des inklusiven Bildungssystems durch die inkludierten Schülerinnen und Schüler mit Beeinträchtigungen weitaus heterogener werden (Schels et al., 2015), sind im Sinne von Prävention und Förderung Erkenntnisse zu Erwerbsverläufen notwendig. Empirische Basis für Erwerbsprozesse oder Kompetenzstufenmodelle sind dabei Längsschnittstudien, an denen es in Deutschland bisher mangelt (Ditton & Krüsken, 2009). Ein Beispiel für eine Längsschnittstudie ist die bayrische Studie „Kompetenzaufbau und Laufbahnen im Schulsystem“ (kurz: KOALA-S) (Ditton, 2007). Die KOALA-S-Studie betrachtet von 2003 bis 2005 den Kompetenzerwerb von 650 Schülerinnen und Schülern von der zweiten bis zu vierten Klassen u. a. im Bereich Mathematik (Grundrechenarten, Zahlenbegriff und -verständnis, Geometrie, sachbezogene Mathematik). Die Ergebnisse zeigen, dass bei allen Schülerinnen und Schülern eine positive Kompetenzentwicklung im Bereich Mathematik stattfindet. Diejenigen aus bildungsfernen Elternhäusern entwickeln sich jedoch schlechter, was sich dann auch bei der Übergangsentscheidung in die Sekundarstufe I abbildet. Es wird seltener das Gymnasium als weiterführende Schule gewählt. Zudem wird ein Schereneffekt am Ende der Klasse zwei festgestellt, der bis zum Ende der Klasse vier stark abnimmt. „Allerdings reicht das nicht aus, um die Leistungsunterschiede völlig auszugleichen. Obwohl sich die Schülerleistungen angleichen, bleiben die Leistungsrangreihen im Wesentlichen erhalten“ (Ditton & Krüsken, 2009, S. 58).

In den folgenden beiden Unterkapiteln werden sowohl aus der pädagogisch-psychologischen als auch aus der neurowissenschaftlichen Sichtweise Modelle zum mathematischen Kompetenzerwerb in der (Grund-)Schulzeit aufgezeigt. Die Berücksichtigung beider Forschungsdisziplinen als Korrelate im Themengebiet des mathematischen Kompetenzerwerbs ermöglicht eine differenzierte theoretische Grundlegung für das Forschungsanliegen dieser Arbeit. Abschließend wird dargelegt, welche Bedeutung die erworbenen Kompetenzen in der Grundschulzeit für den weiteren mathematischen Erwerb haben.

2.3.1 Pädagogisch-Psychologische Perspektive

Zur Untersuchung des Erwerbs mathematischer Kompetenzen über mehrere Schuljahre hinweg gibt es im deutschsprachigen Raum wenige längsschnittliche Untersuchungen. Auch international stellt dies ein Forschungsdesiderat dar (Jacobs & Petermann, 2007, S. 29; Ufer, Reiss & Heinze, 2009, S. 65).

An dieser Stelle liegt die Frage zugrunde, welche Kompetenzen Schülerinnen und Schüler im Laufe der ersten vier Schuljahre erwerben, die Voraussetzung für den weiteren Erwerb in der Sekundarstufe I sind. Aufbauend auf die bereits skizzierten „biologically primary quantitative abilities“ (vgl. Kapitel 2.2) wird im Folgenden der Erwerb der sog. „biologically secondary abilities“ (Geary, 2000, S. 12 f.) dargestellt.

Es werden das psychologisch-pädagogische Modell nach Reiss (2004), die Bildungsstandards für den Primarbereich (Kultusministerkonferenz, 2004) und auch internationale Modelle und Ansätze zum Kompetenzerwerb in der Primarstufe reflektiert.

2.3.1.1 Das Kompetenzstufenmodell nach Reiss (2004)

Das Modell nach Reiss (2004) orientiert sich maßgeblich am Kompetenzmodell der Internationalen Grundschul-Lese-Untersuchung (IGLU) aus dem Jahr 2011, in der auch mathematische Kompetenzen von Viertklässlerinnen und Viertklässler erhoben werden (Walther et al., 2003). Das Modell nach Reiss (2004) besteht aus fünf hierarchisch angeordneten Kompetenzstufen (bzw. später Kompetenzniveaus; Reiss u. a., 2007) (siehe Tabelle 3).

Folgende Tabelle 3 gibt einen Überblick über die Kompetenzstufen.

TABELLE 3 KOMPETENZSTUFENMODELL (REISS, 2004; REISS & WINKELMANN, 2008)

Kompetenzstufen		Klasse 1	Klasse 2	Klasse 3	Klasse 4
I	Numerisches und begriffliches Grundlagen-wissen (Routine-prozeduren)	Zähl-fähigkeiten bis etwa 20	Übertragung der Ergebnisse des kleinen Einspluseins auf Zehnerzahlen	Grundlagen des kleinen Einmaleins	Kontextfreies Rechnen im Zahlenraum bis 100

Kompetenzstufen		Klasse 1	Klasse 2	Klasse 3	Klasse 4
II	Einfache Anwendung von Grundlagen-wissen	Grundlagen des kleinen Einspluseins	Grundlagen des kleinen Einmaleins	Kontextfreies Rechnen im Zahlenraum bis 100	Schriftliche Addition im Zahlenraum bis 1000 ohne Übergänge
III	Sicheres Rechnen in curricularem Umfang und einfaches Modellieren	Zähl-fertigkeiten über 30 hinaus	Gute Beherrschung des kleinen Einmaleins	Halb-schriftliches Rechnen im Zahlen-raum bis 1000	Rechen-ergebnisse überschlagen
IV	Beherrschung der Grundrechenarten unter Nutzung der Dezimalstruktur und begriffliche Modellierung	Addition und Subtraktion mit Zehner-zahlen	Rechnen in Sachkontexten	Informationen in Sach-situationen nutzen und verarbeiten	Sichere Beherrschung der schriftlichen Rechen-verfahren in Sachkontexten
V	Anspruchsvolles Problemlösen im mathematischen Kontext	Beschreiben und Modellieren von Sachsituationen	Problemlösender Umgang mit Größen	Anwendung mehrerer Grundrechen-arten in komplexen Sachsituationen	Bewältigung kombina-torischer Fragestellungen

Das Hauptanliegen des Anfangsunterrichts in der Grundschule besteht darin, die Grundrechenarten Addition, Subtraktion, Multiplikation und Division zu vermitteln. Das Kompetenzstufenmodell zeigt, dass eben diese Vermittlung der arithmetischen Kompetenzen ($+, -, \cdot, :$) in den Klassestufen eins bis vier eine zentrale Aufgabe der Grundschulmathematik darstellt (vgl. Schneider u. a., 2013, S. 100). Die einzelnen Kompetenzstufen bzw. -niveaus sind zunächst inhaltsunspezifisch und jahrgangsübergreifend formuliert. Jedes Kompetenzniveau wird darüber hinaus für die Schuljahre eins bis vier differenziert. Es wird in dem Modell davon ausgegangen, dass das Wissen über Fakten und das Beherrschen einfacher Rechenoperationen wie 5 x 2 = 10 unverzichtbare Grundlage für eine Aufgabe wie 50 x 2 = 100 darstellt und mit diesem Wissen letztlich erst Modellierungsaufgaben gelöst werden können wie „Jeder Schüler erhält 2 Bonbons. Wie viele Bonbons haben 50 Schüler zusammen?“ (vgl. Reiss u. a., 2007, S. 114). Es liegt also sowohl eine vertikale als auch eine horizontale Hierarchie vor. Reiss (2009, S. 196) schreibt zu ihrem Modell: „Es liegt nahe, diese Kompetenzen am Ende der vierten Klasse als Mindeststandards anzusehen, da sie notwendige Voraussetzungen für den Unterricht in der fünften Klasse darstellen.“

Im Rahmen der BIGMATH-Längsschnittstudie ist das Modell empirisch überprüft worden. Bestätigt werden konnte es für die Klassen zwei und drei (Reiss et al., 2007) und auch für die Klasse vier (Ufer et al., 2009), wobei stets Itemrevisionen in Folge der statistischen Analysen stattfinden mussten. Trotz empirischer Überprüfungen muss das Modell derzeitig als vorläufig angesehen werden (Reiss & Winkelmann, 2009, S. 140). Reiss (2004, S. 644) kommt mit ihrem Modell der Forderung nach einer „Konkretisierung“ allgemeiner Kompetenzen als Basis für die Formulierung von Zielen mathematischer Schulbildung nach. „Normative Vorgaben sagen [...] nichts darüber aus, ob und, wenn ja, wie Schülerinnen und Schüler tatsächlich die gewünschten

Kompetenzen erreichen können“ (Reiss, 2009, S. 117). Differenzierte und empirisch gesicherte Modelle können Zusammenhänge und Abläufe im mathematischen Erwerb aufzeigen, zur Identifizierung von Lernrückständen dienen und die Entwicklung entsprechender Unterrichtsgestaltung bzw. Förderung (Reiss et al., 2007, S. 114) ermöglichen. Letztlich unterstützt ein Kompetenzstufenmodell Lehrkräfte darin, die Schwierigkeit und den Anforderungsgehalt von Aufgaben besser einschätzen zu können und Aufgaben abzuleiten, die zuverlässig bestimmte Erwerbsschritte abbilden.

2.3.1.2 Bildungsstandards für den Primarbereich

Die Bildungsstandards für die Primarstufe sind im Jahr 2004 durch die Kultusministerkonferenz festgelegt worden. Die Bildungsstandards beschreiben fachbezogene Kompetenzen, die zu einem bestimmten Zeitpunkt, hier am Ende der Primarstufe, von den Schülerinnen und Schülern erreicht werden sollten. Neben den Bildungsstandards für das Fach Mathematik in der Primarstufe gibt es zusätzlich Standards für weitere Fächer sowie für den mittleren Schulabschluss und den Sekundarstufen-II-Abschluss. Der Unterschied zu den curricularen Lehrplänen bzw. Kerncurricula der Bundesländer liegt darin, dass die Bildungsstandards *keine* Unterrichtsstandards darstellen sollen. Zentrale Ziele der Bildungsstandards sind „Orientierung und Evaluation“ (Blum, 2006, S. 16). Neben einer Transparenz für alle am Bildungsprozess beteiligten Akteure darüber, welche Kompetenzen beim Übergang in die Sekundarstufe I vorliegen sollen (Orientierung) ermöglichen die Bildungsstandards durch konkrete Leistungsüberprüfungen eine Abbildung des Kompetenzstands deutscher Schülerinnen und Schüler in spezifischen Domänen (Evaluation) (Blum, 2006, S. 16).

Unterschieden werden in den Bildungsstandards des Fachs Mathematik für die Primarstufe allgemeine mathematische Kompetenzen und inhaltsbezogene mathematische Kompetenzen.

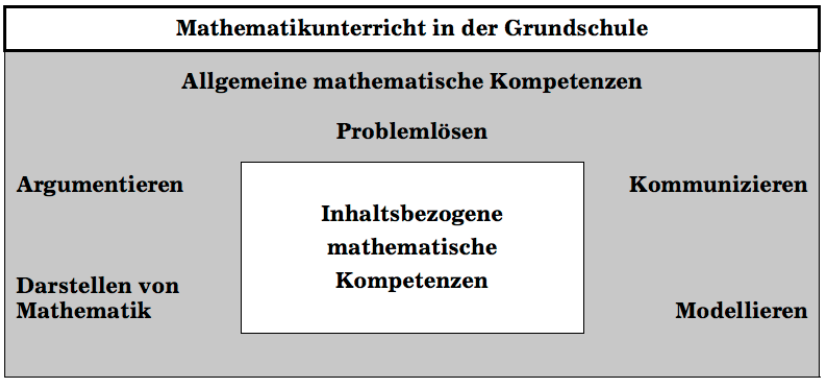


ABBILDUNG 5 ÜBERBLICK ÜBER DIE BEREICHE DER BILDUNGSSTANDARDS FÜR DEN PRIMARBEREICH (KULTUSMINISTERKONFERENZ, 2004, S. 7)

Zu den allgemeinen mathematischen Kompetenzen zählen: Modellieren, Kommunizieren, Problemlösen, Argumentieren und das Darstellen von Mathematik. Die inhaltsbezogenen mathematischen Kompetenzen hingegen dienen dem Ziel einer „innermathematischen Strukturorientierung“ (Kultusministerkonferenz, 2004, S. 7) und sind für den weiteren Kompetenzerwerb von „fundamentaler Bedeutung“ (Kultusministerkonferenz, 2004, S. 8). Die allgemeinen mathematischen Kompetenzen richten sich eher auf die Anwendung bzw. die Nutzung mathematischen Wissens mit dem Ziel einer „lebendigen Auseinandersetzung mit Mathematik“

(Kultusministerkonferenz, 2004, S. 7). Diese Kompetenzen gewinnen in einem Kontext der eher „außermathematischen Anwendungsorientierung“ (Kultusministerkonferenz, 2004, S. 7) Bedeutung.

Die folgende Tabelle 4 zeigt die inhaltsbezogenen mathematischen Kompetenzen der Bildungsstandards aufgeschlüsselt.

TABELLE 4 INHALTSBEZOGENE KOMPETENZEN DER BILDUNGSSTANDARDS MATHEMATIK FÜR DEN PRIMARBEREICH (KULTUSMINISTERKONFERENZ, 2004, S. 9 FF.)

Inhaltsbereich	Inhaltliche Konkretisierung für 4. Klasse
Zahlen und Operationen	Zahlendarstellungen und -beziehungen im Zahlenraum bis eine Million verstehen, Rechenoperationen verstehen und beherrschen sowie in Kontexten rechnen
Raum und Form	Orientierung im Raum, geometrische Figuren und einfache geometrische Abbildungen erkennen, benennen und darstellen, Flächen und Rauminhalte vergleichen und messen
Muster und Strukturen	Gesetzmäßigkeiten bei Mustern sowie funktionale Beziehungen in Sachsituationen und Tabellen erkennen, beschreiben und darstellen
Größen und Messen	Größeneinheiten kennen, vergleichen, schätzen, messen und umwandeln können, mit Größen in Sachsituationen umgehen können
Daten, Häufigkeit und Wahrscheinlichkeit	Daten erfassen und darstellen können, Wahrscheinlichkeiten einfacher Zufallsexperimente einschätzen und vergleichen können

Die inhaltsbezogenen mathematischen Kompetenzen stellen in den Bildungsstandards eine Art Fundament dar, die für die Bewältigung anwendungsbezogener mathematischer Aufgaben notwendig sind. Beide Komponenten bilden letztlich die notwendige Grundlage für den weiteren Kompetenzerwerb in der Sekundarstufe I.

Die Verzahnung der beiden Bereiche spiegelt sich auch in dem Kompetenzstufenmodell nach Reiss (2004) wider, in dem die Autorin auf den Stufen I, II und III Grundlagenwissen, Grundfertigkeiten und sicheres Rechnen verortet, bevor auf Stufe IV und V das Modellieren und Problemlösen erwartet wird. Letztlich zeigen das differenzierte Modell nach Reiss (2004) und die Bildungsstandards für das Fach Mathematik in der Primarstufe als zwei sich ergänzende Korrelate, welche Kompetenzen die Schülerinnen und Schüler im Laufe der Primarstufe im Bereich mathematischer Bildung erwerben und wie die spezifischen mathematischen Inhalte dabei sowohl horizontal (Inhalte im Laufe der Schuljahre) als auch vertikal (Komplexität der Inhalte) hierarchisch aufeinander aufbauen.

2.3.1.3 Internationale Perspektive

Es ist bereits problematisiert, dass in Deutschland kein allgemeingültiges Modell bezüglich des allgemeinen mathematischen Kompetenzerwerbs vorliegt (Vom Hofe, Hafner, Blum & Pekrun, 2009; Jacobs & Petermann, 2007; Reiss, 2004). Dies gibt Anlass, den Fokus darauf zu richten, wie der Erwerb mathematischer Kompetenzen international modelliert und inwiefern das Erreichen spezifischer Kompetenzen an bestimmten Schlüsselstellen der Schullaufbahn curricular erwartet wird. Reiss und Pekrun (2007, S. 112) stellen fest: „Auch im internationalen Bereich

gibt es nur wenige Studien zur mathematischen Kompetenz, die eine längsschnittliche Perspektive haben und damit auf die Entwicklung mathematischer Kompetenz fokussieren". Zudem ist anzumerken, dass die internationale Vergleichbarkeit nur eingeschränkt möglich ist, da die Grundschulzeit in verschiedenen Ländern variiert und eine große Spannweite haben kann (Schmitt, 2001). Zudem ist ein Vergleich des mathematischen Erwerbs bei europäischen Kindern mit Kindern aus dem asiatischen Sprachraum erschwert, da es in Europa keine gänzliche Übereinstimmung zwischen dem dekadischen System der arabischen Zahlenform und der sprachlichen Wortform gibt (Jacobs & Petermann, 2007, S. 29). So wird die Zahl 21 zum Beispiel in Deutschland „Einundzwanzig" ausgesprochen, in Asien hingegen „zwei-zehn-eins" (Landerl & Kaufmannn, 2013, S. 69). „Die Bildungsregel unserer Zahlwörter stellt also eine ernst zu nehmende Hürde dar. Eine perfekte Regelhaftigkeit von Zahlennamen dürfte das Rechnen weniger fehleranfällig machen [...]" (Schneider et al., 2013, S. 97).

Bezogen auf Curricula, wie die Lehrpläne der Bundesländer oder die Bildungsstandards in Deutschland, sollen hier exemplarisch die aus den USA stammenden „Principles and Standards of School Mathematics" (kurz: *Standards*) vorgestellt werden, die schon in den 1980er Jahren von dem „National Council of Teachers of Mathematics" (NCTM) entwickelt worden sind und im Jahr 2000 eine umfassende Überarbeitung erfahren haben. Daneben wird das australische Early Numeracy Research Projekt vorgestellt.

Principles and Standards for School Mathematics (USA)

In den *Standards* werden sowohl Inhalte als auch Methoden für einen möglichst effektiven Mathematikunterricht festgelegt. Differenziert wird in den Anforderungen zwischen den Spannen Vorschule bis zur zweiten Klasse, dritte bis fünfte Klasse, sechste bis achte Klasse und neunte bis zwölfte Klasse. Vergleichbar mit den Stufen in dem bereits vorgestellten Modell nach Reiss (2004) werden in den *Standards* für jede Altersspanne grundlegend die gleichen Inhalte festgelegt (Number, Algebra, Geometry, Measurement, Data Analysis and Probability), deren Anforderungsniveau dann im Laufe der Schuljahr steigt (i. S. eines Spiralcurriculums) (National Council of Teachers of Mathematics, 2000).

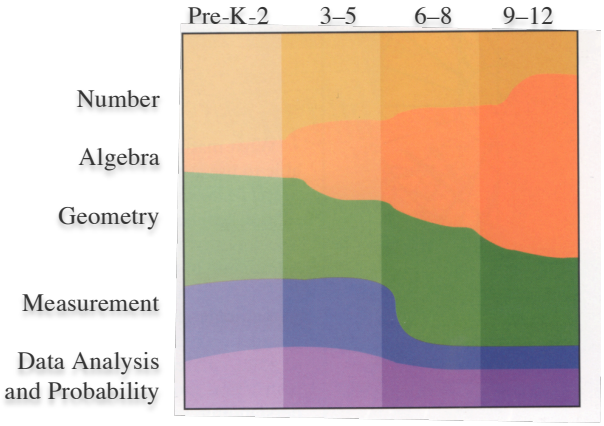


Abbildung 6 Principles and Standards for School Mathematics (NCTM, 2000)

Das wellenartige Modell in der Abbildung 6, dessen Farben sich von links nach rechts intensivieren, verdeutlicht, dass den Inhalten „Number" und „Measurement" zu Beginn der Schullaufbahn eine hohe inhaltliche Relevanz zugeschrieben wird, die jedoch im Laufe der Zeit abnimmt.

Der Bereich „Algebra“ bekommt stetig mehr Gewicht. Die Bereiche „Geometry“ und „Data Analysis and Probability“ haben in den *Standards* über alle Schuljahre hinweg ungefähr ein unverändert starkes Gewicht. Die Abbildung 6 stellt hier nur eine schematische Übersicht der Bereiche dar, die in den *Standards* deutlich differenzierter ausgearbeitet sind. Die *Standards* üben seit der Überarbeitung im Jahr 2000 einen großen Einfluss auf die schulischen Mathematikprogramme aus. In der überarbeiteten Fassung finden sich auch Hinweise zum Umgang mit Schülerinnen und Schülern mit Unterstützungsbedarfen (Verwendung leichter Sprache, mehr Zeit, eher mündliche statt schriftliche Aufgaben, Peer-Tutoring). Zudem stützt sich die zweite Auflage wesentlich umfassender auf Forschungsliteratur und nicht lediglich auf methodisch-didaktische Überlegungen (Mastropieri, Scruggs, Davidson & Rana, 2008, S. 309 f.).

Early Numeracy Research Projekt (Australien)

Eine Untersuchung zum Erwerb mathematischer Kompetenzen wird im australischen Early Numeracy Research Project (ENRP) von 1999 bis 2002 umgesetzt. Das Vorgehen der Untersuchung ähnelt dem von Reiss (2004) mit dem Ziel, abbilden zu können, wie der mathematische Kompetenzerwerb ab dem Eintritt in die Schule verläuft. Clarke (2001, S. 19) untersucht, wie viele der 34.398 Schülerinnen und Schüler Aufgaben bei einem ansteigenden Anforderungsniveau vom einfachen Zählen (Count) bis hin zu Anwendungsaufgaben (Extending and Applying) bereits vor dem Schuleintritt und bis zur Klasse vier lösen können. Die Abbildung 7 zeigt auf grafischer Ebene, welche Kompetenzen von wie vielen Schülerinnen und Schülern (in %) zu welchem Zeitpunkt beherrscht werden. Es steigt die prozentuale Anzahl von Jahrgangsstufe zu Jahrgangsstufe bei denjenigen Schülerinnen und Schülern, die die spezifischen Anforderungen mit dem steigenden Anforderungsniveau der Aufgaben bewältigen können. Es ist erkennbar, dass die Schülerinnen und Schüler im Laufe der Jahre ihre Kompetenzen ausbauen und sich in den verschiedenen Bereichen bis zum Ende der Klasse vier grundsätzlich positiv entwickeln

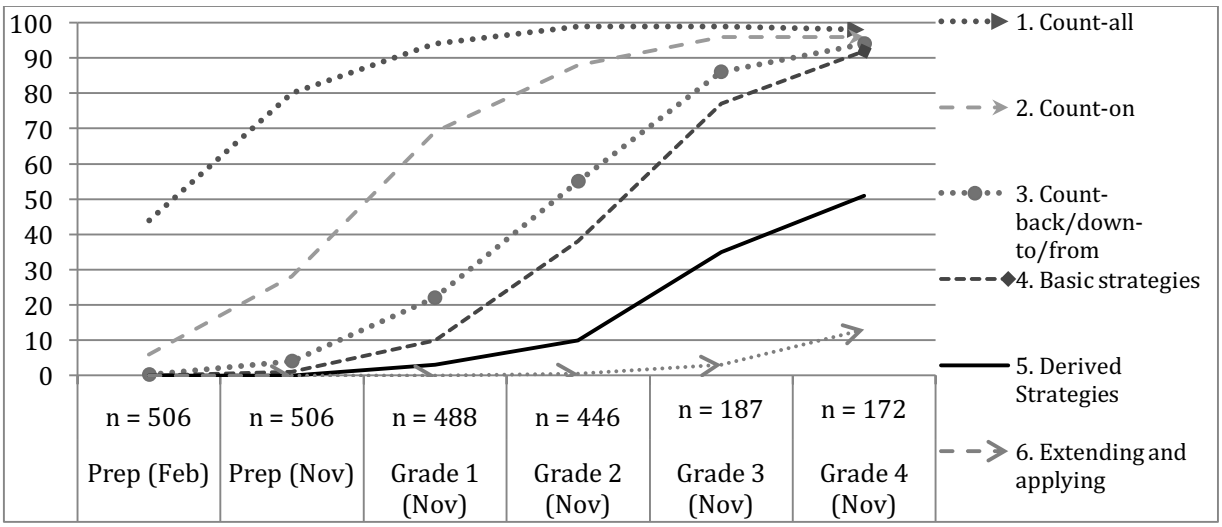


Abbildung 7 Ergebnisse des Early Numeracy Research Projects (Clarke, 2001)

Etwas weniger stark verläuft der Kompetenzerwerb bei Anwendungsaufgaben (Extending and Applying) und bei den Aufgaben, bei denen eigene Strategien abgeleitet werden müssen (Derived Strategies). Die Schülerinnen und Schüler sind am Ende der Klasse vier in diesen beiden Bereichen weniger stark als in den Bereichen der Rechenoperationen. Berücksichtigt werden

muss, dass es sich hier nicht um eine Entwicklung handelt, wie sie in Kapitel 2.2 bzgl. der Vorläuferkompetenzen dargestellt ist, sondern um einen Kompetenzerwerb, der durch Vorgaben schulischer Curricula gesteuert wird. Die Autoren haben von den Ergebnissen bisher kein Kompetenzstufenmodell abgeleitet (vgl. Reiss & Winkelmann, 2008).

Im folgenden Kapitel wird die kognitiv-neurowissenschaftliche Perspektive auf den schulischen mathematischen Kompetenzerwerb fokussiert.

2.3.2 Kognitiv-Neurowissenschaftliche Perspektive

Die Entwicklung und der Erwerb mathematischer Kompetenzen und die Bewältigung komplexer mathematischer Aufgaben finden auch im Bereich der kognitiv-neurowissenschaftlichen und medizinischen Forschung Aufmerksamkeit. Mit einem Blick auf *innere kognitive Prozesse*, auf Funktionen des Gehirns und Denkprozesse wird versucht, Teilbereiche im Gehirn aufzudecken, die bei bestimmten mathematische Anforderungen angesprochen werden. Im Zuge der „Kognitiven Wende“ (Dember, 1974) in den 1960er und 1970er Jahren entstanden einige Modelle zur Zahlenverarbeitung, die sich in Aufbau und Komplexität der Verarbeitungswege unterscheiden (Schneider et al., 2013). Als Grundlage der später vorgestellten „Entwicklungsschritte der Rechenfähigkeit“ nach Jacobs und Petermann (2007) wird vorerst in wenigen Zügen das Modell der drei Repräsentationsformen nach Dehaene (1992) skizziert.

Das Modell der drei Repräsentationsformen nach Dehaene (1992)

Das „Triple-Code-Modell“ ist das bisher umfassendste und das empirisch am weitesten abgesicherte Modell zu Prozessen der Zahlenverarbeitung (Schneider et al., 2013, S. 43). Innerhalb des Modells gibt es drei Repräsentationsformen von Zahlen und Mengen, die als „Codes“ bezeichnet werden. Die drei Verarbeitungskomponenten lauten:

1. Visuell-arabische Repräsentation (Ziffernform)
2. Auditiv-verbale Repräsentation (Wortform)
3. Analoge Größenrepräsentation

Eingehende Informationen können demnach in drei unterschiedlichen Modulen aufgenommen werden (vgl. Abbildung 8). Jedes Modul stellt dabei die Schnittstelle zwischen einem Input und einem Output dar (Dehaene, 1992, S. 32).

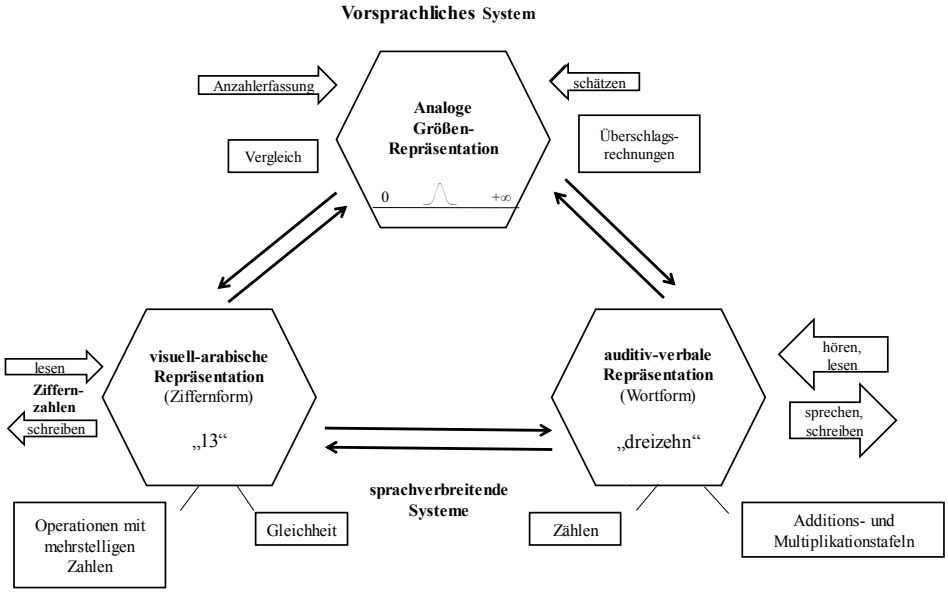


Abbildung 8 Triple-Code-Modell nach Dehaene (1992)

Im Modul der visuell-arabischen Repräsentation (visual arabic number form) erfolgt die Zahlenverarbeitung von geschriebenen Ziffern wie zum Beispiel „13“. Diese Verarbeitung impliziert das Lesen und Schreiben von Ziffernzahlen und beinhaltet auch Rechenoperationen mit mehrstelligen Zahlen. In dem Modul der auditiv-verbalen Repräsentation werden Zahlen in Wortform (zum Beispiel „dreizehn“) durch Hören oder Lesen aufgenommen. In diesem „Zahlwortlexikon“ (Schneider et al., 2013, S. 44) erfolgt die Verarbeitung gesprochener und geschriebener Wörter, das Zählen und die Anwendung von Additions- und Multiplikationsfakten durch das Abrufen von Informationen aus dem Langzeitgedächtnis. Im dritten Modul, der analogen Größenrepräsentation, erfolgen Größenvergleiche von Mengen. Es handelt sich eher um eine unscharfe Repräsentation von Mengen, sodass hier auch Prozesse des Schätzens und der mentale Zahlenstrahl verortet werden. Die drei Module kooperieren untereinander. Diese neuronalen Verknüpfungen erfolgen im Entwicklungsverlauf stets sicherer und schneller. Nach Dehaene (1992) kann jedes der drei Module spezifisch beeinträchtigt sein, wodurch aufgeklärt werden kann, weshalb Probleme auch nur in einzelnen Bereichen unabhängig voneinander auftreten können.

Die Entwicklungsschritte der Rechenfähigkeit nach Jacobs & Petermann (2007)

Jacobs und Petermann (2007) unterscheiden ihre Entwicklungsschritte der Rechenfähigkeit wie auch Geary (2000) in „Primäre vorschulische Fähigkeiten“ und „Sekundäre, in der Schule erworbene Fähigkeiten“. In der folgenden Darstellung findet nur der zweite Teil des Modells mit dem schulischen Bezug Berücksichtigung.

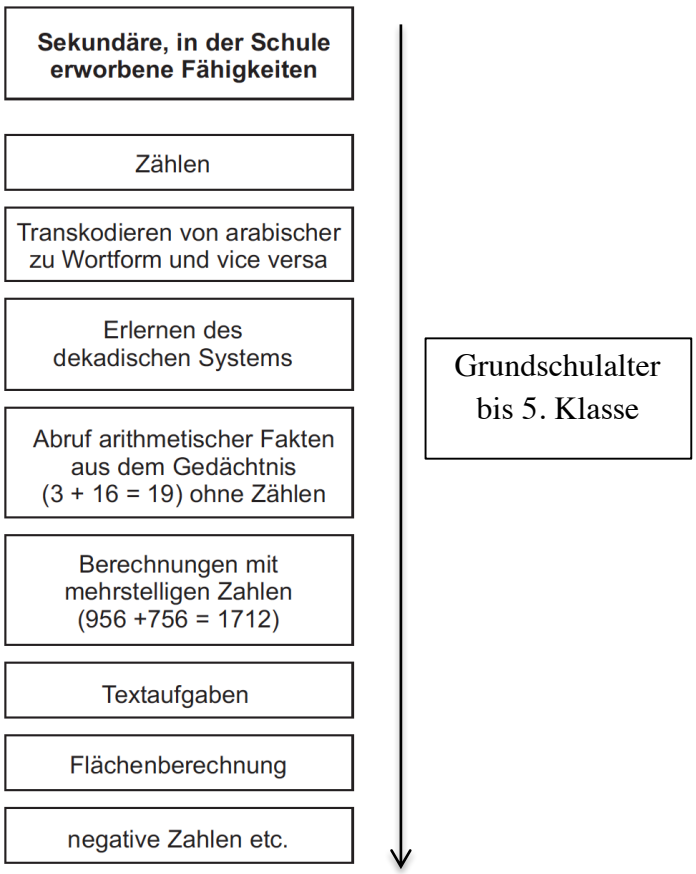


Abbildung 9 Entwicklungsschritte der Rechenfähigkeit (Jacobs & Petermann, 2007)

Das Modell in Abbildung 9 zeigt keine Kompetenzniveaus, wie zum Beispiel differenziert für einzelne Jahrgangsstufen bei Reiss (2004), sondern verdeutlicht, welche Prozesse beim Erwerb der Rechenfähigkeit eher kognitiv-neuronal ablaufen. Zu Beginn der Schulzeit sollte das Zählen sowie die Transkodierungen zwischen dem auditiv-sprachlichen und dem visuell-arabischen Repräsentationsmodus erlernt und gefestigt werden. Die auditiv-verbale Repräsentation wird durch Zählübungen und Kopfrechenaufgaben und die visuell-arabische Repräsentation durch schriftliche Rechenverfahren gefördert und differenziert. Die sichere Transformation zwischen den beiden Modi gelingt im Regelfall bis zum Ende der vierten Klasse. Im Verlauf der Grundschulzeit kommen Überschlagsrechnungen oder Schätzaufgaben hinzu, die zu einer Weiterentwicklung der analogen Größenrepräsentation führen (Jacobs & Petermann, 2007, S. 30). Dieses Modul mit seinem mentalen Zahlenstrahl entwickelt sich im Vergleich zu dem auditiv-sprachlichen und visuell-arabischen Modul erst beim Eintritt in die Schule bzw. in den ersten drei Schuljahren.

Bei komplexen mathematischen Aufgaben kooperieren die Module untereinander. Wenn Aufgaben mit großen Zahlen wie dreihundertfünfundfünfzig plus zweitausenddreiundvierzig verbal vorgegeben werden (Aufnahme in der auditiv-verbalen Repräsentation), wird sie vor dem geistigen Auge in Zahlen umgewandelt (355 + 2043; visuell-arabische Repräsentation) und die Angemessenheit des Ergebnisses (≈ 2400) letztlich durch die analoge Größenrepräsentation geprüft (Dehaene, 1992, S. 31).

Zunehmende Rechenfertigkeiten führen letztlich dazu, dass Schülerinnen und Schüler schneller und akkurater rechnen können. Dabei zeigt sich, dass zum Beispiel beim Zählen das prozedurale Wissen (wie?) dem konzeptuellen Wissen (warum?) voraus sein kann. Diese Reihenfolge kann aber bei anderen Aufgaben – wie beim Addieren mehrstelliger Zahlen – andersherum sein, wie Rittle-Johnson und Siegler (1998) in ihrem Review zeigen.

In dem Modell ist die Wahl des Begriffs „Entwicklungsschritte“ etwas unpräzise gewählt, da es sich eher um einen curricular gesteuerten Kompetenz*erwerb* handelt und weniger um eine Kompetenz*entwicklung*. Zudem bleibt die Frage offen, auf welcher empirischen Grundlage die Autoren die Entwicklungsschritte ableiten und begründen.

Das Arbeitsgedächtnis und die Die Cognitive-Load-Theorie

Weitere Hinweise zu kognitiven Prozessen im Rahmen mathematischer Entwicklung geben das Modell des Arbeitsgedächtnisses und die Cognitive-Load-Theorie.

Das Arbeitsgedächtnis dient dazu, Informationen aufzunehmen, zu speichern und zu verarbeiten (Gold, 2011, S. 34). Es tritt zum Beispiel in Kraft, wenn eine Zahlenreihe gemerkt werden soll und nebenbei nach einem Notizblock und einem Stift gesucht wird, um diese zu notieren (Gerrig, 2015, S. 248). Baddeley unterscheidet in seinem „weithin akzeptierte[n]“ (Gold, 2011, S. 117) Modell des Arbeitsgedächtnisses vier Komponenten: die phonologische Schleife, den visuell-räumlichen Notizblock, den episodischen Puffer und die zentrale Exekutive (Baddeley, 2010a, 2010b).

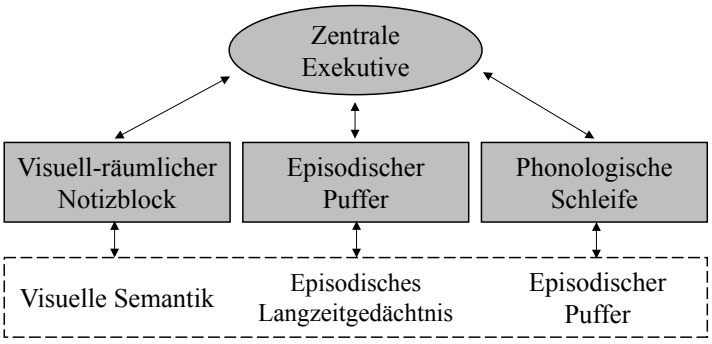


Abbildung 10 Arbeitsgedächtnismodell nach Baddeley (Baddeley, 2010a, 2010b) (grau: Arbeitsgedächtniskomponenten)

Die phonologische Schleife besteht aus einem phonologischen Speicher und einem Mechanismus des stillen Wiederholens (Rehearsal) (Baddeley, 2010a, S. 27). Die phonologische Schleife verarbeitet sprachbasierte Informationen, die durch den phonologischen Speicher aufgenommen und weitergeleitet werden. Um zu verhindern, dass die aufgenommenen Informationen zerfallen, können sie durch ein inneres Sprechen wiederholt und aufrechterhalten werden (Rehearsal). Gesprochene Informationen können dabei unmittelbare in die phonologische Schleife aufgenommen werden, geschriebene Informationen (zum Beispiel „5“), müssen jedoch zuvor

in sprachliche Informationen kodiert werden. Die phonologische Schleife spielt in den Prozessen des Arbeitsgedächtnisses eine große Rolle. „It is […] indeed the most important component of working memory" (Baddeley, 2010b, S. 48).

Der visuell-räumliche Notizblock ist die zweite Komponente des Arbeitsgedächtnismodells. Dieser ist für die Verarbeitung visueller und räumlichen Informationen zuständig. Der visuell-räumliche Notizblock besteht aus zwei Arbeitsbereichen: Einem visuellen Speicher für die Verarbeitung, zum Beispiel von Formen und Farben, und einem Mechanismus für die Aufnahme räumlicher Bewegungssequenzen (Baddeley, 2010a, S. 33).

Eine dritte Komponente ist der episodische Puffer. Er wird von der zentralen Exekutive gesteuert und ruft episodische Informationen aus dem Langzeitgedächtnis ab (zum Beispiel die Erinnerung an eine Mathematikklausur). Er speichert multimodal visuelle und phonologische Informationen und ermöglicht einen Vergleich der Informationen zur aktuellen Situation mit denen aus dem Langzeitgedächtnis. Durch die Kombination der Informationen wird eine integrierte Interpretation der vorliegenden Situation, zum Beispiel im Sinne einer Problemlösung, möglich (Gerrig, 2015, S. 248)

Die vierte Komponente ist die zentrale Exekutive. Sie ist als übergeordnete Instanz für die Kontrolle der Aufmerksamkeit zuständig und verknüpft die Informationen aus der phonologischen Schleife und dem räumlich-visuellen Notizblock. Insbesondere wenn Informationen parallel verarbeitet werden, ist die zentrale Exekutive im Einsatz (Gerrig, 2015, S. 247). Die zentrale Exekutive kann als diskriminierender Filter verstanden werden, der die – für die vorliegende Situation – relevanten Informationen an die Teilbereiche des Arbeitsgedächtnisses weiterleitet.

Das Modell des Arbeitsgedächtnisses ist in Fragen des gelingenden Erwerbs mathematischer Kompetenzen sehr bedeutsam (Lambert, 2015, S. 81). Die Darstellung der kognitiven Prozesse im Arbeitsgedächtnis verdeutlicht, dass sehr komplexe Vorgänge zwischen Informationsaufnahme, Verarbeitung und der endgültigen Handlung geschehen. Studien zur Gedächtnisleistung von Schülerinnen und Schülern mit einer Beeinträchtigung im Bereich des Lernens weisen Probleme in allen Bereichen des Arbeitsgedächtnisses nach (Alloway, Gathercole, Adams & Willis, 2005; Gathercole & Pickering, 2001; Mähler, 2007; Schuchardt, Gebhardt & Mähler, 2010). Aber auch Schülerinnen und Schüler mit Schwierigkeiten im Erwerb mathematischer Kompetenzen zeigen oftmals Defizite im Arbeitsgedächtnis (Lambert, 2015, S. 85). Hinweise auf erfolgreiches mathematisches Lernen kann die Cognitive-Load-Theorie geben, die mit dem Modell des Arbeitsgedächtnisses zusammenhängt. Die Cognitive-Load-Theorie beschäftigt sich mit der kognitiven Belastung beim (mathematischen) Lernen und spricht dabei dem Arbeitsgedächtnis eine zentrale Rolle mit begrenzter Kapazität für die Prozesse der Informationsverarbeitung zu.

Unterschieden wird zwischen der intrinsischen Cognitive Load, der extrinsischen und der germanen Cognitive Load (Rey & Nieding, 2010, S. 70). Der extrinsische Cognitive Load, auch als lernirrelevante kognitive Belastung bekannt, bezieht sich auf die Art der Darbietung des Lernmaterials, das durch den visuell-räumlichen Notizblock aufgenommen wird. Eine hohe extrinsische kognitive Belastung kann dabei zum Beispiel durch überfrachtet gestaltete Arbeits- und Lernmaterialien erfolgen (Born & Oehler, 2013; Busch, Oranu, Schmidt & Grube, 2013). Der intrinsische Cognitive Load bezieht sich hingegen auf das Lernmaterial selbst. Einmaleins-Reihen stellen zum Beispiel einen geringen intrinsischen Cognitive Load dar, mehrschrittige Problemlöseaufgaben hingegen einen hohen. Das Nicht-Abrufen-Können von vorhandenen

Strategien oder Schemata zum Anknüpfen neuen Wissens durch die zentrale Exekutive führt zu einer kognitiven Belastung und hemmt die Lernfähigkeit (Lebens & Lauth, 2014).

> „Gute mathematische Kompetenzen, die mit effektiven Strategien und einem schnellen Abruf von Rechenfakten einhergehen, beanspruchen die Speicher- und Verarbeitungskapazität deutlich weniger. Dagegen wird die Kapazität des Arbeitsgedächtnisses stark belastet oder sogar überfordert, wenn effektive Strategien fehlen. Dies auch dann, wenn grundsätzlich gute Arbeitsgedächtniskapazitäten vorhanden sind." (Fritz & Ricken, 2008, S. 21)

Der germane Cognitive Load versucht den extrinsischen Cognitive Load so weit wie möglich zu reduzieren und ist für die Konstruktion und Automatisierung von Schemata ins Langzeitgedächtnis zuständig.

Die Cognitive-Load-Theorie kann letztlich Hinweise darüber geben, wie mathematisches Lernen gerade bei Schülerinnen und Schülern mit Lern- und/oder Rechenschwierigkeiten erleichtert werden kann.

2.3.3 Bedeutung mathematischer Basiskompetenzen für den weiteren Kompetenzerwerb

International und auch im deutschsprachigen Raum fehlen aktuell Langzeitstudien, die die Bedeutung mathematischer Lerninhalte der Grundschule für den weiteren Kompetenzerwerb über den Zeitraum der Primarstufe *hinaus* differenziert betrachten (Pekrun et al., 2006). Zahlreiche Studien betrachten den Kompetenzerwerb oftmals *vor* bzw. *nach* und nicht *inklusive* des Übergangs in die Sekundarstufe I (Ufer, 2009, S. 93). Konsens besteht darüber, dass der Mathematikunterricht der Primarstufe ein wichtiges Fundament für den weiteren Kompetenzerwerb darstellt. „Ein weiterführendes mathematisches Verständnis [ist] das Ergebnis eines kumulativen Lernprozesses" (Humbach, 2008, S. 67), sodass ein gefestigtes Verständnis der Grundschulmathematik unumgänglich für einen gelingenden mathematischen Kompetenzerwerb in der Sekundarstufe I ist. Walther und Kollegen (2003, S. 189 f.) betonen:

> „Die in der Grundschule zu erwerbenden grundlegenden mathematischen Konzepte und Prozesse besitzen einen hohen Stellenwert für den darauf aufbauenden Mathematikunterricht von weiterführenden Schulen. In den beiden TIMS-Studien und in PISA wurden zum Teil Besorgnis erregende Defizite in den mathematischen Leistungen von Schülern des Sekundarbereichs zu Tage gebracht. In diesem Zusammenhang wurde die Vermutung geäußert, dass diese Schwächen nicht allein aus der Sekundarstufe resultieren, sondern ihre Wurzeln zum Teil bereits in der Grundschule haben."

International wird diese Auffassung geteilt, wenn Duncan und Kollegen (2007, S. 1429) schreiben: „Academic achievement is a cumulative process involving both mastering new skills and improving already existing skills". Das Verständnis des systematischen Rechnens und das Regelwissen in der Grundschulzeit bilden eine Grundlage für komplexe Rechenoperationen im weiteren Mathematikunterricht (Gaidoschik, 2008; Grube, 2005; Moser Opitz, 2007).

Das Kompetenzstufenmodell nach Reiss (2004) und die o. a. Bildungsstandards zeigen, dass die Vermittlung der arithmetischen Kompetenzen ($+, -, \cdot, :$) in den Klassenstufen eins bis vier eine zentrale Aufgabe der Grundschulmathematik darstellt (vgl. Schneider u. a., 2013, S. 100). Im Laufe ihrer Entwicklung sollten die Schülerinnen und Schüler in der Lage sein, Additions- und Subtraktionsaufgaben zu lösen, ohne stets bei 1 mit dem Zählen anfangen zu müssen. In

der Grundschule werden verschiedene Strategien geübt und gefestigt und „die spezifische Wissensbasis entwickelt und ausgebaut“ (Schneider et al., 2013, S. 101). In einem nächsten Entwicklungsschritt werden Lösungen aus dem Gedächtnis abgerufen ohne zu rechnen, was ein Indiz für die Automatisierung des Zahlenbereiches darstellt. Auch das schriftliche Rechnen hat in der Grundschule eine große Bedeutung. „Die schriftlichen Rechenverfahren gehören zu den ersten Beispielen für Algorithmen im schulischen Mathematikunterricht und sind somit von exemplarischer Bedeutung für den gesamten algorithmisch-kalkülhaften Teil der Schulmathematik“ (Gerster, 2009, S. 269). In der Grundschule werden auch bereits Flächen- und Raumvorstellungen angebahnt. In der Sekundarstufe werden diese durch „konkrete Mess- und Vergleichsaktivitäten“ (Schneider et al., 2013, S. 115) sowie durch Formelanwendungen vertiefend betrachtet. „Das Verständnis des Dezimalsystems und der Grundoperationen müssen als Grundlage für ein verständiges Rechnen angesehen werden, gleichzeitig zeigt sich das *Verständnis* der Grundoperationen als bedeutsamer Schlüssel für einen kompetenten Umgang mit Sachaufgaben“ (Freesemann, 2014, S. 91 Herv. d. Verf.).

Gefestigte mathematische Basiskompetenzen spielen nicht nur aus Gründen des kumulativen Lernens eine bedeutsame Rolle, sondern insbesondere hinsichtlich der Entlastung des Arbeitsgedächtnisses für die Anforderungen in der Sekundarstufe I. Ennemoser und Krajewski (2013, S. 230) weisen darauf hin, dass ein ungesichertes Konventions- und Regelwissen gerade für leistungsschwache Schülerinnen und Schüler hinsichtlich komplexer Aufgabenanforderungen in der weiterführenden Schule ein Risiko darstellt. Schwache Rechnerinnen und Rechner weisen oftmals „leistungslimitierende Funktionen des Arbeitsgedächtnisses“ auf (Ennemoser & Krajewski, 2013, S. 230). Das für eine Aufgabe notwendige Regel- und Konventionswissen muss mühevoll abgerufen werden, sofern es nicht gefestigt ist (siehe auch Grube, 2006; Landerl & Kaufmannn, 2013, S. 146).

> „Da die gebundenen Arbeitsgedächtnisressourcen nicht für anspruchsvolle Prozesse, etwa im Sinne mathematischen Problemlösens, zur Verfügung stehen, kann dies zu einer Beeinträchtigung des eigentlich anvisierten Lernerfolgs führen.“ (Ennemoser & Krajewski, 2013, S. 230)

Ennemoser und Krajewski (2013) beziehen sich mit ihrer Aussage auf die oben vorgestellte Cognitive-Load-Theorie (Ayres & Paas, 2012). Ebenso im Sinne der Cognitive-Load-Theorie lassen sich Ergebnisse von Vom Hofe und Kollegen (2009) einordnen. Die Autoren betrachten die Modellierungskompetenz von Sekundarstufen-I-Schülerinnen und -Schüler und stellen fest: „Insgesamt lassen sich etwa die Hälfte aller analysierten Schüler-Fehler auf unvollständige ausgebildete oder fehlende Grundvorstellungen zurückführen“ (Vom Hofe et al., 2009, S. 143). Unter „Grundvorstellungen“ fassen die Autoren etwa die Vorstellung und Wahl von Operationen. Die Untersuchung bekräftigt die Notwendigkeit gut ausgebildeter Basiskompetenzen in der Grundschulzeit. Ennemoser und Krajewski (2013, S. 230) bezeichnen das Konventions- und Regelwissen als „vernachlässigte Basiskompetenz“, die einen großen Einfluss auf den weiteren mathematische Kompetenzerwerb hat.

Siegler und Kollegen (2012) ziehen für die Frage, inwiefern frühe schulische mathematische Leistungen gar die Kompetenzen in der High-School vorhersagen können, zwei umfassende Längsschnittstudien heran: Die British Cohort Study (BCS) und die Panel Study of Income Dynamics-Child Development Supplement (PSID-CDS). In den beiden Studien werden jeweils mathematische Kompetenzen von Schülerinnen und Schülern im Alter von ca. 10 bis 16 Jahren erhoben. Die BCS umfasst dabei 3677 und die PSID-CDS 599 Schülerinnen und Schüler. Die

Forschergruppe findet in ihrer Analyse heraus, dass Kompetenzen im Bruchrechnen und im Geteilt-Rechnen mit ganzen Zahlen spätere Leistungen am besten vorhersagen. Verglichen werden sie dabei mit weiteren Prädiktoren wie das Plus-, Minus- oder Malrechnen, die Intelligenz, die Arbeitsgedächtniskapazitäten oder familiäre bzw. soziale Faktoren. Begründet werden die Erkenntnisse durch hohe Korrelationen in der Highschool zwischen allgemeinen mathematischen Kompetenzen und Bruchrechnen (r = .81 (BCS) bzw. r = .87 (PSID-CDS)) und ebenso hohen Korrelationen zwischen Bruchrechnen und den Kompetenzen im Bereich Algebra (r = .68 (BCS) bzw. r = .65 (PSID-CDS)). Auch Reiss und Kollegen (2007, S. 111) stellen heraus, dass die mathematische Kompetenz in der Grundschulzeit ein guter Prädiktor für die mathematische Kompetenz zu einem späteren Zeitpunkt der Schullaufbahn ist. Gravierend ist hingegen, dass die mathematischen Kompetenzen am Ende der Primarstufe in einer Klasse oftmals schon weit auseinanderklaffen (Bos & Gröhlich, 2010; Mittelberg, 2004; Van Ophuysen & Wendt, 2009; Pekrun et al., 2006) und sich diese Unterschiede in der Sekundarstufe I manifestieren (Becker, Lüdtke, Trautwein & Baumert, 2006).

Durlak (1997a, S. 74) betont auch die Bedeutsamkeit guter Schulleistungen für die langfristig stabile emotional-soziale Entwicklung: „Finally, there is a growing evidence that good academic performance is a protective factor in young people's lives that can lessen the future likelihood of negative outcomes such as crime and delinquency and various psychological and behavioral problems". Fundiert ausgebildete mathematische Basiskompetenzen sind folglich sowohl für die akademische als auch für die Persönlichkeitsentwicklung von großer Bedeutung.

Nachdem ein umfassender Überblick über die Forschung bezüglich des mathematischen Kompetenzerwerbs in der Grundschulzeit gegeben ist, wird im Folgenden ein Fokus auf den Erwerb mathematischer Kompetenzen im Sekundarstufen-I-Bereich gelegt.

2.4 Erwerb mathematischer Kompetenzen in der Sekundarstufe I

Die Betrachtung des Erwerbs mathematischer Kompetenzen in der Sekundarstufe I ist in Deutschland „spärlich" (Ufer, 2009, S. 92) und muss als Forschungsdesiderat bezeichnet werden (Ehlert et al., 2013; Krajewski & Ennemoser, 2010). Derzeit gibt es kein empirisch abgesichertes Erwerbsmodell oder Kompetenzstufenmodell für die Sekundarstufe I. Einen Überblick über die zu erwerbenden Kompetenzen am Ende der Sekundarstufe I ermöglichen zwar auch hier die Bildungsstandards für den Mittleren Schulabschluss (Kultusministerkonferenz, 2003), jedoch ist bisher kein jahrgangsübergreifendes, an die Primarstufe anknüpfendes Kompetenzstufenmodell (z. B. an Reiss & Winkelmann, 2008) entwickelt. Erschwert wird dies u. a. dadurch, dass bundesweit in der Sekundarstufe I unterschiedliche Schulformen besucht werden, für deren Unterricht Lehrkräfte spezifisch und auf Grundlage verschiedener Kerncurricula ausgebildet werden (Ufer, 2009, S. 94).

Hinweise auf Erwerbsprozesse mathematischer Kompetenzen in der Sekundarstufe I stammen von Krajewski und Ennemoser (2010). Ausgehend von der Annahme, dass die Ausbildung der frühen Vorläuferfertigkeiten gemäß des ZGV-Modells (vgl. 2.2.2) Grundlage für den weiteren Kompetenzerwerb ist, stellen die Autoren fest, dass in der Sekundarstufe I in höheren Zahlenräumen eben diese drei Ebenen noch einmal durchlaufen werden. Sprich: Wenn ein Schüler oder eine Schülerin in der Vorschulzeit exakte Zahlenfolgen benennen kann, bedeutet dies nicht, dass dies auch automatisch in einem Zahlenraum bis 1.000.000 beherrscht wird (Ebene

1) oder dass auch Größer- und Kleiner-Vergleiche im großem Zahlenraum ebenso gut gelingen (Ebene 2). Krajewski und Ennemoser (2010) sichern diese Überlegungen durch eine Studie mit 442 Schülerinnen und Schülern der fünften bis achten Klasse ab.

Im Folgenden werden die mathematischen Kompetenzen in der Sekundarstufe I im Zentrum der Betrachtung stehen. Nachdem nationale und internationale Erkenntnisse zum Übergang in die Sekundarstufe I dargelegt sind, werden sowohl Längsschnitt- als auch Querschnittstudien angeführt, die sich um eine Abbildung des Kompetenzerwerbs in der Sekundarstufe I bemühen. Dabei wird der Frage nachgegangen, ob sich der sog. Matthäus-Effekt („Denn wer da hat, dem wird gegeben“) (siehe Reichenbach, 2010) auch hier bewahrheitet.

2.4.1 Der Prozess der Transition in die Sekundarstufe I

Der Zeitpunkt des Übergangs in die Sekundarstufe I divergiert zwischen vielen Ländern, was die Vergleichbarkeit von Studien zum Prozess der Transition insgesamt erschwert. In Deutschland erfolgt der Übergang in die Sekundarstufe I in der Regel nach der Primarstufe, die mit der vierten Klasse endet. Ausnahme sind Berlin und Brandenburg, wo sie sechs Jahre umfasst. In den USA dauert die Primary School bis zur sechsten Klasse, in Irland oder in den Niederlanden dauert sie acht Schuljahre und in Belgien wiederum sechs Schuljahre (Schmitt, 2001). Der Schwerpunkt dieses Kapitels 2.4.1 wird primär auf Erkenntnissen zum Prozess der Transition im *deutschen* Schulsystem liegen. Dabei werden jedoch auch internationale Erkenntnisse einbezogen.

Aktuell besteht eine differenzierte Forschungsgrundlage zum Prozess des Übergangs in die Sekundarstufe I, in der verschiedene Facetten wie zum Beispiel die akademische Leistungsentwicklung, die Selbsterwartungshaltung der Schülerinnen und Schüler oder auch methodisch-didaktische Praktiken der Lehrkräfte untersucht werden. Dabei werden oftmals Problemfelder benannt, die im Übergang auftauchen und diesen erschweren. So schreibt Weißbach (1986) bereits vor einigen Jahrzehnten vom sog. „Sekundarstufenschock“ im deutschen Schulsystem.

Im Begriff der Transition sieht Braun (2011, S. 43) drei zu unterscheidende Dimensionen: die räumliche, die qualitative (sprich: akademische) und die soziale Dimension. Im Folgenden soll der Schwerpunkt auf der qualitativen Dimension, auf der schulischen Domäne Mathematik, liegen, „wenn es darum geht, vorhandene Fähigkeiten auf einem anderen Niveau als bislang weiterzuentwickeln“ (Braun, 2011, S. 43). Dabei darf den anderen beiden Dimensionen, die sich zum Beispiel mit der Rolle von Freundschaft und Peers beschäftigen (Krüger, Köhler, Pfaff & Zschach, 2007), nicht weniger Aufmerksamkeit zukommen.

Griebel und Niesel (2011) sehen im Kontext der Transition in die Sekundarstufe I auf drei Ebenen umfassende Entwicklungsaufgaben (siehe Abbildung 11), die sie auf Grundlage aktueller Forschungen ableiten: Auf der kontextuellen Ebene beziehen sich die Entwicklungsaufgaben des Schülers oder der Schülerin auf die Einstellung auf der sich grundlegend veränderten Rahmenbedingungen (größere Schule, größere Klassenstärke, anderer Schulweg, neue Regeln). Zudem lernt der Schüler oder die Schülerin das Fachlehrerprinzip kennen und muss größere Anforderungen bzgl. Hausaufgaben und Leistungstests bewältigen. Auf der interaktionalen Ebene wird er oder sie sich auf veränderte Beziehungen in der Schülergruppe einstellen, soziale Netz-

werke aufbauen und pflegen, sich mit seiner oder ihrer neuen Rolle im Gruppengefüge auseinandersetzen und eine größere Selbstständigkeit im Arbeitsverhalten aufbauen müssen. Auf der individuellen Ebene sollte der Schüler oder die Schülerin ein problemlösendes Bewältigungsverhalten einüben, seine bzw. ihre Schlüsselkompetenzen optimieren (emotionale Kompetenzen, Kommunikations- und Kooperationsfähigkeit), die Identität als Schüler bzw. Schülerin der Sekundarstufe I erwerben und das Zugehörigkeitsgefühl dazu entwickeln.

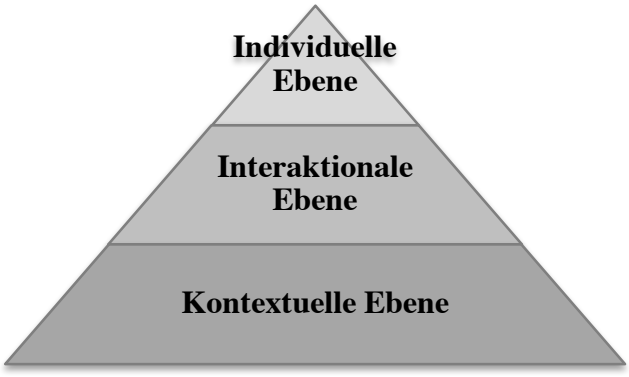


ABBILDUNG 11 ENTWICKLUNGSAUFGABEN IN DER TRANSITION IN DIE SEKUNDARSTUFE I (GRIEBEL & NIESEL, 2011, S. 208 FF)

Entwicklungsaufgaben sehen Griebel und Niesel (2011) jedoch nicht nur auf Seiten des Schülers oder der Schülerin, sondern auch auf der Seite der Eltern und der Lehrkräfte als im sozialen System beteiligte Akteure. „Nur durch Kommunikation und Beteiligung des Kindes und aller Personen, die an seiner Bildung und Erziehung beteiligt sind, kann ein Kind dahin gelangen, dass es sich in der neuen Umgebung wohlfühlt und die Bildungsangebote bestmöglich nutzt“ (Griebel & Niesel, 2011, S. 207).

Prognostisch zeigen zahlreiche ältere und auch jüngere Studien, dass Schülerinnen und Schüler ihre Noten der Primarstufe in der Sekundarstufe I oftmals nicht halten können (Böck & Migutsch, 1974; Eder, 1995; Isakson & Jarvis, 1999; McGee, Ward, Gibbons & Harlow, 2003; Sirsch, 2000). Dies scheint unabhängig davon zu sein, zu welchem Zeitpunkt der Übergang stattfindet. McGee und Kollegen (2003, S. 3) schreiben in ihrem Literatur-Review:

> „Some studies reported on transition at 10 or 11 years old [...], others with a transition at 13 or 14 years old [...], and they have similar findings, a drop in academic achievement. Students who make two transitions, to middle school and then to secondary school, appear to experience the drop twice. [...] There appears to be evidence that any transition may cause a drop in achievement regardless of the age at which it takes place.“

Die Gründe für eine Verschlechterung der Noten sind vielfältig und erweisen sich als konstant. „Vieles trifft die Kinder beim Übergang vom vierten zum fünften Schuljahr unerwartet: neue Maßstäbe, auf die sie anscheinend nicht genügend vorbereitet sind, neue soziale Beziehungen u. a.“ (Krause, 1987, S. 60). International wird oftmals die Pubertät bzw. das Erwachsenwerden als Grund für eine Verschlechterung der Schulleistungen nach der Transition herangezogen (Potter, 2001), da die Primarstufe i. R. bis zur sechsten oder gar achten Klasse dauert (siehe oben). Es liegen hingegen auch Studien vor, die keine Verschlechterung der schulischen Noten nach dem Übergang feststellen (Anderman & Midgley, 1997; Sirsch, 2000).

Die Notenentwicklung in der Sekundarstufe I steht jedoch selten zuverlässig für den tatsächlichen Kompetenzerwerb. Oftmals besuchen Schülerinnen und Schüler weiterführende Schulen,

die ein zu hohes Anforderungsniveau für sie haben (Liegmann, 2014), sodass sie dort eher schlechtere Noten erhalten. Zudem bemüht sich nicht jede Lehrkraft um eine diagnostische Erfassung der Lernausgangslage, was sich auch in der weit auseinanderklaffenden Vorstellung von Differenzierung widerspiegelt (Racherbäumer, 2014).

Wie schätzen Lehrkräfte die Kompetenzen der Schülerinnen und Schüler nach der Transition ein? Racherbäumer und Kohnen (2014) befragen Sekundarstufen-I-Lehrkräfte zu den Lernvoraussetzungen ihrer Schülerinnen und Schüler im Fach Mathematik zu Beginn der fünften Jahrgangsstufe. Hinsichtlich der Grundrechenarten schätzen die befragten Lehrkräfte ($n_{Hauptschule} = 30$; $n_{Realschule} = 42$), dass lediglich ca. die Hälfte der Hauptschülerinnen und -schüler und nahezu 75 % der Realschülerinnen und -schüler Aufgaben aller vier Grundrechenarten lösen können. Noch geringere Lernvoraussetzungen schreiben die Lehrkräfte den Schülerinnen und Schülern in den Bereichen Geometrie, Größen und Einheiten als auch Mengenerfassung im Zahlenraum bis 1.000.000 zu. Die Lehrkräfte gehen offenbar davon aus, dass ein nicht geringer Teil der Schülerinnen und Schüler zu Beginn der Sekundarstufe I *nicht* über die notwendigen mathematischen Basiskompetenzen verfügt. In einer weiteren Untersuchung von Racherbäumer (2014) geben 40 % der Hauptschullehrkräfte (n = 57), 60 % der Realschullehrkräfte (n = 101) und 30 % der Gymnasiallehrkräfte (n = 248) an, die Lernausgangslage ihrer Schülerinnen und Schüler im Fach Mathematik zu Beginn der fünften Klasse durch standardisierte oder nicht standardisierte Verfahren zu erheben. Ufer (2009, S. 94) fordert von Lehrkräften eine kohärente Sichtweise zwischen mathematischer Kompetenz und Mathematikunterricht, um eine „wichtige Grundlage für eine kontinuierliche Entwicklung mathematischer Kompetenzen beim Übertritt von der Primarstufe in die Sekundarstufe“ zu bilden. Jedoch stellt er zudem fest: „[...] bereits die Ausbildung der Lehrkräfte [ist] oft stark auf die Schulstufe und Schulform fokussiert [...], in der die Lehrkräfte später eingesetzt werden sollen. So ist es keine Selbstverständlichkeit, dass Gymnasial- und Realschullehrkräfte einen Überblick über Inhalte und Anspruchsniveau des Grundschullehrplans haben“ (Ufer, 2009, S. 94).

Mit dem Übergang in die Sekundarstufe I aus Sicht der Schülerinnen, Schüler *und* Lehrer beschäftigt sich 1986 das Hagener Übergangsprojekt. In Form von Kleingruppeninterviews werden Problemstellen des Übergangs in die Sekundarstufe I identifiziert. Die interviewten Schülerinnen und Schüler in der fünften Klasse benennen u. a. folgende Faktoren als „übergangserschwerend“: das Fachlehrerprinzip, Gruppenbildung und Rollenprozesse, die große Zahl der neuen Fächer, die Häufung von Klassenarbeiten, inhaltliche Brüche, die ungewohnte Größe des neuen Schulsystems, die ungewohnte Schülerzahl, die Nicht-Gestaltung der Schulräume, die längere Zeit des Schultages und auch der größere Umfang an Hausaufgaben. Insgesamt zeigt sich jedoch, dass der Großteil der Schülerinnen und Schüler den Übergang gut bewältigt. Schwieriger fällt der Übergang hingegen den eher unsicheren Schülerinnen und Schülern und denjenigen, die alleine (ohne bekannte soziale Kontakte) in eine neue Klasse kommen (Mitzlaff & Wiederhold, 1989). Hinsichtlich des Übergangs aus Lehrersicht wird deutlich, dass sich das Problembewusstsein interindividuell massiv unterscheidet. Einige Lehrkräfte bemühen sich dabei um einen intensiven und kontinuierlichen Bildungsprozess der Schülerinnen und Schüler, andere hingegen leugnen die Übergangsprobleme oder „[erklären] übergangsbedingte Anpassungskriterien und Diskontinuitäten zu ‚natürlichen' Erscheinungen oder zu gesunden Selektionsinstrumenten“ (Mitzlaff & Wiederhold, 1989, S. 31). Die sich an das Hagener Übergangsprojekt anschließende Marburger Übergangsstudie (Koch, 2001) stellt erneut den Übergang aus

Sicht der Lehrkräfte in das Zentrum. Auf Grundlage von 22 Interviews ($n_{Grundschullehrkräfte} = 12$; $n_{Sekundarstufenlehrkräfte} = 10$) und 191 Fragebögen ($n_{Grundschullehrkräfte} = 87$; $n_{Sekundarstufenlehrkräfte} = 104$) stellt Koch (2001) – entgegen den Ergebnissen des Hagener Übergangsprojekts – heraus, dass Grundschul- und Sekundarstufenlehrkräfte durchaus eine vergleichbare Vorstellung davon haben, wie der Übergang in die Sekundarstufe I gestaltet werden sollte und welcher pädagogischer Auftrag dabei auf beiden Seiten zu realisieren ist (Zusammenarbeit, „sanfte" Gestaltung des Übergangs, routinierte Vorbereitung u. a.). Es zeigt sich, dass der Unterricht in Grundschule und Sekundarstufe I unterschiedlich gestaltet ist und ein unterschiedliches Erziehungsverständnis vorliegt. Koch (2001, S. 151) zieht aus ihrer Forschung das Fazit „Den Übergang gemeinsam gestalten!", damit dieser für die Schülerinnen und Schüler ein weniger einschneidendes Erlebnis wird.

2.4.2 Erwerb mathematischer Kompetenzen in der Sekundarstufe I im Längsschnitt

Wie verläuft der mathematische Kompetenzerwerb nach der Transition in die Sekundarstufe I? Es liegen sowohl Befunde dazu vor, dass die Transition mit Leistungsrückgängen einhergehen kann, als auch Hinweise dafür, dass die Transition ohne Probleme vollzogen wird (siehe Kapitel 2.4.1). Eindeutig ist die Befundlage dahingehend, dass der Übergang für einen Teil der Schülerinnen und Schüler ein kritisches Ereignis in der Lerngeschichte darstellt. Das folgende Unterkapitel gibt einen Überblick über nationale und internationale Längsschnittuntersuchungen zum Erwerb mathematischer Kompetenzen beim Übertritt in die Sekundarstufe I.

In diesem Kapitel 2.4.2 handelt es sich eher um bildungswissenschaftlich angelegte Studien, die einen allgemeinen (quantitativen) Überblick über den mathematischen Kompetenzerwerb geben. Stärker qualitative Befunde zu den Kompetenzen werden im Kapitel 2.4.3 (Nationale Querschnittstudien) dargestellt.

2.4.2.1 Nationale Studien

Aktuell sind einige deutschsprachige Längsschnittstudien zu verzeichnen, die sich mit dem Kompetenzerwerb im Bereich Mathematik *während* bzw. *nach* dem Übergang in die Sekundarstufe I beschäftigen. Die hier vorgestellten Studien umfassen den Zeitraum von 1984 (Beginn der PALMA-Studie) bis 2009 (Ende der KESS-Studie).

Die einbezogenen Studien sind in der Abbildung 12 grafisch in ihrem zeitlichen Verlauf dargestellt. Beginnt eine Datenmessung auf der X-Achse bereits bei 0, meint dies, dass die Datenerhebung bereits in der Vorschulzeit beginnt.

Die vorliegenden Studien erfüllen folgende Kriterien:

- Das Design entspricht einem längsschnittlichen Design (mindestens zwei aufeinander folgende Datenpunkte).
- Es wird der Übergang von der Primar- in die Sekundarstufe I berücksichtigt.
- Es werden – auch – mathematische Kompetenzen erhoben.
- Es werden Schülerinnen und Schüler an sog. „nicht gymnasialen Schularten" (Hammer et al., 2016, S. 237) berücksichtigt.

- Die Untersuchung selbst stammt aus Deutschland oder Deutschland ist an ihr beteiligt.

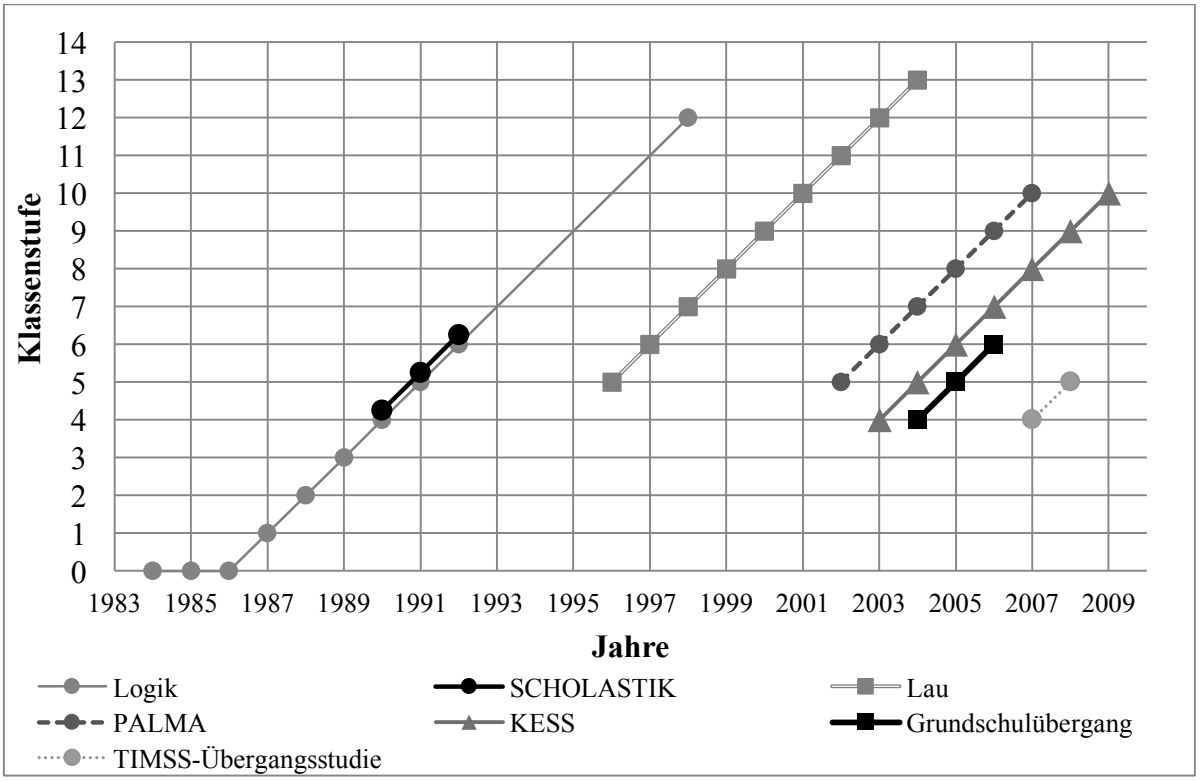


ABBILDUNG 12 ÜBERBLICK ÜBER NATIONALE LÄNGSSCHNITTSTUDIEN ZUM ERWERB MATHEMATISCHER KOMPETENZEN IN DER SEKUNDARSTUFE I

Nicht einbezogen sind in diesem Kapitel die BIJU-Studie (Bildungsverläufe und psychosoziale Entwicklung im Jugendalter), da diese mit der Erhebung mathematischer Kompetenzen erst in der Klasse sieben startet und deswegen nur wenige Aussagen zum Übergang enthält. Ebenso nur erwähnt werden soll an dieser Stelle die Berliner ELEMENT-Studie (Erhebung zum Lese- und Mathematikverständnis in den Jahrgangsstufen vier bis sechs) (Lehmann & Lenkeit, 2008), die hier jedoch nicht weiter beschrieben wird, da sie ausschließlich Gymnasialschülerinnen und -schüler berücksichtigt. Ebenso nicht berücksichtigt wird die PING-Untersuchung (Pilotprojekt Inklusive Grundschule) (Friedrich & Laag, 2015), die nur den Kompetenzerwerb innerhalb der Grundschulzeit betrachtet.

In einem späteren Kapitel werden deutschsprachige Querschnittstudien angeführt, die sich speziell mit Kompetenzen in verschiedenen Jahrgangsstufen der Sekundarstufe I beschäftigen.

Einen Überblick über die Inhalte, die Zielgruppe und die zentralen Ergebnisse der Studien ermöglicht die folgende Tabelle 5. Die Angabe in der Spalte N bezieht sich auf die Stichprobe zum jeweils ersten Messzeitpunkt und berücksichtigt keinen Drop-Out für den Verlauf der Erhebungen. Die Studien sind chronologisch entsprechend des letzten Erhebungsjahres aufgeführt, sodass oben die jüngste Studie zu finden ist.

TABELLE 5 DETAILANGABEN ZU NATIONALEN LÄNGSSCHNITTSTUDIEN ZUM ERWERB MATHEMATISCHER KOMPETENZEN IN DER SEKUNDARSTUFE I

Studie	Literatur-verweis	Lauf-zeit	Klassen-stufen	N	Variablen	Zentrale Ergebnisse
Kompetenzen und Einstellungen von Schülerinnen und Schülern (KESS)	Bos & Gröhlich (2010)	2003 – 2009	4–10/11	14.000	• Mathematikleistung • Deutschleistung • Sachunterrichtleistung • Englischleistung • Fachbezogene Einstellungen	• Schereneffekt: Hauptschüler in Klasse acht liegen 1 ½ Jahre hinter den Realschülern und 2–3 Jahre hinter den Gymnasialschülern
TIMS-Übergangsstudie	Bundesministerium für Bildung und Forschung (2010)	2006 – 2008	4 & 5	5.712	• Mathematikleistungen • Naturwiss. Leistungen • Deutschleistungen • Kognitive Leistungsfähigkeit • Hintergrundmerkmale	• Substanzielle Gruppenunterschiede u. a. im Bereich Mathematik zwischen den Schulformen • Herausstellung eher homogener Lerngruppen • Große Überlappungsbereiche in den math. Kompetenzen zwischen den Schulformen
Projekt zur Analyse der Leistungs-entwicklung in Mathematik (PALMA)	Pekrun u. a. (2006)	2002 – 2008	5–10	2.070	• Mathematikleistung • Mathematische Emotionen	• Schereneffekt (sowohl bei Kalkül- als auch bei Modellierungsaufgaben): Gymnasialschüler entwickeln sich stärker als Realschüler und diese besser als Hauptschüler • Zahl an „Risikoschülern“ steigt in der Hauptschule im Laufe der Zeit massiv und am stärksten an • Schüler aller drei Schulformen entwickeln sich über die Schuljahre hinweg grundsätzlich positiv
Dortmunder Projekt Grundschul-übergang	Van Ophuysen & Wendt (2009)	2004 – 2006	4–6	630	• Mathematikleistung • Leseleistung • Schulfreude • Fähigkeitsselbstkonzept	• Grundsätzlich positive Entwicklung der Schüler • Schereneffekt: Gymnasialschüler entwickeln sich stärker als IGS-Schüler, diese besser als Realschüler und diese besser als Hauptschüler

Studie	Literatur-verweis	Lauf-zeit	Klassen-stufen	N	Variablen	Zentrale Ergebnisse
Aspekte der Lernausgangslage und der Lernentwicklung (LAU 5–13)	Lehmann u. a. (2001)	1996 – 2005	5–13	13.099	• Mathematikleistung • Leseverständnis • Sprachverständnis • Rechtschreibung • Leistungen im Fach Englisch • Soziale/demografische Einflüsse	• Anfang Jahrgang 5 befinden sich insgesamt 6 4,5 % auf Kompetenzstufe I & II (von V) • Anfang Jahrgang 7 befinden sich insgesamt ca. 20 % auf allen fünf Kompetenzstufen; Haupt- und Realschüler sind jedoch noch zu 64,7 % auf Stufe I & II • Schwache Schüler erzielen absolut den stärkeren math. Leistungszuwachs • Schüler mit vergleichbarem Niveau entwickeln sich auf dem Gymnasium stärker
Longitudinal-studie zur Genese individueller Kompetenzen (LOGIK)	Weber & Stefanek (1998) Stern (2003) Stern (2008) Stern (2013)	1984 – 1992	Vorschule–Klasse 6 (Follow-Ups: Jugendalter; 24 Jahre; 30 Jahre)	205	• Mathematische Kompetenzen • Schriftsprachkompetenzen • Kognitive Entwicklung • Motorik • Moralische Entwicklung • u. a.	• Hoher Zusammenhang zwischen dem Lösen von Textaufgaben in der zweiten Klasse und dem Lösen von Textaufgaben in der elften Klasse • Mittlerer Zusammenhang zwischen dem numerischen Verständnis im Alter von 5 Jahren und der mathematischen Intelligenz mit 23 Jahren • Es gibt kaum Schüler, die in der zweiten Klasse leistungsschwach und in der elften Klasse leistungsstark, jedoch welche, die in der zweiten Klasse leistungsstark und in der elften leistungsschwach sind
Schulorganisierte Lernangebote und Sozialisation von Talenten, Interessen und Kompetenzen (SCHOLASTIK)		1986 – 1992	1–4 (z. T. 1–6)	1.150	• Mathematikleistung • Leseleistung • Rechtschreibleistung • Motivationale Entwicklung • Selbstkonzept • Unterrichtsmerkmale & Kontextfaktoren	

Anmerkung: Schüler bedeutet hier zugleich Schüler und Schülerinnen.

Insgesamt weisen die aufgeführten Längsschnittstudien darauf hin, dass sich die Leistungen im Bereich Mathematik bei den Schülerinnen und Schülern in der Sekundarstufe I grundsätzlich positiv entwickeln. Die PALMA-Untersuchung zeigt, dass Schülerinnen und Schüler der Haupt- und Realschule sowie des Gymnasiums über die begleiteten sieben Schuljahre hinweg im Regensburger Mathematikleistungstest für 5. bis 10. Klassen eine signifikante Leistungsentwicklung in ihren mathematischen Kompetenzen aufweisen (Pekrun et al., 2006). Neben der PALMA-Studie bestätigt auch das Dortmunder Projekt Grundschulübergang, dass sich die Schülerinnen und Schüler in ihren mathematischen Leistungen im Laufe der Sekundarstufe I insgesamt signifikant verbessern (ES $d_{4\text{-}6} = 1{,}031$) (Van Ophuysen & Wendt, 2009, S. 320). In der LAU-Studie ist jedoch deutlich erkennbar, dass sich in der Klasse fünf ein Großteil (64,5 %) der Schülerinnen und -schüler auf den Kompetenzstufen I und II (von V) befindet, am Anfang der Klasse sieben hingegen verteilt sich die Stichprobe auf alle fünf Niveaustufen mit ungefähr 20 % (+/– 4 %) (Lehmann, Husfeldt, & Peek, 2001, S. 36 ff.). Zur Erhebung der Mathematikleistungen wird der Subtest Mathematik des Hamburger Schulleistungstests für 4. und 5. Klassen (HST 4/5) verwendet.

Trotz einer grundsätzlich positiven Entwicklung sind die Unterschiede zwischen den Schulformen immens. In der TIMS-Übergangsstudie zeigen die Gymnasialschülerinnen und -schüler nach dem Übergang in Klasse fünf die besten Mathematikleistungen und -noten, gefolgt von der Realschulkohorte, den Schulen mit mehreren Bildungsgängen, den Gesamtschulen und abschließend den Hauptschulen (Jonkmann, Maaz, Neumann & Gresch, 2010, S. 142). Auch in der PALMA-Untersuchung zeigen die Gymnasialschülerinnen und -schüler die beste Entwicklung, gefolgt von der Real- und abschließend von der Hauptschulkohorte. Die schulformabhängigen Unterschiede bestätigen sich hier sowohl in der differenzierten Betrachtung der „Kalkülaufgaben“ (Aufgaben, die in erster Linie das Anwenden von Rechenregeln erfordern) als auch in den „Modellierungsaufgaben“ (Pekrun et al., 2006, S. 36). Das Dortmunder Projekt Grundschulübergang findet ebenso einen Schereneffekt zwischen den Schulformen in der Sekundarstufe I (ES $d_{4\text{–}6\ \text{Hauptschule}} = 0{,}956$, ES $d_{4\text{–}6\ \text{IGS}} = 0{,}100$, ES $d_{4\text{–}6\ \text{Realschule}} = 1{,}180$, ES $d_{4\text{–}6\ \text{Gymnasium}} = 1{,}1332$). „80 % der Varianz zwischen den Klassen ist [...] durch die Schulform erklärbar“ (Van Ophuysen & Wendt, 2009, S. 320). Die LAU-Ergebnisse zeigen zudem, dass Schülerinnen und Schüler mit einem gleichen Ausgangsniveau am Anfang der Klasse fünf auf dem Gymnasium einen höheren Kompetenzzuwachs erfahren als – im Lernniveau vergleichbare – Schülerinnen und Schüler in den anderen Schulformen. „Dies könnte ein Hinweis darauf sein, dass gerade auch bei schwächeren Schülern unter günstigen Bedingungen ein höheres Leistungsniveau erreicht werden kann, als dies an den Haupt- und Realschulen sowie an Gesamtschulen gegenwärtig der Fall ist“ (Lehmann, Husfeldt, et al., 2001, S. 43). Die TIMS-Übergangsstudie weist zudem darauf hin, dass es massive Überschneidungen der Leistungen zwischen den Schulformen gibt. Es sind darin zum Beispiel leistungsstarke Hauptschülerinnen und -schüler zu identifizieren, die sich mit ihren Leistungen im durchschnittlichen Gymnasialniveau bewegen (Jonkmann et al., 2010, S. 142).

In der Betrachtung der „Risikoschülerinnen und -schüler“, die für den Beruf bzw. das alltägliche Leben relevante Aufgaben auf dem Niveau der jeweiligen Jahrgangsstufe nur mit 50-prozentiger Wahrscheinlichkeit lösen können (Pekrun et al., 2006), stammen in der PALMA-Untersuchung 85 % aus der Hauptschule. Dabei steigt die Anzahl der Risikoschülerinnen und

-schüler in dieser Schulform im Laufe der Schuljahre massiv (5. Schuljahr: 34, 8%; 6. Schuljahr: 42,5 %, 7. Schuljahr: 50,5 %). Festgestellt wird zudem in der KESS-Untersuchung, dass Hauptschülerinnen und -schüler in Klasse acht mit ihren Kompetenzen ca. 1 ½ Jahre hinter den Realschülerinnen und -schülern liegen und zudem der Abstand zwischen den Realschülerinnen und -schülern und den Gymnasiasten sogar zwei bis drei Schuljahre beträgt (Bos & Gröhlich, 2010, S. 38). Auch die LAU-Untersuchung zeigt dahingehend gravierende Ergebnisse. Auch wenn sich alle Schülerinnen und Schüler dort zu Beginn der siebten Klasse auf alle fünf Niveastufen verteilen (s. o.), befindet sich mehr als zwei Drittel (64,7 %) der Haupt- und Realschülerinnen und -schüler auf der Niveaustufe I und II (von V) (Lehmann, Husfeldt, & Peek, 2001, S. 36 ff.). Ein ähnliches Bild zeigt sich in der LAU-Untersuchung bei den Gesamtschülerinnen und -schülern. Dort beträgt der Anteil derjenigen auf Stufe I und II 57,8 %. Die 29,9 % der Haupt- und Realschülerinnen und -schüler sowie die 25,0 % Gesamtschülerinnen und -schüler, die sich *ausschließlich* auf der Niveaustufe I bewegen, liegen mit ihrem Leistungsvermögen (maximal 72 Punkte) signifikant unter dem Gesamtdurchschnitt (M = 100; SD = 25).

Dass sich aber offenbar gerade die schwachen Schülerinnen und Schüler (zumindest in absoluten Werten) stärker entwickeln, zeigt ebenso die LAU-Untersuchung. Diejenigen Schülerinnen und Schüler, die am Anfang der Klasse fünf nur wenige Punkte im Hamburger Schulleistungstest 4/5 erreichen, verzeichnen bis zum Ende der Klasse sechs den größten Lernzuwachs. Bei den leistungsstarken Schülerinnen und Schülern ist ein deutlich geringerer Lernzuwachs zu verzeichnen. Es gelingt denjenigen Haupt- und Realschülerinnen und -schülern, die am Anfang der Klasse fünf unter der schulformspezifischen durchschnittlichen Lernausgangslage liegen (M = 81,2 Punkte), am Ende der sechsten Jahrgangsstufe durchschnittlich 88,2 Punkte zu erreichen, sodass sie dann an der Schwelle zur Kompetenzstufe II stehen. Eine vergleichbare Entwicklung durchlaufen die leistungsschwachen Gesamtschülerinnen und -schüler (M = 83,3 Punkte), die am Ende der sechsten Jahrgangsstufe durchschnittlich 91,6 Punkte erreichen und sich damit schon innerhalb der Kompetenzstufe II befinden (Lehmann, Husfeldt, et al., 2001, S. 42).

In der Frage, inwiefern mathematische Leistungen in der Grundschulzeit spätere mathematische Kompetenzen *vorhersagen* können, sind die Münchener „Longitudinalstudie zur Genese individueller Kompetenzen“ (kurz: LOGIK-Studie) und die „Schulorganisierte Lernangebote und Sozialisation von Talenten, Interessen und Kompetenzen“-Studie (kurz: SCHOLASTIK-Studie) zentral. Die SCHOLASTIK-Studie stellt eine Ergänzung der LOGIK-Studie dar: Die LOGIK-Untersuchung beschränkt sich ausschließlich auf den Zeitraum der Grundschulzeit und durch die SCHOLASTIK-Studie werden dann beim Eintritt in die Grundschule die neu dazu gekommenden Klassenkameraden und -kameradinnen in die Testungen einbezogen.

Stern (2013) kann aufzeigen, dass das Lösen von Textaufgaben in der zweiten Klasse mit dem Lösen von Textaufgaben in der elften Klasse mit einer Korrelation von $r = .58$ zusammenhängt. Es liegt somit ein deutlicher Zusammenhang vor (nach Von Hehn, 2014).

TABELLE 6 KORRELATION ZWISCHEN DEN MATHEMATIKLEISTUNG IN DEN ERSTEN SECHS SCHULJAHREN MIT DENEN IN DER ELFTEN KLASSE (STERN, 2013)

Klasse	Lösen von Textaufgaben	Rechnen
2	.58**	.22
3	.45**	.21
4	.42**	.25
5	.46**	-
6	.49**	-

Besorgniserregend ist dabei die Feststellung Sterns (2013), dass es Schülerinnen und Schüler gibt, die in der zweiten Klasse leistungsstark sind, die guten Leistungen aber dann in der elften Klasse nicht mehr zeigen, es hingegen aber keine Schülerinnen und Schüler gibt, die in der zweiten Klasse leistungsschwach und in der elften Klasse leistungsstark sind.

> „Die LOGIK-Daten zeigen sogar, dass frühe Versäumnisse im Erwerb der kulturellen Mathematik später nicht mehr kompensiert werden können. Dieses Ergebnis könnte fatalistische Einstellungen auslösen. So könnte man meinen, wenn die ‚kritische Periode' für den Zugang zur kulturellen Mathematik versäumt wurde, sei ‚der Zug abgefahren'." (Stern, 2003, S. 128)

Zudem zeigt bereits Abbildung 4 in Kapitel 2.2.3, dass Stern (2008) eine Korrelation von $r = .43$ zwischen dem numerischen Verständnis im Alter von 5 Jahren und der mathematischen Intelligenz (zum Beispiel gemessen durch die Aufgabe „Zahlenreihen fortführen") im Alter von 23 Jahren belegen kann.

Insgesamt zeigen vor allem die LOGIK- und SCHOLASTIK-Studie, dass die frühe und gesicherte Anbahnung mathematischer Kompetenzen für den weiteren Entwicklungsverlauf bis ins Erwachsenenalter hinein von großer Bedeutung ist. Fehlendes mathematisches Wissen kann nicht durch Intelligenz kompensiert werden. Dennoch betont Stern (2008, S. 201), dass „für den Übergang von der intuitiven zur akademischen Mathematik [die; d. Verf.] schulische Unterstützung unabdingbar [ist]". Damit stützt sie die These von Schneider und Kollegen (2013, S. 95 f.), die darlegen, dass „der Erwerb mathematischer Kompetenz von außerschulischen Erfahrungen beeinflusst wird, [...] dennoch außer Frage [steht], dass der wichtigste Lernkontext natürlich der Mathematikunterricht ist". Geführte Interviews mit den Schülerinnen und Schülern der PALMA-Untersuchung zeigen, „dass unzureichend ausgebildete Grundvorstellungen eine zentrale Rolle für die Schwierigkeiten von Schülerinnen und Schülern mit mathematischen Aufgaben zukommt; eine stärkere Förderung dieser Vorstellungen könnte ungünstigen Lernverläufen entgegenwirken" (Pekrun et al., 2006, S. 36).

Die dargelegten Längsschnittuntersuchungen spiegeln insgesamt ein stark einheitliches Bild bzw. die durchschnittlichen Erwerbsverläufe der Schülerinnenund Schüler an unterschiedlichen Schulformen wider. Festzuhalten ist, dass die Schülerinnen und Schüler grundsätzlich eine positive mathematische Entwicklung durchlaufen. Ein Blick ins Detail zeigt jedoch, dass es einen nicht geringen Anteil an Schülerinnen und Schülern gibt, der auch in der Sekundarstufe I massive Defizite zeigt. Diese befinden sich oftmals an Haupt- oder Gesamtschulen, sodass diese beiden Schulformen in Vergleichsstudien eher das Schlusslicht bilden. Die dargestellten Studien sind dabei bildungswissenschaftlich angelegt, die allgemeine Entwicklungstrends zulassen, qualitative Ergebnisse jedoch eher offen bleiben.

Zu berücksichtigen ist hier die Überschneidungsmenge der in den Studien herangezogenen Mathematikaufgaben bzw. Items. So werden zum Beispiel im Projekt Dortmunder Grundschulübergang Aufgaben aus der LAU- oder PALMA-Untersuchung gewählt. Die KESS- und die LAU-Untersuchung sowie die TIMS-Übergangsstudie ziehen Aufgaben aus der TIMS-Studie hinzu oder lehnen ihre daran an. Die Zirkularität der verwendeten Items mag ein Hinweis auf die insgesamt eher vergleichbaren Forschungsergebnisse sein. Dies sollte in Fragen der Aussagekraft der Studien berücksichtigt werden.

2.4.2.2 Internationale Studien

Es ist bereits problematisiert worden, dass der mathematische Kompetenzerwerb von Schülerinnen und Schülern aufgrund divergierender Schulsysteme und Curricula bzw. Syllabi international nur schwer zu vergleichen ist. Zudem ist auch international ein Desiderat an Längsschnittstudien festzustellen, die sich mit der Kompetenzentwicklung nach dem Übergang in die weiterführende Schule beschäftigen. Holden-Pitt (2007, S. 27) stellt in diesem Zusammenhang jedoch fest:

> „Although mathematics still receives less attention than reading, interest in it has been growing in recent years. [...] The increased emphasis on mathematics also results from recognition of its role in an increasingly technical and international economy."

Ähnlich dem Vorgehen bei den nationalen Längsschnittstudien gibt die Abbildung 13 einen ersten grafischen Überblick über ausgewählte internationale Längsschnittstudien, die den Erwerb mathematischer Kompetenzen im Kontext der Transition in die Sekundarstufe I betrachten.

Die einbezogenen Studien erfüllen folgende Kriterien:

- Das Design entspricht einem längsschnittlichen Design (mindestens zwei aufeinander folgende Datenpunkte).
- Es wird der Übergang von der Primar- in die Sekundarstufe I berücksichtigt.
- Es werden – auch – mathematische Kompetenzen erhoben.
- Die Untersuchung stammt nicht aus Deutschland.

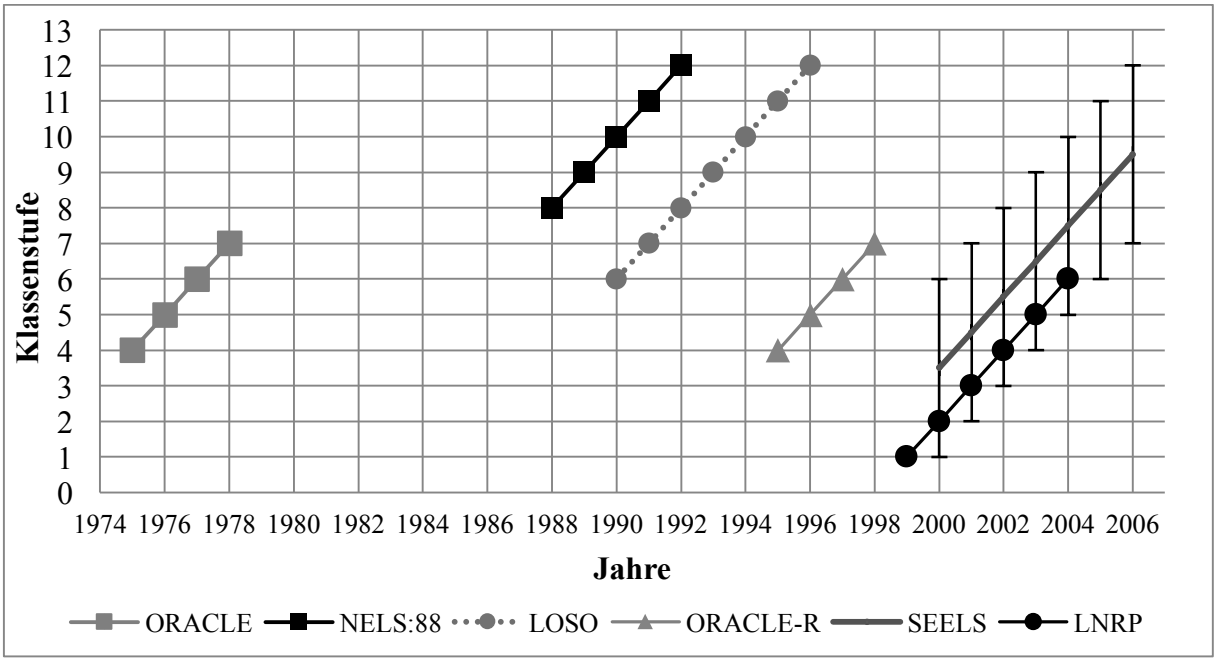


ABBILDUNG 13 ÜBERBLICK ÜBER INTERNATIONALE LÄNGSSCHNITTSTUDIEN ZUM ERWERB MATHEMATISCHER KOMPETENZEN IN DER SEKUNDARSTUFE I

Trotz schwieriger Forschungslage wird versucht, einen weiten internationalen Überblick zu geben und dabei verschiedene Länder zu berücksichtigen. Betont werden muss an dieser Stelle, dass der Übergang in die Sekundarstufe international im Regelfall nicht nach der vierten Klasse, sondern später, zum Beispiel nach der Klasse sechs erfolgt, sodass die Studien im Vergleich zu Kapitel 2.4.2.1 höhere Jahrgänge zum Inhalt haben.

Nicht einbezogen ist in diesem Kapitel die National Longitudinal Transition Study 2 (NLTS2) des US-amerikanischen National Center für Special Education Research, die von 2000 bis 2010 die Entwicklung von Schülerinnen und Schüler mit Unterstützungsbedarfen im Übergang in den *beruflichen* Kontext begleitet (siehe dazu Lenz & Deshler, 2008, S. 447).

Einen detaillierten Überblick über die Studien ermöglicht hier die folgende Tabelle 7. Es gilt ebenso: Die Angabe in der Spalte N bezieht sich auf die Stichprobe zum jeweils ersten Messzeitpunkt und berücksichtigt keinen Drop-Out für den Verlauf der Erhebungen. Zudem ist das Land angegeben, aus dem die Studie stammt. Auch hier sind die Studien chronologisch entsprechend des letzten Erhebungsjahres aufgeführt, sodass oben die jüngste Studie zu finden ist.

TABELLE 7 DETAILANGABEN ZU INTERNATIONALEN LÄNGSSCHNITTSTUDIEN ZUM ERWERB MATHEMATISCHER KOMPETENZEN IN DER SEKUNDARSTUFE I

Studie	Literaturverweis	Lauf-zeit	Klassen-stufen	N	Variablen	Zentrale Ergebnisse bzgl. math. Erwerb
USA: Special Education Elementary Longitudinal Study (SEELS)	Holden-Pitt (2007)	2000 – 2006	1–12	~ 11.000	• School engagement • Academic performance • Social adjustment • independence	• Schüler mit einem Unterstützungsbedarf entwickeln sich schwach aber grundsätzlich positiv • Schüler mit einer starken geistigen Beeinträchtigung oder einer komplexen Beeinträchtigung zeigen die schwächste Entwicklung • Einige Schüler zeigen Leistungsrückgänge im Laufe der Zeit
UK: Levelhume Numeracy Research Programme (LNRP)	Brown u. a. (2003)	1998 – 2004	Vor-schule – 6	~ 1.700	• understanding the number system • methods of computation • solving numerical problems	• Arithemtische Kompetenzen entwickeln und festigen sich ab Klasse zwei bis Klasse sechs immer stärker • Arithmetische Kompetenzen steigen auch in Klasse fünf und sechs weiter an
USA: National Education Longitudinal Study of 1988 (NELS:88)	Kaufman & Bradbury (1992)	1988 – 2000	8 – Berufs-tätigkeit	~ 25.000	• Math • Science • Reading • History • Student background • School characteristics • Lehrer- und Eltern-fragebögen	• 19 % der Schüler befinden sich zu Beginn der Klasse acht auf Stufe I (von IV) • Risikofaktoren bei unterdurschnittlichen Mathematikleistungen: alleinerziehende Eltern, häufiger Schulwechsel, keine Erledigung von Hausarbeit, schlechte Einschätzug der Lehrkraft, Schulabsentismus, Schulen mit vielen Minderheiten-Schülern

Studie	Literaturverweis	Laufzeit	Klassenstufen	N	Variablen	Zentrale Ergebnisse bzgl. math. Erwerb
Belgien: Longitudinal Research Secondary Education (LOSO)	Opdenakker & Van Damme (2006) Opdenakker & Van Damme (2007) Van Damme et al. (2002)	1990 – 2000	6 – Berufstätigkeit	6.411	• effort concerning educational tasks • relationship with the teacher • academic self-concept • general school well being	• Leistungsstärke der Lerngruppe hat einen signifikanten Einfluss auf die Leistungsentwicklung des einzelnen Schülers • Ebenso einen Einfluss nehmen das kognitive Niveau der Schülerschaft, die Schulgröße, die Kooperation unter Lehrkräften und das Schulklima
UK: Observational and Classroom-Learning Evaluation-Study-Replication ORACLE-R	Galton, Comber & Pell (2002)	1995 – 1997	4–7	476	• Mathematics • English • Reading • Teaching style	• Schüler entwickeln sich in Rechenoperationen und Problemlöseaufgaben grundsätzlich positiv • Leistungsrückgänge jeweils zu Beginn eines neuen Schuljahres möglich • 10,6 % weniger richtig gelöste Aufgaben als in der ORACLE-Studie (1975–1980)

Anmerkung: Schüler bedeutet zugleich Schüler und Schülerinnen

Auch international weisen die hier vorgestellten Längsschnittstudien darauf hin, dass Schülerinnen und Schüler in der Sekundarstufe I grundsätzlich eine positiv verlaufende mathematische Entwicklung aufweisen. Dies zeigt zum Beispiel die ORACLE-Replikationsstudie. Brown und Kollegen (2003, S. 10) beschreiben darin eine ansteigende und stärker werdende mathematische Entwicklung der Schülerinnen und Schüler von Klasse zwei bis sechs: „Generally they rose slowly from zero to 10 % or 20 %, then more quickly, at a rate of 10 % to 20 % per year, then gradually more slowly again, plateauing out towards 100 %“ (siehe auch Abbildung 14). Die SEELS-Untersuchung zeigt, dass dieser positive Trend auch für Schülerinnen und Schülern mit einem Unterstützungsbedarf bestätigt werden kann. Dort erzielen die Schülerinnen und Schüler (overall) zum ersten Messzeitpunkt (2001) im Durchschnitt 497 Punkte (range 477–501) und im weiteren Messzeitpunkt (2004) 514 Punkte (range 490–514) (Holden-Pitt, 2007, S. 28). Die Ergebnisse werden von den Autoren nicht bezüglich der konkreten Altersspannen veröffentlicht, sodass in den Ergebnissen Schülerinnen und Schüler aus der Klasse eins bis sechs gemeinsam berücksichtigt sind.

Ergebnisse über eine allgemeine Entwicklungsbeschreibung hinaus stammen aus der ORACLE-R- und der LNRP-Studie. Die Autoren der ORACLE-R-Studie veröffentlichen Detailergebnisse bezogen auf die Bereiche Rechenoperationen und Problemlösen (siehe Abbildung 14) und die Autoren der LRNP-Studie gar bezogen auf die konkret gestellten Rechenaufgaben (siehe Abbildung 15).

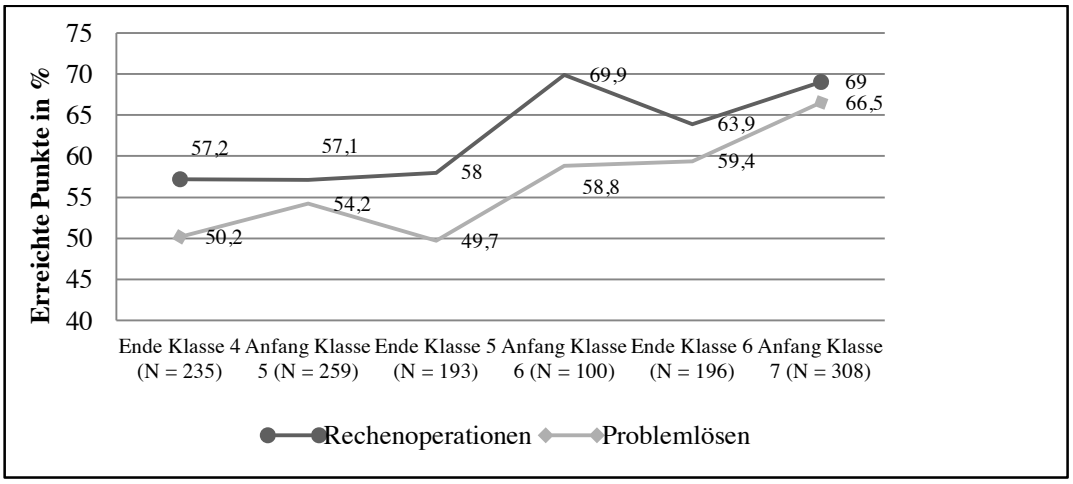


ABBILDUNG 14 ERWERB MATHEMATISCHER KOMPETENZEN IN DER ORACLE-REPLIKATIONSSTUDIE (GALTON ET AL., 2002, S. 147)

In Abbildung 14 wird der grundsätzlich positive Trend durch den tendenziell steigenden Punktezuwachs im Bereich Rechenoperationen und Problemlösen deutlich. Doch ein Blick ins Detail zeigt hier, dass ausschließlich *innerhalb* eines Schuljahres Leistungsrückgänge stattfinden (zum Beispiel Problemlösen Anfang Klasse fünf vs. Ende Klasse fünf). Im Übergang von der vierten zur fünften Klasse erzielen im Bereich Mathematik 45,8 % der Schülerinnen und Schüler einen Leistungsrückgang und nur 54,2 % einen Leistungszuwachs. Im Übergang von der fünften zur sechsten Klasse betragen die Quoten 15,7 % (Rückgang), 5,9 % (keine Veränderung) bzw. 78,4 % (Zuwachs). Von der sechsten bis zur siebten Klasse liegen die Werte bei 23,7 % (Rückgang), 10,9 % (keine Veränderung) bzw. 65,4 % (Zuwachs).

Zudem scheint die Kompetenzentwicklung im Bereich der Rechenoperationen stärker zu verlaufen, als im Bereich Problemlösen. Dies wird auch durch die Befunde des oben bereits vorgestellten Early Numeracy Research Projekts (Australien) bestätigt (siehe Kapitel 2.3.1.3).

Im Rahmen der LNRP-Studie werden Ergebnisse auf Itemebene dargestellt. Exemplarisch werden hier die Ergebnisse zur Kompetenzentwicklung im Bereich der Addition vorgestellt.

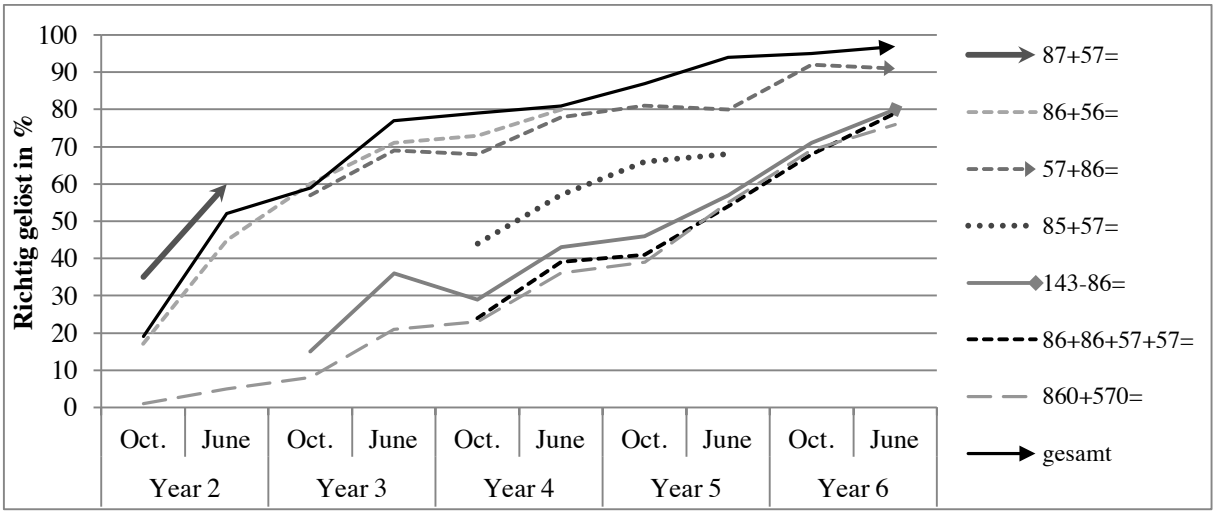


ABBILDUNG 15 ERGEBNISSE DES LEVERHULME NUMERACY RESEARCH-PROGRAMMS IM BEREICH „METHODS OF COMPUTATION" (ADDITION) (BROWN ET AL., 2003)

Auch hier ist in den Klassen fünf und sechs durchschnittlich ein Kompetenzzuwachs festzustellen. Die Addition und Subtraktion von z. T. mehrstelligen Zahlen (860+570=; 143–86=) fällt den Schülerinnen und Schülern in der Klasse zwei und drei erwartungsgemäß noch schwerer und wird im Laufe der Schulzeit stetig leichter. Am Ende der Klasse sechs zeichnet sich ein Deckeneffekt ab, was ein Hinweis dafür ist, dass die arithmetischen Aufgaben verstärkt von den meisten Schülerinnen und Schülern richtig gelöst werden können.

Einen Blick auf die Risikoschülerinnen und -schüler, die nicht in der Lage sind, die grundlegenden Rechenoperationen mit ganzen Zahlen auszuführen (Kaufman & Bradbury, 1992), richtet die NELS:88-Studie, die zu den bedeutsamsten Längsschnittstudien der USA im Sekundarstufenbereich zählt (Kristen, Römmer, Müller & Kalter, 2005, S. 28). In der Untersuchung befinden sich zu Beginn der Klasse acht nahezu 19 % der Schülerinnen und Schüler auf dem Niveau „below basic" (I von IV). Detailanalysen zeigen u. a. folgende Risikofaktoren auf, die diese Schülerinnen und Schüler auf dem below-basic-Niveau im Bereich Mathematik kennzeichnen (Kaufman & Bradbury, 1992, S. 51): alleinerziehende Eltern, häufiger Schulwechsel, Klassenwiederholung, schlechte Noten in Mathematik, keine Vorbereitung auf den Unterricht zum Beispiel durch Hausarbeiten, Schulabsentismus, schlechte Einschätzung durch die Lehrkraft oder Unterrichtung an Schulen mit einer großen Anzahl an Minderheiten. Auch Schülerinnen und Schüler mit einem Unterstützungsbedarf haben oftmals Schwierigkeiten im Erwerb mathematischer Kompetenzen. Die SEELS-Untersuchung nimmt diese Schülerinnen und Schüler in den Blick (Blackorby, Chorost, Garza & Guzman, 2004). Dabei erzielen die Schülerinnen und Schüler mit einer starken geistigen Beeinträchtigung (cognitive) und einer komplexen Beeinträchtigung (severe) in allen Erhebungen die geringsten Punktewerte und weisen dabei zugleich die höchste Streuung der Werte auf. Trotz eines Leistungszuwachses der Schülerinnen und Schüler innerhalb von sechs Jahren liegen die Leistungen immer noch deutlich unter denen

der Schülerinnen und Schüler ohne einen Unterstützungsbedarf. Insgesamt zeigt die SEELS-Untersuchung, dass die meisten Schülerinnen und Schüler mit Unterstützungsbedarfen im Bereich Mathematik niedrige Prozentränge einnehmen. 63 % der Gesamtgruppe erreicht in der *ersten* Erhebung in den Problemlöseaufgaben einen Prozentrang von 0–25, im Bereich Rechenoperationen waren es 40 %.

Die Forschergruppe der SEELS-Untersuchung berichtet zudem davon, dass in jeder Gruppe der Schülerinnen und Schüler mit einem Unterstützungsbedarf jene aufzufinden sind, die sich in ihrer Entwicklung verschlechtern (Holden-Pitt, 2007, S. 29). Dass zwischenzeitige Leistungsrückgänge erfolgen können, zeigt auch die ORACLE-R-Untersuchung, wie oben bereits abgebildet (siehe Abbildung 14).

Die vorliegenden Studien geben zudem Hinweise darauf, welche Faktoren den mathematischen Kompetenzerwerb beeinflussen können. Opdenakker und Van Damme (2007) finden in der LOSO-Untersuchung heraus, dass die Leistungsstärke der Lerngruppe einen signifikanten Einfluss auf die mathematische Entwicklung jedes einzelnen Schülers bzw. jeder Schülerin hat (sog. Kompositionseffekt). Ist die Lerngruppe insgesamt leistungsstark, erzielen auch die einzelnen Schülerinnen und Schüler bessere Leistungen, als befänden sie sich in einer leistungsschwachen Klasse. Ebenso Einfluss auf den mathematischen Kompetenzerwerb nimmt die durchschnittliche Intelligenz der gesamten Schülerschaft („school composition“) und die Schulgröße („school context“). Der Einfluss dieser Variablen wird dabei durch die „school practice“ vermittelt, worunter zum Beispiel die Kooperation unter den Lehrkräften oder das Schulklima fällt (Opdenakker & Van Damme, 2007, S. 17). 19 % der Unterschiede in der Varianz der mathematischen Leistungen konnten durch Aspekte auf Schulebene aufgeklärt werden (zum Beispiel group composition, social and learning climate, opportunities to learn) (Opdenakker & Van Damme, 2006).

Die internationalen Längsschnittstudien sind – im Vergleich zu den nationalen – schwer zu vergleichen bzw. zu synthetisieren. Gründe dafür sind zum Beispiel die unterschiedlichen Forschungsziele, die sich hinter den Studien befinden oder auch die unterschiedlichen Designs und Operationalisierungen verbunden mit unterschiedlichen Schulsystemen. Grundsätzlich erkenntlich ist auch hier, dass die mathematische Kompetenzentwicklung im Regelfall positiv verläuft. Jedoch wird auch ersichtlich, dass Schülerinnen und Schüler an verschiedenen Stellen (Schuljahresbeginn, innerhalb des Schuljahres) Rückgänge in ihren mathematischen Leistungen zeigen. In den Fokus geraten hier vor allem Schülerinnen und Schüler mit einem sonderpädagogischen Unterstützungsbedarf und diejenigen, die einer sozioökonomischen Benachteiligung unterliegen. Die Darstellung eines übergeordneten internationalen Vergleichs mathematischen Kompetenzerwerbs stellt bisher ein Forschungsdesiderat dar, dem durch eine Metaanalyse begegnet werden kann.

2.4.3 Erwerb mathematischer Kompetenzen in der Sekundarstufe I im Querschnitt

Neben den bisher aufgezeigten Längsschnittstudien zum Kompetenzerwerb im Bereich Mathematik sind ebenso Querschnittstudien zu verzeichnen, die sich mit Kompetenzständen in der Sekundarstufe I in verschiedenen Jahrgangsstufen beschäftigen. Folgende Tabelle 8 gibt in chronologischer Reihenfolge einen Überblick über die im Folgenden vorgestellten Studien. Somit ist auch hier die jüngste Studie oben zu finden.

Die vorliegenden Studien erfüllen folgende Kriterien:

- Das Design entspricht einem querschnittlichen Design.
- Die Schülerinnen und Schüler befinden sich am Ende der Primarstufe oder in der Sekundarstufe I.
- Es werden – auch – mathematische Kompetenzen erhoben.
- Es werden auch Schülerinnen und Schüler an sog. „nicht gymnasialen Schularten“ (Hammer et al., 2016, S. 237) berücksichtigt.
- Die Untersuchung stammt aus Deutschland oder Deutschland ist an ihr beteiligt.

TABELLE 8 DETAILANGABEN ZU NATIONALEN QUERSCHNITTSTUDIEN ZUM ERWERB MATHEMATISCHER KOMPETENZEN IN DER SEKUNDARSTUFE I

Autoren/Studie	Jahr	Klasse	N	Variablen	Zentrale Ergebnisse
Gebhardt, Oelkrug & Tretter	2013	5–9	124	• Mathematische Kompetenzen (Mengen-Zahlen-Kompetenz, Konventions- und Regelwissen)	• Betrachtet Schüler mit sonderpäd. Untertstützungsbedarf • Grundsätzlich positive Entwicklung • Große Streuungen in den Leistungen • Abschlussschüler erfüllen nicht die Anforderungen
Ehlert, Fritz, Arndt & Leutner	2013	5, 6, 7	3.807	• Teil-Teil-Ganzes-Konzept	• Leistungen in allen Schulformen entwickeln sich grundsätzlich positiv • Gesamtschule zeigt keine sign. Entwicklung • Haupt- und Gesamtschüler zeigen große Defizite im Teil-Teil-Ganzes-Konzept
Programme for International Student Assessment (PISA)	2012	15jähr.	5.001	• Veränderungen und Beziehungen • Raum und Form • Quantität • Unsicherheit und Daten	• Deutschland liegt signifikant über dem OECD-Durchschnitt • 17,7 bzw. 17,2 % der Schüler auf oder unter Stufe I (von VI) • keine Verbesserung zwischen 2012 und 2015
	2015	15jähr.	6.504		
Trends in International Mathematics and Science Study (TIMSS)	2011	4	3.995	• Arithmetik • Geometrie/Messen • Umgang mit Daten	• 2011 & 2015: Deutschland liegt signifikant über dem OECD-Durchschnitt • Zwischen 2007 und 2015 keine sign. Veränderungen • 2011: 19,2 % der Schüler auf Stufe I oder II (von V), 2015: 23,6 %
	2015		3.948		
Ennemoser, Krajewski & Schmidt	2011	5–9	1.782	• Mengen-Zahlen-Kompetenz • Regel- und Konventions-wissen	• Grundsätzlich positive Entwicklung in der Sek. I (Quasi-Querschnitt) • Hauptschüler zeigen schlechtere Leistung als Realschüler, diese schlechtere als Gymnasiasten • Festigung der Inhalte dauert bis zum Ende der Sek. I

Autoren/Studie	Jahr	Klasse	N	Variablen	Zentrale Ergebnisse
Humbach	2008	10	458	• Arithmetische Basiskompetenzen	• Haupt- und Gesamtschüler zeigen schwache Leistung in arith. Kompetenzen • Hohe Korrelation zwischen arith. Kompetenzen und weiterführenden Aufgaben
Moser-Opitz	2005	5 & 8	4.025	• Arithmetische Kompetenzen • Regel- und Konventions-wissen	• Kompetenzdefizite in basalen arithmetischen Aufgaben sowohl bei Schülern mit als auch ohne eine Rechenschwäche
Mittelberg	2004	7 & 8	204	• Arithmetische Kompetenzen	• Schwierigkeiten liegen nicht nur an Hauptschulen, sondern ebenfalls an Gymnasien vor
Berliner Erhebung der Lernausgangslagen arbeitsrelevanter Basiskompetenzen von Schülerinnen und Schülern mit Förderbedarf (BELLA)	2006 – 2007	7–10	4.481	• Mathematikleistung • Unterrichts- und lernbezogene Merkmale • schulische und außerschulische Kontextbedingungen	• Insgesamt schwache math. Leistungen der Schüler mit einem sonderpäd. Unterstützungsbedarf • Breite Streuung im Leistungsspektrum • Bimodale Verteilung der Werte • Inklusive Beschulung hängt nur bedingt mit besseren math. Kompetenzen zusammen
Mathematik-Gesamterhebung Rheinland-Pfalz: Kompetenzen, Unterrichtsmerkmale, Schulkontext (MARKUS)	1999 – 2002	8	37.520	• Mathematikleistung • Unterrichts- und lernbezogene Merkmale • schulische und außerschulische Kontextbedingungen	• Signifikante Unterschiede zwischen Schulformen • Bedeutsam sind Classroom-Management der Lehrkraft, die Mitarbeit im Unterricht (Lernhaltung der Klasse) und die Bildungsnähe der Schüler
Qualitätsuntersuchung an Schulen zum Unterricht in Mathematik (QuaSUM)	1999	5 & 9	12.500	• Leistungen im Fach Mathematik • Inner- und außerschulische Lernbedingungen	• Leistungen können aufgeklärt werden durch kognitive Voraussetzungen, Sach- und Fachinteresse für Mathematik, Selbstvertrauen, Schulform und mathematisches Fachinteresse

Anmerkung: Schüler bedeutet zugleich Schüler und Schülerinnen.

Die beiden größten und umfangreichsten Querschnittstudien sind hier die PISA- und die TIMS-Studie. Sie ermöglichen einen allgemeinen Überblick über die mathematischen Kompetenzen von Schülerinnen und Schülern und vergleichen diese international. Die Trends in International Mathematics and Science Study (kurz: TIMSS) betrachtet dabei die Kompetenzen *am Ende* der Primarstufe, mit denen die Schülerinnen und Schüler dann in die Sekundarstufe I übergehen. Im Jahr 2011 erreichen die Schülerinnen und Schüler in ihren mathematischen Kompetenzen 528 Punkte, im Jahr 2015 522 Punkte (Selter, Walter, Walther & Wendt, 2016, S. 106) und liegen damit in beiden Jahren signifikant über dem OECD-Durchschnitt (Selter et al., 2012). 2011 befinden sich jedoch 2,7 % der deutschen Viertklässler auf der Kompetenzstufe I (von V), 2015 sind es sogar 4 %. Diese Schülerinnen und Schüler verfügen damit nur über rudimentäres mathematisches Anfangswissen. Weitere 16,5 % der Stichprobe befinden sich 2011 auf Stufe II (2015: 19,6 %), sodass diese Schülerinnen und Schüler lediglich elementares mathematisches Wissen haben und damit zum Beispiel nur leichte Additions- und Subtraktionsaufgaben lösen können (Selter et al., 2012, S. 104). Insgesamt zeigen die deutschen Viertklässler die größten Defizite im Bereich der Arithmetik (Selter et al., 2012, S. 106).

Die mathematischen Kompetenzen in der Sekundarstufe I untersucht das Programme for International Student Assessment (PISA). In der Studie werden international u. a. die mathematischen Kompetenzen 15-jähriger Schülerinnen und Schüler erhoben. Für Deutschland zeigt sich in der Untersuchung im Jahr 2012, dass sich im mathematischen Teil 17,7 % der Stichprobe unter oder auf der Kompetenzstufe I (von VI) befindet. Im Jahr 2015 sind es noch 17,2 % der Schülerinnen und Schüler (Hammer et al., 2016, S. 233). Dies bedeutet, dass „jeder sechste Jugendliche in Deutschland die Mindestanforderungen für ein anschlussfähiges mathematisches Verständnis nicht erreicht und erhebliche Probleme haben dürfte, einen Ausbildungsplatz zu finden beziehungsweise eine Ausbildung erfolgreich abzuschließen sowie anspruchsvollere mathematische Anforderungen im Alltag zu bewältigen“ (Prenzel et al., 2012, S. 75). Die PISA-Ergebnisse aus dem Jahr 2015 zeigen zudem, dass keine signifikanten Verbesserungen in den mathematischen Kompetenzen im Vergleich zum Jahr 2012 vorliegen.

Doch wie sehen die Fähigkeiten und Schwächen der Schülerinnen und Schüler konkret aus? Hinweise dazu gibt die Untersuchung von Moser Opitz (2005) über die mathematischen Kompetenzen in den Klassenstufen fünf und acht, die die alarmierenden Ergebnisse der TIMS- und PISA-Studie bestätigt. In ihrer Untersuchung lösen die Aufgabe 12 : 4 = lediglich 19,1 % der identifizierten rechenschwachen Fünftklässler, jedoch auch nur 51,1 % der durchschnittlich mathematisch begabten Schüler. Im achten Schuljahr beträgt die Lösungsquote bei dieser Aufgabe 31,8 % (rechenschwach) bzw. 63,8 % (durchschnittlich mathematisch-begabt). Die größten Defizite zeigen sich im Bereich der Division und weitere Defizite darüber hinaus u. a. in der Strategieverwendung und dem halbschriftlichen Rechnen (Moser Opitz, 2007a). Auch Mittelberg (2004) untersucht die mathematischen Kompetenzen in der Sekundarstufe I von 7. und 8.-Klässlern an Hauptschulen und Gymnasien ($n_{Gymnasium} = 44$; $n_{Hauptschule} = 160$). Der Autor stellt den Schülerinnen und Schülern Aufgaben aus dem Bereich der Grundschulmathematik. Folgende Abbildung 16 zeigt die Fehlerquote bei Aufgaben zur Arithmetik zusammengefasst für die siebte und achte Klasse.

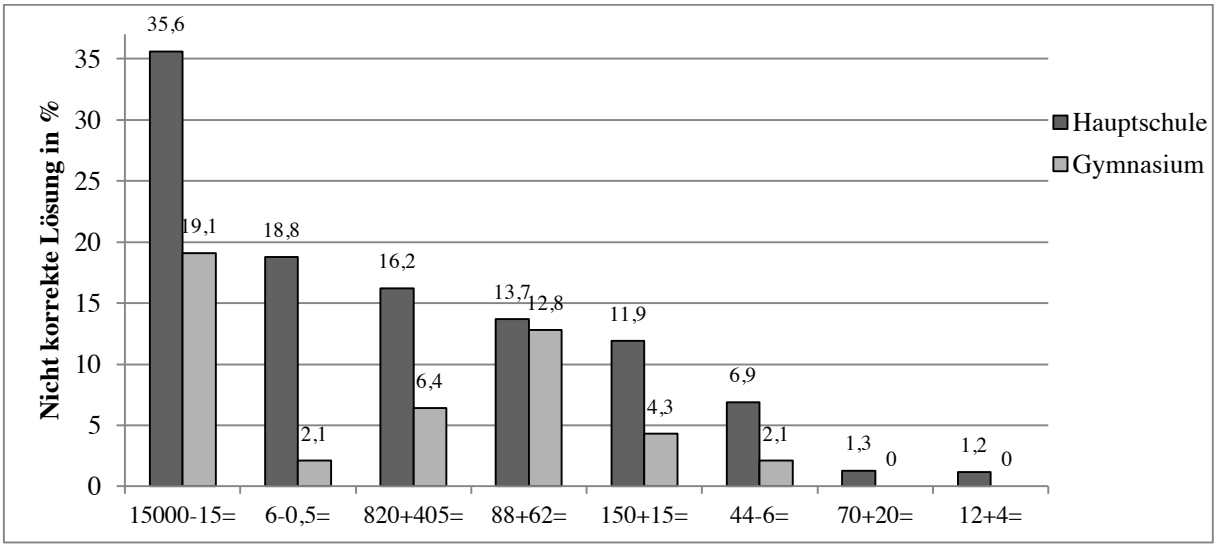


ABBILDUNG 16 ARITHMETISCHE KOMPETENZEN VON 7- UND 8-KLÄSSLERN (MITTELBERG, 2004, S. 68)

Wie die Abbildung 16 grafisch verdeutlicht, werden noch in Klasse sieben und acht basale Rechenaufgaben von Schülerinnen und Schülern sowohl in der Hauptschule als auch am Gymnasium falsch gelöst. Je größer die Zahlen, umso weniger Strategien bzw. Ideen zum Lösen der Aufgaben liegen offenbar vor, sodass weniger Aufgaben richtig gelöst werden. Diese Befunde bestätigt im Längsschnitt auch das oben bereits beschriebene Leverhulme Numeracy Research Programme (Brown et al., 2003) (siehe Kapitel 2.4.2.2.).

Neben der Betrachtung der mathematischen Kompetenzen der allgemeinen Schülerschaft liegen auch Befunde zu denjenigen mit einem sonderpädagogischen Unterstützungsbedarf im Bereich Lernen vor. Gebhardt und Kollegen (2013) zeigen, dass innerhalb der einzelnen Jahrgangsstufen fünf bis neun große Streuungen in den mathematischen Kompetenzen bestehen. Erwartungsgemäß zeigt sich jedoch, dass Schülerinnen und Schüler in den höheren Jahrgangsstufen in allen Bereichen mehr Aufgaben richtig lösen als in den unteren Jahrgangsstufen. Das Niveau der älteren Schülerinnen und Schülern ist dabei trotzdem sehr niedrig. In der neunten Klasse können zum Beispiel lediglich 47 % der Schülerinnen und Schüler die Aufgabe „__ : 6 = 8“ lösen. In der fünften Klasse liegt die Quote bei 9 %. Bei der Additionsaufgabe „16 + __ = 43“ liegt die Lösungsquote bei 32 % (Klasse fünf), 37 % (Klasse sieben) und bei 79 % (Klasse neun).

> „Eine weitere Erkenntnis aus den hier vorliegenden Ergebnissen ist, dass die untersuchten Abschlussschülerinnen und -schüler nicht über die erforderlichen mathematischen Anforderungen verfügten, welche für den Besuch der Berufsschule notwendig sind. Bedenkt man, dass in dieser Untersuchung nicht die mathematischen Anforderungen der jeweiligen Klassenstufe überprüft wurden, sondern Basiskompetenzen, welche am Ende der Grundschulstufe beherrscht werden müssten, so lässt sich sogar konstatieren: Etliche Schülerinnen und Schüler mit dem SPF Lernen verlassen das Förderzentrum, ohne dass sie die Kulturtechnik Mathematik auf dem Stand der Grundschule in ausreichendem Maße erlernt hätten.“ (Gebhardt et al., 2013, S. 140)

Eine weitere Untersuchung zu Schülerinnen und Schülern mit einem sonderpädagogischen Unterstützungsbedarf liegt mit der BELLA-Studie vor. Zum einen zeigen auch hier die Ergebnisse, dass das Leistungsspektrum der Schülerinnen und Schüler eine breite Streuung aufweist. Die

Verteilung der Werte ist dabei eher bimodal, sodass die größte Leistungsgruppe nicht im mittleren Anforderungsbereich liegt (wie bei einer Normalverteilung), sondern eine große Leistungsgruppe über und unter dem durchschnittlichen Anforderungsbereich aufzufinden ist. Dies weist auf eine hohe Leistungsheterogenität hin. Wie die TIMS-Studie werden die Aufgaben in fünf Anforderungsbereiche eingeteilt. Auf Stufe I befinden sich 5,6 %, auf Stufe II 32,4 %, auf Stufe III 36,9 %, auf Stufe IV 19,1 % und auf Stufe V 6,0 % der Schülerinnen und Schüler. Zwar sind es also immerhin 6,0 % der Schülerinnen und Schüler, die Aufgaben aus dem höchsten Anforderungsniveau (Daten) lösen können. Dennoch sind es 38,0 %, die zum Beispiel keine einfachen Divisionsaufgaben auf Niveau III lösen können („Drei Kinder wollen 72 Sticker gleichmäßig aufteilen. Wie viele bekommt jedes Kind?“).

Die Ergebnisse nach Beschulungsform zeigen, dass eine inklusive Beschulung in Klasse sieben zu besseren Leistungsergebnissen führt ($d = -.11$). Dieser Trend ändert sich jedoch unter der Kontrolle der kognitiven Voraussetzungen ab Klasse neun (siehe Abbildung 17).

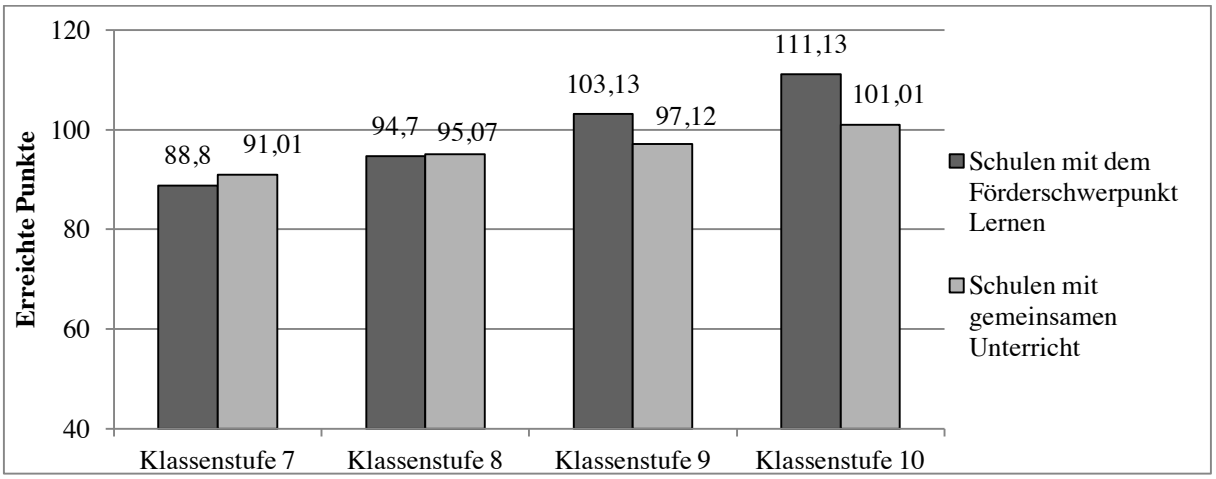


ABBILDUNG 17 VERGLEICH DER FACHLEISTUNGEN IM BEREICH MATHEMATIK NACH BESCHULUNG (HOFFMANN, 2009, S. 36)

Gravierend sind die Ergebnisse der BELLA-Studie dahingehend, dass sich der eingesetzte Mathematiktest am Lernstandsniveau von Viert- und Fünftklässlern der allgemeinen Schule orientiert. Vergleicht man folglich Berliner Schülerinnen und Schüler mit einem Unterstützungsbedarf im Bereich Lernen mit Viert- und Fünftklässlern der allgemeinen Schule, zeigt sich, dass sich diese zu einem nicht geringen Anteil unter dem Niveau der Jüngeren befinden.

Wie auch die Analyse der Längsschnittuntersuchungen bereits zeigt, geraten bei der Betrachtung mathematischer Kompetenzen neben Schülerinnen und Schüler mit einem sonderpädagogischen Unterstützungsbedarf vor allem die Haupt- und Gesamtschulen in besonderen Blick. Mittelberg (2004) kommt im Rahmen seiner Untersuchung (s. o.) zu dem Schluss, dass zu viele Hauptschülerinnen und -schüler in der Sekundarstufe I die grundlegenden Basiskompetenzen nicht beherrschen. „Durch diese Schwierigkeiten werden sie zu *rechenschwachen Schülern* der Hauptschule. Die Ergebnisse weisen darauf hin, dass *ein Großteil der untersuchten Hauptschüler* nicht in der Lage ist, den Anforderungen des aktuellen Mathematikunterrichts gerecht zu werden.“ (Mittelberg, 2004, S. 197 Herv. i. Orig.). In der PISA-Untersuchung 2012 erzielen neben den Hauptschulen auch die Integrierten Gesamtschulen und Schulen mit mehreren Bildungsgängen (die sog. „nicht gymnasialen Schularten" nach Hammer et al., 2016, S. 237) signifikant schlechtere Ergebnisse. Die mathematischen Leistungen dieser Schülerinnen und

Schüler liegen signifikant unter dem deutschen Durchschnitt. Ca. 50 % der Hauptschülerinnen und -schüler liegen unter oder auf dem ersten Kompetenzbereich, was „Anlass zu Besorgnis" (Prenzel et al., 2012, S. 87) geben muss. Auch Ehlert und Kollegen (2013) kommen in ihrer Untersuchung zur Teil-Teil-Ganzes-Kompetenz im Bereich der Grundrechenarten zu ähnlich gravierenden Erkenntnissen. In der Untersuchung zeigt sich, dass die Leistungen der Schülerinnen und Schüler auf dem Gymnasium, in der Real- und Hauptschule im Laufe der Schuljahre ansteigen (um ca. 15 Punkte). Die Leistungsentwicklung ist ausschließlich bei der Gesamtschulgruppe nicht signifikant; die Hauptschülerinnen und -schüler verzeichnen im Schulformvergleich den größten Leistungszuwachs. Dennoch können 25 % der Hauptschülerinnen und -schüler und 50 % der Gesamtschülerinnen und -schüler in der siebten Klasse lediglich Aufgaben auf Basis des kleinen Einmaleins lösen. In den Haupt- und Gesamtschulen in der fünften Klasse können ca. 30 % der Schülerinnen und Schüler Teil-Teil-Ganzes-Aufgaben zur Addition (a + b = c) und Subtraktion (a – b = c) im kleinen Zahlenraum sowie Aufgaben zur Multiplikation nur im sehr leichten Schwierigkeitsbereich lösen. 50 % der Schülerinnen und Schüler der Haupt- und Gesamtschulen und knapp 75 % derjenigen aus den Realschulen in der fünften Klasse können Teil-Teil-Ganzes-Aufgaben zur Addition und Subtraktion im Zahlenraum bis 1000 nicht lösen. „Für diese Schülergruppe ist anzunehmen, dass sie mit hoher Wahrscheinlichkeit große Probleme in ihrem weiteren Ausbildungs- und Berufsleben haben werden [sic!]" (Ehlert et al., 2013, S. 259). Ehlert und Kollegen kommen zu dem Schluss, dass in der Sekundarstufe I massive Defizite in den mathematischen Basiskompetenzen bestehen und diesen Schwierigkeiten entgegengewirkt werden muss. Bestätigt wird dies durch die Untersuchung von Ennemoser, Krajewski und Schmidt (2011). Die Mengen-Zahlen-Kompetenz ist bei Schülerinnen und Schülern in der Hauptschule über die Klassen fünf bis neun konstant signifikant schlechter ausgebildet als in der Realschule und dort wiederum signifikant schlechter als im Gymnasium. Zudem befinden sich Hauptschülerinnen und -schüler in der Mengen-Zahlen-Kompetenz in der neunten Klasse unter dem Niveau der Gymnasialschülerinnen und -schüler der fünften Klasse. Signifikante Unterschiede zwischen den Schulformen zeigen sich auch im Bereich des Regel- und Konventionswissens (zum Beispiel Vorzeichen- und Klammerrechnung). Es zeigt sich zudem, dass der Erwerb und die Festigung der einzelnen Kompetenzbereiche bis zum Ende der Sekundarstufe I andauert, da eine steigende Leistungsentwicklung in diesen basalen Bereichen zu erkennen ist. Auch Balzer und Jäger (2001) können in der MARKUS-Untersuchung ihre Annahme bestätigen, dass es signifikante Unterschiede zwischen den Schulformen gibt. „Gymnasiasten schneiden besser ab als Realschüler, diese wiederum sind besser als die Hauptschüler im A-Kurs, welche wiederum besser sind als Hauptschüler aus dem G-Kurs. Es ist jedoch zu erkennen, dass es erhebliche Überschneidungen im Fähigkeitsspektrum der Schüler der verschiedenen Bildungsgänge gibt" (Helmke et al., 2000, S. 4).

Dies zeigt auch die PISA-Untersuchung aus dem Jahr 2015. „An nicht gymnasialen Schularten übertreffen 10 Prozent der Schülerinnen und Schüler den Mittelwert, der für das Gymnasium festgestellt wurde. Außerdem verfügt das leistungsstärkste Viertel der Schülerinnen und Schüler an nicht gymnasialen Schularten über eine höhere mathematische Kompetenz als das leistungsschwächste Viertel der Jugendlichen am Gymnasium" (Hammer et al., 2016, S. 238). Humbach (2008) betrachtet die arithmetischen Basiskompetenzen in der Klasse zehn und zeigt, dass sich Haupt- und Gesamtschülerinnen und -schüler in arithmetischen Aufgaben hauptsächlich auf den Kompetenzstufen I und II (in der Studie von insgesamt IV) bewegen und diese

Schülerinnen und Schüler nur wenige Aufgaben des weiterführenden Schulstoffs lösen können. Diejenigen Schülerinnen und Schüler jedoch, die sich in den arithmetischen Grundlagen auf einer hohen Kompetenzstufe befinden, zeigen auch in den weiterführenden Aufgaben bessere Ergebnisse (Humbach, 2008, S. 175). Der Zusammenhang zwischen den Ergebnissen im Arithmetik-Test und dem *RT 9+* liegt bei r = .71 (Humbach, 2008, S. 105).

Es stellt sich auch die Frage nach dem Einfluss von Kontextfaktoren auf den mathematischen Kompetenzerwerb. Helmke und Kollegen (2000) können belegen, dass es keinen signifikanten Zusammenhang zwischen der Klassengröße und dem Ausmaß an Unterrichtsausfall und dem mathematischen Kompetenzstand der Schülerinnen und Schüler gibt. Ausschlaggebend hingegen ist das Classroom-Management der Lehrkraft, die Mitarbeit im Unterricht (Lernhaltung der Klasse) und die Bildungsnähe der Schülerinnen und Schüler. In der QuaSUM-Studie ist es gelungen, 48 % der Zwischen-Schüler-Varianz durch die erfassten Merkmale aufzuklären (Lehmann, Peek, et al., 2001). Die Leistungen im Fach Mathematik werden in Klasse fünf am stärksten durch die allgemeinen kognitiven Voraussetzungen aufgeklärt (gemessen über den CFT-20; $\beta = 0{,}40$). Zudem wirken sich ein ausgeprägtes Sach- und Fachinteresse für Mathematik ($\beta = 0{,}22$) und ein starkes Selbstvertrauen ($\beta = 0{,}16$) günstig auf die Mathematikleistungen aus. Der Einfluss allgemeiner kognitiver Fähigkeiten ist dort niedriger und liegt bei $\beta = 0{,}26$. Der größte Anteil der Varianz mit $\beta = 0{,}60$ wird über die Schulform bzw. das Kursniveau aufgeklärt. Das mathematische Fachinteresse wirkt mit $\beta = 0{,}22$ auf die erreichten Leistungen im Fach Mathematik.

Auch die Querschnittsuntersuchungen identifizieren die Schülerinnen und Schüler an Hauptschulen als diejenigen, die die größten Schwierigkeiten in der mathematischen Kompetenzentwicklung zeigen. Werden von den Autoren Lösungswahscheinlichkeiten bezogen auf konkrete Aufgaben genannt, sind oftmals massive Probleme im Grundschulstoff im Bereich der mathematischen Basiskompetenzen festzustellen. Bei den dargelegten Querschnittsuntersuchungen ist zu berücksichtigen, dass sich einige Autoren am PISA- oder TIMSS-Vorgehen orientieren und sich bemühen, die erreichten mathematischen Kompetenzen in Kompetenzstufen einzuordnen (siehe z. B. Hoffmann, 2009; Humbach, 2008). Die Kompetenzen werden erhoben und der Output bestimmten Kompetenzniveaus zugeordnet. Dies unterstreicht auch hier, dass der Erwerb mathematischer Kompetenzen in der Sekundarstufe stark curricular gesteuert ist und deswegen vom mathematischen *Erwerb* als von mathematischer *Entwicklung* gesprochen werden muss. Die Zirkularität der Klassifikation von Kompetenzstufen kann ein Grund für die insgesamt vergleichbaren Forschungsbefunde darstellen. Im Hinblick auf ein inklusives Schulsystem ist zu erwarten, dass Kompetenzstufen weiter ausdifferenziert werden müssen, um auch sehr schwache Schülerinnen und Schüler verorten und vergleichen zu können.

Folgende Zusammenfassung gibt einen Überblick über die zentralen Inhalte und Zusammenhänge des dritten Kapitels.

2.5 Zusammenfassung

Die Ausführungen zur Entwicklung *früher* mathematischer Vorläuferfertigkeiten in Kapitel 2.2 zeigen, dass mathematisches Lernen bereits ab der Geburt beginnt. Mathematische Kompetenzen im Vorschulalter können bereits Hinweise auf spätere mathematische Leistungen geben

(siehe z. B. Lerkkanen et al., 2005; Siegler et al., 2012; Stern, 2008) (siehe Kapitel 2.2.3). Zudem scheint es, dass die Erwerbsprozesse in der frühen mathematischen Entwicklung auf den späteren Kompetenzerwerb in der Sekundarstufe I übertragen werden können, sobald große Zahlenräume erschlossen werden müssen (Krajewski & Ennemoser, 2010). Die mathematischen Lernprozesse, die im jungen Alter eher unbewusst erfolgen, werden im schulischen Kontext durch Curricula und Bildungsstandards systematisiert und gesteuert, sodass eher von einem mathematischen Kompetenz*erwerb* gesprochen werden sollte. Der zentrale Ort mathematischen Lernens ist dann der Schulunterricht. Es besteht aktuell ein großes Forschungsdefizit, diesen Kompetenzerwerb in der Primar- und Sekundarstufe in Form von anschlussfähigen Kompetenzstufen oder Erwerbsmodellen empirisch abgesichert abzubilden. Bisherige Kompetenzmodelle nehmen Curricula als theoretische Grundlage und überprüfen anschließend durch den Output, ob diese „Entwicklung" entlang der Kompetenzniveaus und -stufen durch die Schülerinnen und Schüler durchlaufen wird (siehe z. B. Hammer et al., 2016; Hoffmann, 2009; Humbach, 2008; Reiss & Winkelmann, 2008; Wendt, Bos, Selter & Köller, 2012). Neben pädagogisch-psychologischen Ansätzen zum Erwerb mathematischer Kompetenzen versucht die kognitiv-neurowissenschaftliche Perspektive einen Einblick in die *internen* Prozesse beim mathematischen Lernen zu geben (siehe Kapitel 2.3.2). Der Einblick in die interne Verarbeitung von Zahlen auf verschiedenen Repräsentationsebenen (Dehaene, 1992) und Möglichkeiten zur Vermeidung einer Überlastung des Arbeitsgedächtnisses (Cognitive-Load-Theorie, siehe Ayres & Paas, 2012) können wertvolle Hinweise auf die Förderung mathematischer Kompetenzen geben.

Umfassende und gesicherte Forschungsergebnisse gibt es dahingehend, dass viele Schülerinnen und Schüler Defizite in den mathematischen Basiskompetenzen aufweisen (siehe Kapitel 2.3). Dies gilt bereits für den Verlauf der Grundschulzeit, nach deren Ende sich die Schwierigkeit in der Sekundarstufe I manifestieren und es zu Schwierigkeiten im Erwerb mathematischer Kompetenzen kommt (siehe Kapitel 2.4). Befunde aus der Transitionsforschung zeigen zudem, dass Schülerinnen und Schüler einen Leistungsrückgang nach dem Übergang in die Sekundarstufe I erfahren können (Böck & Migutsch, 1974; Eder, 1995; Isakson & Jarvis, 1999; McGee et al., 2003; Sirsch, 2000). Dabei scheint es bzgl. der Leistungsrückgänge keine Unterschiede zwischen einer vierjährigen oder längeren Grundschulzeit zu geben (McGee et al., 2003). Es gibt jedoch auch Studien, die eine stabile Leistungsentwicklung nach der Transition belegen. Die Befunde sind dahingehend uneinheitlich oder gar widersprüchlich (siehe Kapitel 2.4.1).

Einige internationale und nationale Längsschnittstudien, die sich mit dem mathematischen Kompetenzerwerb im Übergang Primar- und Sekundarstufe I beschäftigen, zeigen, dass die Kompetenzentwicklung in der Sekundarstufe I stetig voranschreitet und ein Ausbau der mathematischen Basiskompetenzen (v. a. im Bereich Rechenoperationen) auch über die weiteren Schuljahre innerhalb der Sekundarstufe I erfolgt (s Kapitel 2.4). Nationale Studien fokussieren oftmals auf den Erwerb mathematischer Kompetenzen in Abhängigkeit verschiedener Schulformen. Es zeigt sich erwartungsgemäß, dass der Lernzuwachs an Gymnasien am stärksten ist und dass Haupt-, Real- oder Gesamtschulen dahingehend weniger effektiv sind (siehe Kapitel 2.4.2.1). Prekär in dem Zusammenhang ist die Erkenntnis, dass dies auch für diejenigen Schülerinnen und Schüler gilt, die am Ende der Grundschulzeit über vergleichbare Kompetenzen verfügen und dann lediglich unterschiedliche Schulformen besuchen (Hammer et al., 2016; Lehmann, Peek, et al., 2001).

Die Untersuchung mathematischer Kompetenzen innerhalb der Sekundarstufe I durch zahlreiche nationale und internationale Querschnittstudien zeigt, dass es gerade in der Lösung von Aufgaben zu den Grundrechenarten eine große Anzahl an Schülerinnen und Schülern gibt, die nicht über diese grundlegenden Kompetenzen verfügt. Es zeigt sich, dass die mathematischen Rückstände offensichtlich nicht frühzeitig erkannt werden und sich diese im Laufe der Schulzeit manifestieren (siehe Kapitel 2.3.3). Die Ergebnisse fordern eine bessere Strukturierung des Übergangs von der Primarstufe in die Sekundarstufe I.

Es stellt sich in der Folge die Frage, wie diese Schwierigkeiten entstehen, welche Faktoren diese Schwierigkeiten hemmen oder gar begünstigen und inwiefern eine Förderung zur Lösung des Problems beitragen kann. Diese Fragen werden im folgenden Kapitel 3 aufgegriffen, in dem das Themenfeld „Schwierigkeiten in der Entwicklung und im Erwerb mathematischer Kompetenzen“ betrachtet wird. Das Kapitel 3 versucht auf Basis der vorangestellten Forschungsergebnisse Antworten auf die Frage des erschwerten Kompetenzerwerbs in der Sekundarstufe I zu geben. Es gilt zu klären, inwieweit endogene und exogene Faktoren der Schülerinnen und Schüler Einfluss nehmen auf deren mathematischen Kompetenzerwerb. Es dient als Grundlage für das darauffolgende Kapitel 4, in dem sachlogisch aufbauend fundierte Methoden und Programme zur wirksamen Förderung mathematischer Kompetenzen dargestellt werden.

3 Schwierigkeiten in der Entwicklung und im Erwerb mathematischer Kompetenzen

„Kannst du mir sagen, Sokrates, ob die Rechenschwäche geheilt werden kann?
Oder ob sie nicht geheilt, sondern durch Übung eliminiert werden kann?
Oder ob sie weder durch Übung noch durch Heilung
eliminiert wird, sondern von den Menschen
einwohnt [...] und wieder verschwindet?“

(Meyer, 1993; In: Was sagst du zur Rechenschwäche, Sokrates?)

In diesem Kapitel 3 stehen Schwierigkeiten in der Entwicklung und im Erwerb mathematischer Kompetenzen in der Bandbreite ihrer Ausprägungen im Zentrum der Betrachtung. Dafür werden eingangs Begrifflichkeiten geklärt, im Anschluss Angaben zu den Prävalenzraten von Rechenschwierigkeiten und -störungen dargelegt, mögliche Ursachen und Komorbiditäten aufgeschlüsselt und letztlich das diagnostische Vorgehen sowie Grundätze der Förderung vorgestellt.

3.1 Begrifflichkeiten

In der Fachliteratur sind im Bereich der allgemeinen Sonderpädagogik aber auch in der Mathematikdidaktik vielfältige Begrifflichkeiten formuliert worden, die zu beschreiben versuchen, wie der erschwerte Prozess des Mathematiklernens zu bezeichnen ist und welche Phänomene sich dahinter verbergen. In diesem Kapitel 3.1 wird in einem Zweischritt eine begriffliche Annäherung vorgenommen: Zuerst werden zentrale Begriffe im allgemeinen Themenfeld „Schwierigkeiten im Lernen“ definiert, um im Anschluss Termini der „Schwierigkeiten im Rechnen“ einordnen und präzisieren zu können.

3.1.1 Schwierigkeiten im Lernen

Die Diskussion um korrekte und angemessene grundlegende Begrifflichkeiten im Bereich der Sonderpädagogik ist kein Phänomen des 21. Jahrhunderts. So schrieb Zielinski (1980, S. 14) bereits vor mehr als drei Jahrzehnten:

> „Ist Schulversagen [...] ein Schicksal, von dem jährlich Hunderttausende von Schülern betroffen sind, so ist zu fragen, ob und wie das den verschiedenen Formen des Schulversagens Gemeinsame begrifflich zu fassen ist, denn bei der Durchsicht der Literatur fällt auf, dass zur Kennzeichnung dieses Phänomens eine Vielzahl von Begriffen verwendet wird, die teils synonym gebraucht, teils ausdrücklich voneinander abgehoben werden. So ist etwa die Rede von Lernbehinderung (Kobi, 1975), Leistungsversagen (Löwe, 1963), Leistungshemmung (Rösler, 1963) oder Lernstörungen (Correll, 1964); man spricht vom leistungsschwachen Schüler (Kemmler, 1967), vom schlechten Schüler oder einfach vom Schulversager (Höhn, 1967).“

Auch heute noch werden die bei Zielinksi (1980) aufgeführten Begriffe im alltäglichen und wissenschaftlichen Gebrauch uneinheitlich und zum Teil synonym verwendet (Souvignier, 2008). Weitgehende Einigkeit scheint es bei dem Begriff der „Lernschwierigkeit“ zu geben (Gold, 2011; Heimlich, 2009; Lauth, Brunstein & Grünke, 2014; Zielinski, 1998). Gefasst werden unter dem Begriff alle denkbaren Möglichkeiten, bei denen eine „erschwerte oder verhin-

derte Lernzielerreichung“ (Heinecke-Müller, 2014) vorliegt. Der Begriff „Lernschwierigkeiten“ stellt also eine Art Oberbegriff für verschiedene Formen des erschwerten Lernens dar. Die Idee der Präzisierung könnte dabei auf Zielinksi (1998) zurückgehen:

> „Von Lernschwierigkeiten spricht man im Allgemeinen, wenn die Leistungen eines Schülers unterhalb der tolerierbaren Abweichungen von verbindlichen, institutionellen, sozialen und individuellen Bezugsnormen (Standards, Anforderungen, Erwartungen) liegen oder wenn das Erreichen (bzw. Verfehlen) von Standards mit Belastungen verbunden ist, die zu unerwünschten Nebenwirkungen im Verhalten, Erleben oder in der Persönlichkeitsentwicklung des Lernenden führen.“

Die Definition verdeutlicht, dass zum Beispiel auch ein Gymnasialschüler in bestimmten Bereichen (zum Beispiel in der Kurvendiskussion) Lernschwierigkeiten haben kann, obwohl er mit seinen Gesamtleistungen dennoch über der Durchschnittsleistung seiner Alters- bzw. Bezugsgruppe liegt. Die internationale Übersetzung für Lernschwierigkeiten lautet „learning difficulties“ . Dieser Begriff findet gerade im inklusiven Kontext verstärkt Anwendung, da er „[den inklusiven Anspruch] nach Aufhebung der strikten Trennung zwischen Menschen mit und ohne Behinderung [erfüllt]“ (Heimlich & Wember, 2014, S. 54).

Einen Überblick über die verschiedenen Formen von Lernschwierigkeiten geben Lauth und Kollegen (2014) mit ihrer, in der Literatur oft rezipierten, tabellarischen Übersicht. Die Autoren sprechen jedoch nicht von „Lernschwierigkeiten“, sondern wählen als Oberbegriff eher unüblicherweise „Lernstörungen“ – als Synonym für „Minderleistungen absichtsvollen Lernens“ (Lauth et al., 2014, S. 17).

TABELLE 9 ARTEN VON LERNSTÖRUNGEN (LAUTH ET AL., 2014, S. 18)

	Bereichsspezifisch (partiell)	**Allgemein (generell)**
Vorübergehend (passager)	Lernrückstände in Einzelfächern	Schulschwierigkeiten Neurotische Störung
Überdauernd (persistierend)	Lese- Rechtschreibschwäche Rechenschwäche	Lernschwäche Lernbehinderung Lernbeeinträchtigung Geistige Behinderung

Die Einordnung der Lernstörungen erfolgt dabei über zwei Dimensionen: Umfang (bereichsspezifisch vs. allgemein) und Dauer (vorübergehend vs. überdauernd). Als vorübergehend und bereichsspezifisch sehen die Autoren Lernrückstände in Einzelfächern an, wenn zum Beispiel ein bestimmtes Themengebiet (Dreisatz) nicht beherrscht wird. Als allgemein und generell werden Schulschwierigkeiten oder neurotische Störungen gesehen. Denkbar wäre bei dieser Form der Lernstörung zum Beispiel eine Trennung der Eltern, der Verlust einer nahestehenden Person oder Phasen einer Depression. Als überdauernde und bereichsspezifische Störungen werden die (Lese-)Rechtschreibschwäche (ICD-10, F81.0 & F81.1) oder die Rechenschwäche (ICD-10, F81.2) angesehen. Diese Störungen betreffen dabei im Regelfall nur eine spezifische schulische Domäne bei sonstigen eher durchschnittlichen schulischen und kognitiven Leistungen (sog. Diskrepanzkriterium). Darüber hinaus gibt es überdauernde und allgemeine Lernstörungen, die von den Autoren als Lernschwäche (ICD-10, F81.3), Lernbehinderung (ICD-10,

F81.9), Lernbeeinträchtigung oder auch als eine geistige Behinderung (ICD-10, F70) bezeichnet werden.

Heimlich und Wember (2014, S. 51) betiteln, angelehnt an Kanter (1977), die Lese-Rechtschreibstörung und die Rechenstörung (Dyskalkulie) als Lernstörungen. Sie tauchen bei durchschnittlicher Intelligenz nur in einem Lernbereich auf, es liegt nicht mehr als ein Schuljahr Rückstand vor und ihnen kann in der allgemeinen Schule mit einer individuellen Förderung begegnet werden. Eine Lernbehinderung liegt dann vor, wenn mehr als ein Lernbereich bei unterdurchschnittlicher Intelligenz betroffen ist und sich die Schwierigkeiten über einen Zeitraum von mehr als einem Schuljahr erstrecken. Als Oberbegriff für Lernstörungen und Lernbehinderung wählen Heimlich und Wember (2014, S. 51) den Begriff „Lernbeeinträchtigungen".

Schröder (2005, S. 88 f.) bemängelt die Unvollständigkeit dieser Klassifizierung und präferiert, angelehnt an Bach (1977), ein dreidimensionales Modell, in dem die Merkmalsdimensionen „Schwere", „Umfang" und (prognostische) „Dauer" Berücksichtigung finden. „Immer dann, wenn nicht alle drei Merkmale der Lernbehinderung erfüllt sind, d. h. bei sieben Kombinationen, wird man noch von Lernstörungen sprechen müssen" (Schröder, 2005, S. 88).

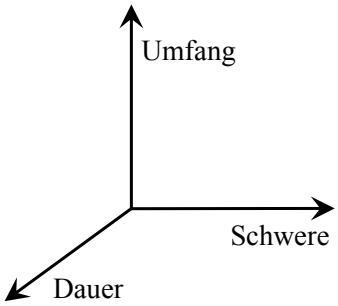


ABBILDUNG 18 DREIDIMENSIONALE TERMINOLOGIE VON LERNBEEINTRÄCHTIGUNGEN (SCHRÖDER, 2005, S. 88 F.; NACH BACH, 1977)

Das dreidimensionale Modell stellt eine Erweiterung des o. g. Modells nach Lauth und Kollegen (2014, S. 18) dar, das aus Sicht Schröders unzureichend ist, da ein Kontinuum nicht linear bestimmbar sein kann und zudem in der Praxis Merkmalskombinationen zu finden sind, die in dem Modell keine Verortung finden können. Das dreidimensionale Modell „hält eine vernünftige Mitte zwischen Begriffslosigkeit und dem [...] [bisher; d. Verf.] dokumentierten Wirrwarr" (Schröder, 2005, S. 89).

Unabhängig der dargelegten Klassifikationen ist den betroffenen Schülerinnen und Schülern gemein, dass „Lernabsichten und Lernanstrengungen [nicht] zum gewünschten Erfolg führen" (Gold, 2011, S. 33). Gold (2011, S. 33 ff.) sieht als zentrale Bedingungen erfolgreichen Lernens die Zuwendung von Aufmerksamkeit, die Funktionsfähigkeit des Arbeitsgedächtnisses, die Aktivierung von Vorwissen, Lernstrategien, Motivation und ein positives Selbstkonzept sowie die Steuerung lernbegleitender Emotionen und der Volition.

Matthes (2006, S. 26) hat in der Frage des Bedingungsgefüges von Lernstörungen ein Handlungsmodell entwickelt (siehe Abbildung 19), in dem als Voraussetzung erfolgreichen Lernens die Bereiche der Informationsverarbeitung und der Lernaktivität unterschieden werden.

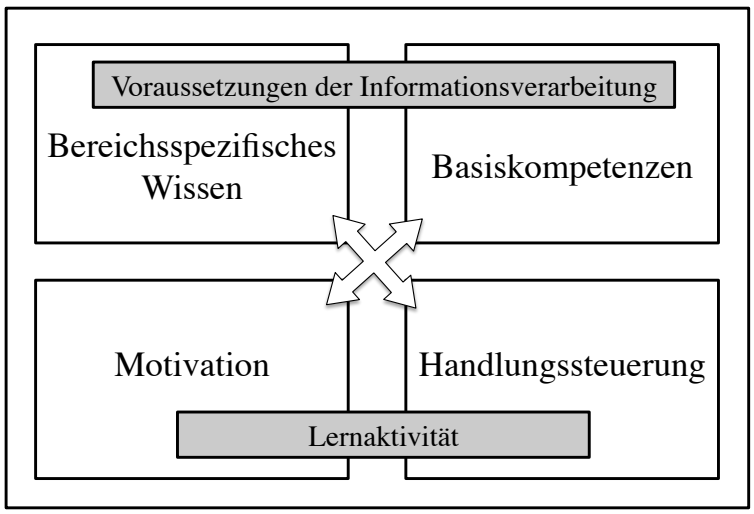


ABBILDUNG 19 HANDLUNGSMODELL BEI LERNSTÖRUNGEN (MATTHES, 2006, S. 26)

Auf Ebene der Informationsverarbeitung geht es um die Voraussetzung bereichsspezifischen Wissens und ausgebildeter Basiskompetenzen für das Behalten und Anwenden wichtiger Lerninhalte. Unter bereichsspezifischem Wissen fasst der Autor zum Beispiel die Wissensbasis im Bereich Mathematik, Lesen, Sachkunde oder anderen Lebensbereichen. Durch das Fehlen bereichsspezifischen Wissens bei Schülerinnen und Schülern mit Lernstörungen ist der Erwerb und die Verknüpfung neuer Inhalte deutlich erschwert. Mit Basiskompetenzen sind „grundlegende Fertigkeiten der Informationsverarbeitung" (Matthes, 2006, S. 27) gemeint, die sich eher auf die funktionale Aspekte der Wahrnehmung beziehen. Gemeint sind zum Beispiel soziale, motorische, visuell-auditive oder kognitive Fertigkeiten.

Der Bereich der Lernaktivität (siehe Abbildung 19) umfasst die Motivation und die Handlungssteuerung. Zum Bereich der Motivation gehört dabei u. a. ein gutes Selbstwertgefühl, Lernfreude, Unlustvermeidung, ein angemessenes Leistungsanspruchsniveau oder auch eine angemessene positive Selbstbewertung. Unter dem Begriff Handlungssteuerung fällt dabei die Metakognition, Impulskontrolle und auch die Erregungsregulation. Die Steuerung des eigenen Handelns oder auch von Teilhandlungen dient dem Erreichen eines anvisierten Ziels. Auch in dem Bereich der Lernaktivität haben Schülerinnen und Schüler mit Lernstörungen oftmals Schwierigkeiten darin, metakognitive Prozesse zu steuern, Emotionen im Lernen zu regulieren, eine angemessene Erwartung von Erfolg und Misserfolg zu zeigen oder auch eine angemessene Selbstbewertung durchzuführen (Matthes, 2006, S. 30 f.).

Das Handlungsmodell verdeutlicht das komplexe Bedingungsgefüge von Lernstörungen, in dem die einzelnen Komponenten miteinander verknüpft sind, sich gegenseitig bedingen oder beeinflussen. Oftmals finden sich die Schwierigkeiten bei Schülerinnen und Schülern mit Lernstörungen in mehreren Bereichen zugleich (Linderkamp & Grünke, 2007).

Die Ausführungen zeigen, dass Schwierigkeiten im Lernen sehr unterschiedlich definiert sein können und sich im Schweregrad unterscheiden. Trotz des Versuchs einer Systematisierung bleiben Überschneidungen und synonyme Verwendungen einzelner Begriffe nicht aus.

3.1.2 Schwierigkeiten im Rechnen

Im folgenden Kapitel wird eine, dem obigen Kapitel vergleichbare Begriffspräzisierung für den Bereich der Schwierigkeiten im Rechnen vorgenommen. Auch hier besteht das Problem, dass viele Begriffe uneinheitlich und unpräzise verwendet werden (Grube, 2008, S. 642; Lambert, 2015, S. 66).

3.1.2.1 Rechenschwierigkeiten

Grube (2008, S. 642) schreibt, dass Rechenschwierigkeiten als Oberbegriff für verschiedene Formen des erschwerten Rechnens auf unterschiedlichen Niveaus gesehen werden können. „Aus der mathematikdidaktischen Perspektive liegen die Schwierigkeiten im Erlenen des Rechnens in einem Kontinuum zwischen mathematischen Spitzenleistungen einerseits und absolutem Versagen beim Rechnen andererseits“ (Schipper, 2010, S. 105). So kann auch ein durchschnittlich guter Rechner in bestimmten mathematischen Teilgebieten die an ihn gerichteten Erwartungen nicht erfüllen. Gründe dafür können u. a. inadäquater Unterricht oder mangelnde Motivation sein. Diese Ausprägung der Rechenschwierigkeit beschreibt Grube (2008, S. 643) als „deskriptiv“, weil sie eher pragmatisch angelegt ist (i. S. der Bestimmung einer „Risikogruppe“) und dabei eher weniger das Eruieren spezifischer Ursachen angestrebt wird. Fritz und Ricken (2008, S. 14) fassen unter den Begriff der Rechenschwierigkeiten „alle Kinder, [...] bei denen sich Probleme beim Rechnenlernen bereits von Beginn der ersten Klasse an zeigen“. Fischbach und Kollegen (2013, S. 66) schreiben, dass eine Abweichung von der Lernleistung der Norm im Lesen, Rechtschreiben und/oder Rechnen *unberücksichtigt* der Intelligenz als „einfaches Diskrepanzkriterium“ zu bezeichnen ist. Fritz und Ricken (2008, S. 9) weisen in diesem Zusammenhang auf die Ergebnisse der vorletzten PISA-Studie aus dem Jahr 2012 hin, die zeigt, dass sich 17,7 % der 15-Jährigen unter oder auf der Kompetenzstufe I von V befinden (Prenzel et al., 2012, S. 75) und sich dabei die Frage auftut, ob diese Schülerinnen und Schüler tatsächlich *alle* von einer schwerwiegenden Rechenstörung (i. S. einer Dyskalkulie) betroffen seien können. Auch Landerl und Kaufmann (2013, S. 96) schreiben, „es macht wenig Sinn, davon auszugehen, dass 25 bis 30 % der Bevölkerung von einer umschriebenen Lernstörung [i. S. einer Dyskalkulie; d. Verf.] betroffen sind, vielmehr sollte hier von unterdurchschnittlichen Rechenleistungen gesprochen werden“.

3.1.2.2 Rechenstörung

Der Begriff Rechenstörung wird oftmals synonym verwendet mit den Begriffen Rechenschwäche, Dyskalkulie und auch Arithmasthenie. Der Begriff Dyskalkulie setzt sich zusammen aus der griechischen Vorsilbe „dys“ (abweichend von der Norm) und „calculare“ (mit Rechensteinen rechnen) (Zimmermann, 2014, S. 33).

Als englischsprachiges Pendant gilt oftmals der Begriff „arithmetic difficulties“ (Grube, 2008, S. 642). In der ICD-10 (Weltgesundheitsorganisation, 2013) wird die Rechenstörung unter der Kennziffer F81 (Umschriebene Entwicklungsstörungen schulischer Fertigkeiten) und explizit unter F81.2 (Rechenstörung; engl. Version „specific disorder of arithmetical skills“) aufgeführt und wie folgt beschrieben:

> „Diese Störung besteht in einer umschriebenen Beeinträchtigung von Rechenfertigkeiten, die nicht allein durch eine allgemeine Intelligenzminderung oder eine unangemessene Beschulung erklärbar ist. Das Defizit betrifft vor allem die Beherrschung grundlegender Rechenfertigkeiten, wie Addition, Subtraktion, Multiplikation und Division, weniger die höheren mathematischen Fertigkeiten, die für Algebra, Trigonometrie, Geometrie oder Differential- und Integralrechnung benötigt werden.

Als eine Unterform der Rechenschwierigkeit wird die Rechenstörung dadurch abgegrenzt, dass die unterdurchschnittlichen mathematischen Leistungen nicht durch zum Beispiel fehlende Motivation, mangelnde Lernmöglichkeiten, Seh- oder Hörbeeinträchtigungen oder eine Intelligenzminderung erklärt werden können. Rechenstörungen treten nicht plötzlich auf, bestehen langfristig und sind auch nur schwer zu beeinflussen (Grube, 2008, S. 643). Bezogen auf die Klassifizierung Lauths und Kollegen (2014) (Tabelle 9) ist eine Rechenstörung als „bereichsspezifisch“ und „überdauernd“ zu bezeichnen. Grube (2008, S. 643) bezeichnet die Rechenstörung als einen „explikativen Begriff“, da es hierbei um die Erforschung spezifischer individueller Ursachen der betroffenen Schülerinnen und Schüler gehe. Die inhaltlichen Probleme bei einer Rechenstörung betreffen weniger höhere mathematische Fertigkeiten wie Algebra oder Differentialrechnung, sondern beziehen sich auf grundlegende Fertigkeiten wie die Rechenfertigkeiten in den Grundrechenarten (Lorenz, 2014).

3.1.3 Begrifflichkeiten international

Der Vergleich deutscher und internationaler Forschungserkenntnisse zum Themengebiet Schwierigkeiten im Lernen ist auf Grund divergierender Begrifflichkeiten stark erschwert. Die Schülergruppe mit einem IQ zwischen 55 und 85 wird hierzulande mit dem Begriff „Lernbehinderung“ umschrieben (Grünke & Grosche, 2014). Der Begriff ist dabei eng verknüpft mit dem bisher (noch) stark ausgebauten Förderschulwesen (Schwerpunkt Lernen), wohingegen Schülerinnen und Schüler mit einem Unterstützungsbedarf im Bereich Lernen international oftmals inklusiv beschult werden (Souvignier, 2008, S. 664). Das englischsprachige Pendant für „Lernbehinderung“ ist in etwa der Begriff „Intellectual Disabilities“, wie ihn zum Beispiel die DSM-V verwendet (American Psychiatric Association, 2013). Die ICD-10 behält sich keine eigene Klassifizierung für den Bereich der Lernbehinderung vor, sodass diese i. e. S. unter F70 (Leichte Intelligenzminderung; IQ liegt zwischen 50 und 69) fallen würde. Liegen sowohl Rechen- als auch Lese- bzw. Rechtschreibschwierigkeiten vor, gibt es zudem in der ICD-10 die Restkategorie F81.3 (Kombinierte Störungen schulischer Fertigkeiten). Der Begriff „Learning Disabilities“ ist international sehr viel weiter verbreitet, sodass darunter auch zum Beispiel die Rechenstörung oder die Lese-Rechtschreibstörung fallen würden (Harwell & Williams Jackson, 2008; Swanson, Harris & Graham, 2006). Diese zählen in Deutschland jedoch nicht zu einer Lernbehinderung, sondern werden als Teilleistungsschwächen bezeichnet. Der am häufigsten synonym verwendeter Begriff zu „Learning Disability“ ist „Learning Disorder“ (Zielinski, 1980, S. 15). Die definitorischen Diskrepanzen geben eine Erklärung dafür, dass die Prävalenzraten von „Learning Disabilities“ weitaus größer sind als die der „Lernbehinderung“. Im Kontext Mathematik schlägt Peard (2010, S. 107) eine Unterscheidung in slow learners, learned disability in mathematics und learning disability in mathematics vor, um die verschiedenen Ursachen von Rechenschwierigkeiten deutlicher herauszustellen.

3.2 Prävalenz und Komorbiditäten

PRÄVALENZ

Da der Fokus in dieser Arbeit auch auf Schülerinnen und Schüler liegt, die allgemein Schwierigkeiten im Rechnen haben, ohne dass eine schwerwiegende Dyskalkulie vorliegt, wird versucht, auch Prävalenzraten für diese allgemeinen Rechenschwierigkeiten aufzustellen. Leider gibt es diesbezüglich nur sehr wenige Angaben (vgl. Kapitel 3.1.2.1). Lorenz (1993) schreibt, dass ca. 15 % der Grundschülerinnen und -schüler förderbedürftige Rechenschwierigkeiten haben. Eine weitere Angabe stammt von Mazzocco und Myers (2003). In ihrer Längsschnittstudie finden die Autoren heraus, dass ca. 17 % der Schülerinnen und Schüler mindestens einmal in den Klassen eins bis vier Rechenschwierigkeiten zeigen. Fischbach und Kollegen (2013) zeigen in ihrer Studie, dass 23,3 % der untersuchten 2.195 Zweitklässlerinnen und -klässler trotz durchschnittlicher Intelligenz (IQ > 85) Lernschwierigkeiten in einem oder mehreren Leistungsbereichen (Lesen, Schreiben, Rechnen) haben. Interessant sind hier die Ergebnisse im Detail: Lediglich 5,0 % der Schülerinnen und Schüler zeigen eine isolierte Rechenschwäche (Rechenschwierigkeiten), bei der die Leistungen im Mathematiktest (DEMAT 2+) unterhalb der Norm liegt. Heimlich (2009) geht hingegen davon aus, dass 6 % bis 15 % der Schülerinnen und Schüler einen Unterstützungsbedarf im Bereich der grundlegenden mathematischen Operationen haben.

Zahlreiche Angaben gibt es hingegen zu den Häufigkeiten von Rechenstörungen (im Sinne einer Dyskalkulie). So kommt ein Großteil der nationalen und internationalen epidemiologischen Studien auf eine Prävalenz von etwa 6 %. Ein „Ausreißer" stellt dabei die Studie von Fischbach und Kollegen (2013) dar, die unter Berücksichtigung des doppelten Diskrepanzkriteriums zu einer Prävalenzrate einer isolierten Rechenstörung von 2,6 % kommen. Auch Peard (2010) statuiert eine geringere Prävalenz von nur 2 %, jedoch im beruflichen Kontext. Folgende Tabelle 10 gibt einen Überblick über Prävalenzraten ab dem Jahr 2000.

TABELLE 10 PRÄVALENZ VON RECHENSTÖRUNGEN (ANGELEHNT AN JACOBS & PETERMANN, 2007; ERWEITERTE ZUSAMMENSTELLUNG D. VERF.)

Autoren	N	Klasse	Land	Prävalenz
Hein, Bzufka & Neumärker (2000)	181	3	Deutschland	6,6 %
Ramaa & Gowramma (2002)	251 1.408	Primarstufe	Indien	5,98 % 5,54 %
Mazzocco & Myers (2003)	209	Vorschule bis Klasse 3	USA	9,6 %
Koumoula u. a. (2004)	240	1–4	Griechenland	6,3 %
M. Van Aster u. a. (2007)	337	Vorschule & Klasse 2	Deutschland	6,0 %
Peard (2010)	2	beruflicher Kontext	Australien	2,0 %
Fischbach u. a. (2013)	2.195	2/3	Deutschland	2,6 %

Grube (2008, S. 647) konstatiert trotz nahezu einheitlicher Forschungslage, dass es für Deutschland grundsätzlich nicht möglich sei, Prävalenzraten zu bestimmen, solange die Diagnosekriterien so unterschiedlich seien (vgl. auch Hasselhorn & Schuchardt, 2006). „Anders als bei vielen Krankheiten kann die Vorkommenshäufigkeit von Rechenschwächen kaum deutlich reduziert

werden, weil diese aufgrund der festgesetzten definierenden Kriterien „nachwachsen“: Es wird immer Kinder geben, die die 3 % am schlechtesten bewerteten Rechenleistungen erzielen“ (Grube, 2008, S. 647). Auch Mazzocco und Myers (2003, S. 249) schreiben „It is not possible to provide a single frequency figure of MD [mathematical disabilities; d. Verf.] because our findings yield a wide range of frequency counts dependent on which MD criteria are used“. Peard (2010, S. 107) schreibt, dass ein Großteil der Schülerinnen und Schüler eher als „slow learners“ bezeichnet werden müssten bzw. eine „learn*ed* disability“ aufweisen, sodass der Anteil derjenigen mit einer neurologisch bedingten Rechenstörung (im Sinne einer learn*ing* disability) bei weniger als 2 % liegen müsste (siehe Tabelle 10).

Neben Prävalenzraten wird oftmals der Zusammenhang zwischen einer Rechenstörung und dem Geschlecht untersucht. Bzgl. allgemeiner Mathematikleistungen zeigt die PISA-Studie aus dem Jahr 2015, dass die Mädchen signifikant geringere Leistungen erzielen als die Jungen (Hammer et al., 2016, S. 240). Hinsichtlich einer Rechenstörung weisen einige Studien jedoch darauf hin, dass es keinen Unterschied in der Häufigkeit zwischen Jungen und Mädchen gibt (Ehlert, Schroeders & Fritz, 2012; Jacobs & Petermann, 2003). Koumoula und Kollegen (2004) (siehe Tabelle 10) berichten für ihre Untersuchung, dass weder durch die Einschätzung der von ihnen befragten Lehrkräfte noch durch Testergebnisse mittels der „Neuropsychological Test Battery for Number Processing and Calculation“ (NUCALC) signifikante Unterschiede in der Häufigkeit einer Rechenstörung im Abhängigkeit vom Geschlecht zu finden sind. Auch Landerl und Kaufmann (2013) sowie Jacobs und Petermann (2007) statuieren, dass offenbar ebenso viele Jungen wie Mädchen betroffen sein können. Zu anderen Ergebnissen kommen wiederum Fischbach und Kollegen (2013), bei denen das ermittelte Verhältnis von Rechenschwierigkeiten bei 1 : 1,80 und bei einer isolierten Rechenstörung bei 3,31 : 1 (Mädchen : Jungen; $p < .01$) liegt. Dieser Trend wird in dem Literaturreview von Krizinger & Günther (2013) und durch weiteren Studien bestätigt (Hein et al., 2000). Die TIMS-Studie spiegelt die Verteilung bezüglich allgemeiner Rechenschwierigkeiten (1 : 1,80) wider (Wendt et al., 2012). Eine Verteilung von 60 : 40 (Mädchen : Jungen) ermittelte vor einigen Jahren Lobeck (1992, S. 91). Die Forschungslage scheint dahingehend als wenig einheitlich.

Komorbiditäten

Viele Studien zeigen zudem, dass neben unzureichenden Mathematikleistungen auch Probleme beim Lesen und Schreiben vorliegen können. Dabei scheint der Zusammenhang zwischen Rechen- und Lesestörungen besser untersucht zu sein als der zwischen Rechen- und Rechtschreibstörungen (Moll & Landerl, 2011). Hasselhorn und Schuchardt (2006, S. 213) gehen davon aus, dass eine Rechenschwäche (hier: Mathematikstörung; MS) bei Jungen und Mädchen grundsätzlich gleich häufig auftaucht (siehe Abbildung 20). Bei der doppelten Problemlage (Komorbidität) hingegen gehen die Autoren davon aus, dass mehr Mädchen betroffen sind.

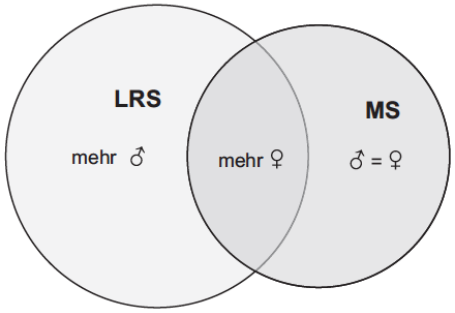


ABBILDUNG 20 GESCHLECHTERUNTERSCHIEDE IM RISIKO DES AUFTRETENS VON LRS UND DYSKALKULIE SOWIE DES GEMEINSAMEN AUFTRETENS (HASSELHORN & SCHUCHARDT, 2006, S. 213)

Van Aster und Kollegen (2007) berechnen in ihrer Untersuchung (N = 337), dass die Prävalenzrate für eine Rechenstörung erwartungsgemäß bei insgesamt 6 % liegt (siehe Tabelle 10). 70 % der identifizierten Schülerinnen und Schüler zeigen in ihrer Untersuchung gleichzeitig eine Kombination einer Lese-Rechtschreibstörung, lediglich 30 % zeigen eine rein isolierte Rechenstörung. Schwenck und Schneider (2003) untersuchen den Zusammenhang zwischen mathematischen und Lese-Rechtschreibkompetenzen bei 115 Erstklässlerinnern und -klässlern. 20 der Schülerinnen und Schüler (17,3 %) zeigen neben unterdurchschnittlichen mathematischen Kompetenzen (im DEMAT1+) auch unterdurchschnittliche Kompetenzen im Lesen (Würzburger Leise Leseprobe) und Rechtschreiben (Wortdiktat), 27 Schülerinnen und Schüler (23,4 %) zeigen durchschnittliche Lese-Rechtschreibleistungen bei unterdurchschnittlichen Mathematikleistungen. Die Prävalenzrate einer kombinierten Lernstörung liegt bei Fischbach und Kollegen (2013) bei lediglich 2,0 %.

Eine geringe Lesekompetenz scheint mit einer unzureichenden mathematischen Kompetenzentwicklung einherzugehen (Gersten, Jordan & Flojo, 2005). Schülerinnen und Schüler mit einer geringen Lesekompetenz *und* Rechenschwierigkeiten zeigen vor allem Schwierigkeiten – ebenso wie die Schülerinnen und Schüler, die ausschließlich Rechenschwierigkeiten aufweisen – im Bereich der Rechenoperationen (Jordan, Hanich & Kaplan, 2003). Dass es weiterhin notwendig ist, Schülerinnen und Schüler mit einer kombinierten Lernstörung und einer isolierten Rechenstörung getrennt zu betrachten, zeigt die Untersuchung von Schuchardt und Mähler (2010). Die Autorinnen vergleichen u. a. die Arbeitsgedächtniskapazitäten verschiedener Gruppen und finden heraus, dass Schülerinnen und Schüler mit einer isolierten Rechenstörung schwache visuell-räumliche Arbeitsgedächtnisleistungen zeigen, wohingegen diejenigen mit einer kombinierten Störung (zusätzlich LRS) auch Defizite in der phonologischen Schleife und der zentralen Exekutive aufweisen. Inwiefern neurologische und/oder umweltbezogene Aspekte ursächlich für eine komorbide Störung sind, ist nicht abschließend geklärt (Moll & Landerl, 2011).

Rechenschwierigkeiten können zudem, wie allgemeine Schwierigkeiten im Lernen, mit Problemen im Arbeits- und/oder Sozialverhalten einhergehen (Stein, 2014, S. 110). Dabei ist „nicht immer zu erkennen [...], welche Störung am Beginn der Fehlentwicklungsprozesses stand, oder ob sich nicht beide Störungen in einem gemeinsamen Prozess manifestierten“ (Myschker, 2005, S. 63). Der Autor schreibt, dass „Lern- und Verhaltensstörungen häufig miteinander (kovariieren)“ (Myschker, 2005, S. 63), wobei zu berücksichtigen ist, dass „Auffälligkeiten im Bereich des Lernens und Auffälligkeiten im Bereich des Verhaltens [...] bei ein und demselben Kind

oder Jugendlichen [...] bislang nur ganz selten der Fokus wissenschaftlicher Untersuchungen [sind]“ (Schröder & Wittrock, 2002, S. 41).

3.3 Ätiologie

Um sich zum einen mit der Diagnostik und zum anderen fundiert mit Möglichkeiten schulischer Förderung bei Rechenschwierigkeiten auseinanderzusetzen, ist es von großer Bedeutung, die konkreten Ursachen für erschwerte Lernprozesse zu ergründen.

Bereits in den vorherigen Kapiteln liegt der Fokus nicht nur auf der schwerwiegende Lernstörung der Dyskalkulie, sondern ebenso auf allgemeinen Lern- und Rechenschwierigkeiten. In diesem Kapitel werden zu Beginn Ursachen von allgemeinen Lernschwierigkeiten erläutert und im Anschluss aktuelle Ergebnisse zur Ursachenforschung einer Rechenstörung dargelegt.

3.3.1 Lern-/Rechenschwierigkeiten

Ein umfassendes und anschauliches Modell der Ursachen von Lernschwierigkeiten stammt von Heimlich (2009, S. 35) und ist in folgender Abbildung 21 dargestellt.

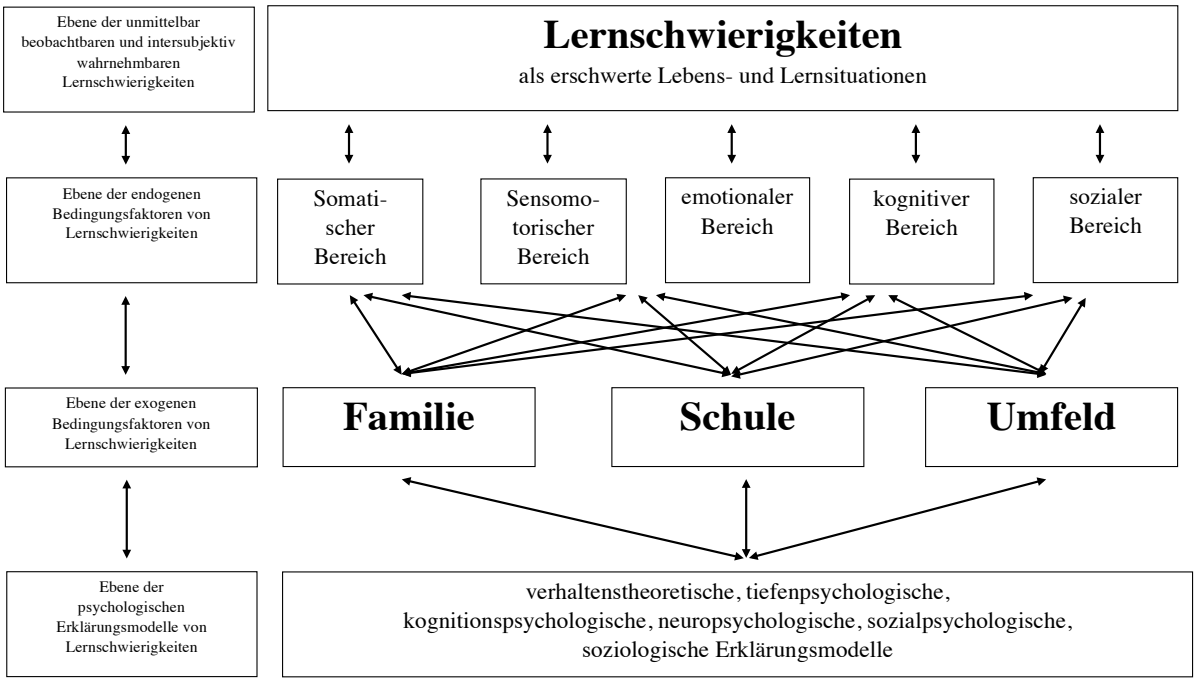


ABBILDUNG 21 ERKLÄRUNGSMODELLE VON LERNSCHWIERIGKEITEN (HEIMLICH, 2009, S. 35)

In dem Erklärungsmodell von Lernschwierigkeiten werden von unten nach oben vier Ebenen unterschieden: Die Ebene der psychologischen Erklärungsmodelle von Lernschwierigkeiten, die Ebene der exogenen Bedingungsfaktoren, die Ebene der endogenen Bedingungsfaktoren und letztlich die unmittelbar beobachtbaren und intersubjektiv wahrnehmbaren Lernschwierigkeiten. Es geht also folglich nicht nur um Kernursachen, sondern ebenso um Bedingungsfaktoren, die das Auftreten von Lernschwierigkeiten erhöhen. Unter Lernschwierigkeiten können hier auch allgemeine Rechenschwierigkeiten, das heißt jegliches Verbleiben der mathematischen Leistungen unter einer bestimmten Norm (siehe z. B. Schipper, 2010, S. 105), verortet werden. Die Ebenen werden im Folgenden kurz erläutert.

Ebene der psychologischen Erklärungsmodelle von Lernschwierigkeiten

Der Versuch, die Entstehung von Lernschwierigkeiten zu ergründen, wird letztlich immer dadurch beeinflusst, aus welcher Perspektive auf die Lernschwierigkeiten geschaut wird. Folgende psychologischen Ansätze sind in diesem Zusammenhang etabliert (Heimlich, 2009):

- Verhaltenstheoretische (Behavioristische) Erklärungsmodelle
- Kognitionspsychologische Erklärungsmodelle
- Tiefenpsychologische Erklärungsmodelle
- Humanistische Erklärungsmodelle
- Neuropsychologische Erklärungsmodelle

An dieser Stelle wäre es zu umfassend, jeden der Ansätze im Detail zu beschreiben. Letztlich versucht jede Perspektive, den (erschwerten) Prozess des Lernens theoretisch und auf bestimmten Prinzipien beruhend zu erklären. Würde zum Beispiel versucht werden zu erklären, weshalb ein Schüler die Mindestanforderungen im Fach Mathematik nicht erfüllt, könnte dies aus der verhaltenstheoretischen Perspektive durch ein fehlendes Modell, an dem Orientierung für gutes Lernen gefunden werden kann, erklärt werden, aus der kognitionspsychologischen Sichtweise durch eine kognitive Überforderung zum Beispiel im Arbeitsgedächtnis, aus tiefenpsychologischer Sicht durch mangelndes Vertrauen in die Lehrkraft, aus humanistischer Perspektive durch ungünstige Attribuierungen beim Lernen und aus der neuropsychologischen Sichtweise durch Blockaden der Synapsen im Gehirn (Heimlich, 2009). Zur Systematisierung können die verschiedenen Perspektiven in zwei Obergruppen geclustert werden: die pädagogisch-psychologisch und kognitiv-neurowissenschaftlich ausgerichtete Perspektive (Von Aster et al., 2007). Dabei sind Überschneidungen unvermeidbar. Aus dem Blickwinkel dieser beiden übergeordneten Perspektiven ist in Kapitel 2.3 der Erwerb mathematischer Basiskompetenzen in der Grundschulzeit beschrieben.

Ebene der exogenen Bedingungsfaktoren von Lernschwierigkeiten

Auf einer weiteren Ebene stehen exogene (umweltbedingte) Bedingungsfaktoren für Lernschwierigkeiten. Die drei zentralen Bereiche sind Familie, Schule und Umwelt (Ortner & Ortner, 2000). Die soziale Herkunft hat in Deutschland immer noch einen großen Einfluss auf die schulische Entwicklung. Die Forschungslage ist dahingehend sehr einheitlich (Lauth et al., 2014). So lag der Zusammenhang zwischen dem sozioökonomischen Status und der Mathematikkompetenz in Deutschland in der PISA-Studie im Jahr 2012 bei $R^2 = 15{,}6$. Deutschland nahm hier den siebtletzten Platz ein und lag signifikant unter dem OECD-Durchschnitt (Müller & Ehmke, 2013, S. 254). Für das PISA-Jahr 2015 liegen dafür keine Zahlen vor. Ditton und Krüsken (2006) zeigen in ihrer Studie, dass sich die Varianz in den schulischen Leistungen von Drittklässlern (N = 582) mit 19 % durch Merkmale der sozialen Herkunft aufklären lassen. Eine Ergänzung der TIMS-Studie im Jahr 2012 mit 4.768 Schülerinnen und Schülern ergibt, dass die Wahrscheinlichkeit für eine Gymnasialempfehlung am Ende der Grundschulzeit bei sozial weniger begünstigten Familien bei p = 2,5 %, bei sozial begünstigten Familien bei p = 72 % und bei sozial deutlich begünstigten Familien sogar p = 96 % beträgt (Maaz & Nagy, 2010; p = Wahrscheinlichkeit).

Im Zusammenhang mit einem geringen sozialen Status der Familie steht oftmals ein prekärer sozialer Wohnraum und ein eher geringer und wenig fördernder Anregungsgehalt durch die

Eltern oder Peers. Stigmatisierungen und Ausgrenzungen verschärfen dabei oftmals die Problematik (Heimlich, 2009). Doch auch Schule kann Lernen nicht nur fördern, sondern auch erschweren. So haben die zahlreichen nationalen Längsschnittstudien in Kapitel 2.4.2.1 gezeigt, dass sich Schülerinnen und Schüler mit einer vergleichbaren Ausgangslage in ihren schulischen Leistungen am Ende der Klasse vier je nach Anregungsgehalt der Klasse und der Schule (Schulform) unterschiedlich stark entwickeln können (Bos & Gröhlich, 2010; Van Ophuysen & Wendt, 2009; Pekrun et al., 2006).

Neben den Einflüssen auf die Entwicklung auf institutioneller Ebene nehmen auch die Qualität des Unterrichts und die Aktualität von Unterrichts-, Diagnostik und Fördermaßnahmen Einfluss auf die Entwicklung der Schülerinnen und Schüler (Križan & Vossen, 2016). Gerade in der Frage, wie ein inklusives Bildungssystem qualitativ hochwertig gestaltet werden kann, damit Problemlagen frühzeitig erkannt und Lernschwierigkeiten präventiv begegnet werden können, nimmt die Forderung nach Evidenzbasierung einen immer höheren Stellenwert ein (Casale, Hennemann & Grosche, 2015; Hillenbrand, 2015; Križan & Vossen, 2016). Der exogene Einflussfaktor Schule erhält folglich in der Diskussion um Evidenzbasierung im Rahmen inklusiver Bildung eine erhebliche Bedeutsamkeit in der Frage der Prävention von Lernschwierigkeiten.

Ebene der endogenen Bedingungsfaktoren von Lernschwierigkeiten

Eine weitere dritte Ebene ist die der endogenen (anlagebedingten) Bedingungsfaktoren. Diese Ebene kann sehr weit aufgefächert werden und umfasst viele verschiedene Aspekte. In der ICD-10 (Weltgesundheitsorganisation, 2013) und der DSM-V (American Psychiatric Association, 2013) werden die genannten Phänomene „Störungsbilder" genannt. Lauth (2014, S. 67) nennt diese Störungsbilder „unzureichende funktionale Lernvoraussetzungen" und fasst darunter Aspekte wie mangelnde Sprachkompetenz, mangelndes Arbeitsgedächtnis, unzureichende selektive Aufmerksamkeit, mangelnde Kapazität des semantischen Gedächtnisses oder kognitive Entwicklungsverzögerungen. Grünke und Grosche (2014, S. 79) beschreiben für die umfassenden und andauernden Fälle der Lernschwierigkeiten (Lernbehinderung), dass kleinere Areale in den parieto-frontalen Hirnarealen für Lern- und Intelligenzleistungen verwendet werden. Eine grundsätzlich andere Gehirnstruktur wird dabei nicht vermutet. Letztlich sind nahezu alle denkbaren endogenen Bedingungsfaktoren, die das Lernen erschweren könnten, in diese Kategorie einzuordnen. Lauth, Brunstein und Grünke (2014, S. 20 ff.) identifizieren auf dieser Ebene Defizite in vier Bereichen: Ein Bereich umfasst fehlende Basisfertigkeiten wie zum Beispiel Prozesse der Informationsverarbeitung, ein weiterer Bereich beschreibt das fehlende Vorwissen in Form von Wissens- und Begriffssystemen, ein dritter Bereich Defizite in metakognitiven Fertigkeiten und ein vierter Bereich eine mangelnde Lern- und Anstrengungsbereitschaft. In der Ebene der endogenen Bedingungsfaktoren sind auch die Aspekte der unzureichenden Informationsverarbeitung und der geringen Lernaktivität aus dem Bedingungsmodells für Lernstörungen nach Matthes (2006) zu verorten (siehe Kapitel 3.1.1).

Ebene der unmittelbar beobachtbaren und intersubjektiv wahrnehmbaren Lernschwierigkeiten

Auf der letzten vierten Ebene werden die Lernschwierigkeiten letztlich vor allem im Bereich Lesen, Schreiben und Rechnen sichtbar. Aber auch Probleme beim „Lernen lernen" können an

die Oberfläche treten. Die „unmittelbar wahrnehmbaren“ Lernschwierigkeiten können durch die vielschichtigen und miteinander verknüpften Bedingungsfaktoren und Ursachen auf den Ebenen 1 bis 3 erklärt werden.

3.3.2 Rechenstörung

Ebenso wie Heimlich (2009) verschiedene Perspektiven anbietet, aus denen auf eine Lernstörung bzw. Lernschwierigkeit geschaut werden kann, gilt diese Multiperspektivität ebenso für den Bereich der Rechenstörung. Auch hier gibt es verschiedene Blickwinkel, aus denen die Störung erklärt werden kann (neuropsychologisch, kognitionspsychologisch, entwicklungspsychologisch u. a.). Zu dieser großen Anzahl an Blickwinkeln kommt hinzu, dass die Forschung im Bereich der Rechenstörung im Vergleich zu der im Bereich der Lese-Rechtschreibstörung auf eine viel jüngere Forschungstradition zurückschauen muss und viele Fragestellungen noch nicht abschließend beantwortet sind – zum Beispiel die der Ätiologie. Grundsätzlich ist festzuhalten, dass derzeit noch nicht empirisch bestätigt ist, ob die mathematische Kompetenzentwicklung bei Rechenstörungen lediglich verzögert abläuft oder ob sich die Entwicklung tatsächlich qualitativ von der Entwicklung nicht beeinträchtigter Kinder unterscheidet (Fritz & Ricken, 2008, S. 48).

Für die Beschreibung möglicher Ursachen einer Rechenstörung wird hier das Modell nach Jabobs und Petermann (2007) gewählt. Die Autoren vertreten einen eher neuropsychologischen Ansatz, doch verknüpfen sie diesen mit allgemeinen Aspekten der Entwicklung. Die Verknüpfung der „sichtbaren“ und „weniger sichtbaren“ Aspekte mathematischer Kompetenzentwicklung ermöglicht die Ableitung passender und individueller Fördermaßnahmen bei einem erschwerten Kompetenzerwerb.

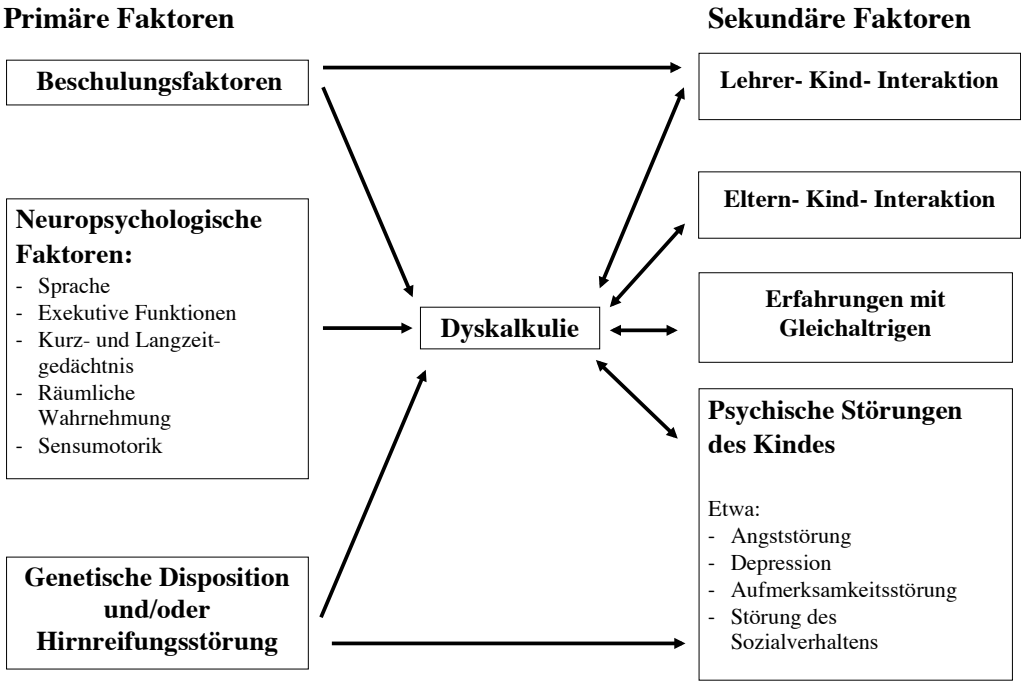


ABBILDUNG 22 MÖGLICHE URSACHEN EINER DYSKALKULIE (JACOBS & PETERMANN, 2007, S. 16)

Die Autoren unterscheiden in ihrem multikausalen Ursachenmodell primäre und sekundäre Faktoren einer Dyskalkulie. Primäre Faktoren (links) werden als ursächlich für die Entstehung

einer Dyskalkulie angesehen. Sekundäre Faktoren (rechts) können in der Folge den Kompetenzerwerb weiter behindern. Bei einer Dyskalkulie ist davon auszugehen, dass sie sich „eher [von] primären Faktoren entwickelt [...] [und das Störungsbild] durch die Wechselwirkung mit sekundären Faktoren zusätzlich negativ beeinflusst werden kann“ (Jacobs & Petermann, 2007, S. 17).

PRIMÄRE FAKTOREN

Zu den primären Faktoren können u. a. genetische Dispositionen, eine Hirnreifungsstörung, neuropsychologische Störungen, aber auch Effekte der Beschulung, die sich auch zusammen auf die Rechenleistung auswirken können, gezählt werden. Lorenz (2014, S. 45 f.) fasst unter den kognitiven Faktoren eine mangelnde Vorstellungsfähigkeit, ein unzureichendes Sprachverständnis und Gedächtnis sowie eine mangelnde Rechts-Links-Unterscheidung. Die Annahme, „dass es sich um eine genetisch determinierte, neurobiologische Störung handelt“ (Landerl & Kaufmannn, 2013, S. 102), beruht auf Erkenntnissen der Neurobiologie. Dass zum Beispiel genetische Dispositionen einen Einfluss darauf haben können, ob ein Kind eine Dyskalkulie entwickelt oder nicht, zeigen verschiedene Zwillings- und Familienstudien. So führen Alarcón und Kollegen (1997) eine der ersten überhaupt vorliegenden Zwillingsstudien durch, in der jeweils ein Kind dyskalkulisch ist ($n_{\text{eineiig}} = 40$; $n_{\text{zweieiig}} = 23$). Sie finden heraus, dass die Wahrscheinlichkeit einer „zweiten“ vorliegenden Dyskalkulie im Vergleich zu allgemeinen Prävalenzraten bei den eineiigen Pärchen um das zwölffache und bei den zweieiigen Pärchen um das achtfache erhöht ist. Mittlerweile gibt es zahlreiche Untersuchungen, die mittels funktioneller Magnetresonanztomographie (fmRT) anatomisch-funktionell nach Ursachen für Rechenstörungen im Gehirn suchen. Viele Studien geben Hinweise darauf, dass Auffälligkeiten im intraparietalen Sulcus (IPS) vorliegen, der im Parietallappen verortet ist. Eine zentrale Studie stammt dabei von Molko und Kollegen (2003) unter Mitarbeit von Stanislas Dehaene (siehe Abbildung 8 Triple-Code-Modell nach Dehaene (1992)). Die Forschergruppe untersucht 14 Probanden im Alter von ungefähr 24 Jahren mit und ohne dem Turner-Syndrom. Da mit dem Turner-Syndrom typischerweise bestimmte Defizite in der Entwicklung der Rechenleistungen einhergehen, soll die Frage beantwortet werden, ob neurologische Unterschiede in den Hirnregionen vorliegen. Die Ergebnisse zeigen, dass bei den Probanden mit dem Turner-Syndrom beim Lösen von Rechenaufgaben eine geringere Aktivierung im intraparietalen Sulcus erkennbar ist, der zudem etwas kürzer und weniger tiefgehend ausgebildet ist. Diese Region wird beim Rechnen jedoch ebenso angesteuert. Dies deckt sich mit weiteren Ergebnissen der Forschung, die zeigen, dass bei Lernstörungen Gehirnstrukturen grundsätzlich nicht anders angeordnet sind, sich aber auf Mikroebene qualitativ unterscheiden können (Kucian & Von Aster, 2013; Schuchardt & Mähler, 2010; Swanson, Cooney & McNamara, 2008).

Ergebnisse weiterer Studien zu Gehirnaktivitäten im Parietallappen bei einer vorliegenden Dyskalkulie (Isaacs, Edmonds, Lucas & Gadian, 2001; Kucian et al., 2006) decken sich im Gros mit den bisherigen Ausführungen.

Ein weiterer Aspekt bei der Beschreibung der primären Faktoren ist die Beschulung. Dazu gehört sicherlich die Wahl der zu dem Kind passenden Schulform, aber auch die Umsetzung guten Mathematikunterrichts. Als eher ungünstige Unterrichtsbedingung fasst Gaidoschik (2004) auf, wenn

- Wahrnehmungsübungen eine höhere Relevanz haben als Rechenübungen,

- Üben bedeutet, lediglich eine große Menge an Aufgaben zu rechnen,
- gedankliche Auseinandersetzungen mit dem Übematerial zu kurz kommen,
- der Fokus auf dem Auswendiglernen liegt und
- unstrukturiertes Anschauungsmaterial verwendet wird.

SEKUNDÄRE FAKTOREN

Neben den primären Faktoren identifizieren Jacobs und Petermann (2007) sekundäre Bedingungsfaktoren. Die Autoren zählen dazu u. a. eine schwierige Lehrer-Kind-Interaktion bzw. Eltern-Kind-Interaktion, negative Erfahrungen mit Gleichaltrigen und auch psychische Störungen des Kindes wie zum Beispiel Angststörungen oder Depressionen. Die primären Faktoren sind vergleichbar mit den endogenen und exogenen Bedingungsfaktoren von Lernschwierigkeiten bei Heimlich (2009). Fritz und Ricken (2008, S. 14) bemängeln, dass es bezüglich des Einflusses motivationaler und emotionaler Faktoren sowie Unterrichtsbedingungen auf die Entwicklung von Rechenschwierigkeiten nur wenige Studien gibt. Oftmals sind der Kern der Problematik die verzögerten Kompetenzen im Rechnen, aus denen dann Lernprobleme, Motivationsprobleme und ein negatives Selbstkonzept folgen können. Die erlebten Misserfolge wirken sich ungünstig auf das Lernverhalten aus, das dann oftmals mit Druck und Sanktionen von Seiten der Familie oder der Schule einhergeht.

3.4 Diagnostik

Aufgezeigt ist bisher, in welcher Form Schwierigkeiten in der Entwicklung und im Ererb mathematischer Kompetenzen vorliegen können. Es stellt sich die Frage, wie verschiedene Ausprägungen von Rechenschwierigkeiten bzw. eine Rechenstörung erkannt und im Rahmen einer Differenzialdiagnostik voneinander abgegrenzt werden können. Erschwert wird eine differenzierte Diagnostik grundsätzlich dadurch, dass Kinder mit einer Rechenschwäche oder Rechenstörung formal und strukturell ähnliche Fehlermuster aufweisen können und oftmals erst deren *Häufigkeit* und die *Manifestation* bzw. die *Resistenz* in Förderphasen Indikatoren für die Ausprägung der Rechenschwierigkeit sind (Fritz & Ricken, 2008; Tröster, 2008). Zudem besteht oftmals die Vorstellung, dass sich lediglich bei der Diagnose „Rechenstörung“ eine Förderphase anschließt. Aktuell besteht jedoch vermehrt die Forderung, von dieser Kategorisierung, die jedoch eng mit einer Ressourcenzuweisung zusammenhängt, abzuweichen. Dies stellt eine pädagogische Herausforderung bezüglich der Förderung dar und zeigt, dass letztlich alle Schülerinnen und Schüler mit Schwierigkeiten im Rechnen jedweder Ausprägung einer Förderung bedürfen (Lambert, 2015, S. 66; Tröster, 2008).

> „Der mehr pädagogische Zugang zum Problem Rechenschwäche fordert einen Verzicht auf die eindeutige Definition der Dyskalkulie. Rechenschwierigkeiten sollen differenziert analysiert werden, um der Heterogenität der entsprechenden Lernschwäche gerecht zu werden und vor allem den pädagogischen Förderbedarf mehr in den Vordergrund zu stellen.“ (Born & Oehler, 2013, S. 4)

In diesem Kapitel wird versucht, eine dimensionale Perspektive der Diagnostik zu eröffnen und weniger den kategorialen Ansatz (Rechenstörung liegt vor vs. Rechenstörung liegt nicht vor) zu vertreten.

In der Literatur liegen diverse Vorschläge zur Diagnostik von Rechenschwierigkeiten vor. Die Ansätze unterscheiden sich insgesamt bezüglich ihrer diagnostischen Tiefe (Abstufung der Schweregrade) und Weite (wird zum Beispiel die Förderung und/oder eine Verlaufsdiagnostik in den diagnostischen Prozess einbezogen?). So schlagen Fritz und Ricken (2008) in ihrem zyklischen diagnostischen Prozess die Bearbeitung von vier Fragen vor: 1. Liegt eine Beeinträchtigung in der Entwicklung vor? 2. In welchen Bereichen zeigen sich die Schwierigkeiten? 3. Wie verändern sich die Kompetenzen unter den Bedingungen der Förderung? 4. Welche Kompetenzen hat das Kind während der Förderung entwickelt? Jacobs und Petermann (2012) stellen mit einem Fließschema einen weitaus differenzierteren Weg, bestehend aus vier Ebenen, vor: 1. Anamnese, 2. Basisdiagnostik, 3. Erfassung von Funktionsbeeinträchtigungen, 4. Befundgespräch (siehe Abbildung 24). Ein weiterer Vorschlag stammt von Lorenz (2014), der eine Diagnostik in sechs Schritten vorschlägt: 1. Durchführung eines Rechentests, 2. Fehleranalyse, 3. Überprüfung der Gedächtniskapazität, 4. Anamnese, 5. Beobachtungen zur motorischen Entwicklung, 6. Psychometrische Bestimmung der kognitiven Leistungsfähigkeit. Letztlich ist auch der dreischrittige diagnostische Prozess nach Bundschuh (2014) mit der Phase der Vorinformation, der psychologisch/pädagogisch/medizinischen Informationsphase und dem Förderungsprozess eine Möglichkeit zur Diagnostik einer Rechenstörung.

Im Kontext dieses Kapitels wird eine Herangehensweise verfolgt, die die bestehenden Ansätze vereint, um sowohl Tiefe als auch Weite des diagnostischen Prozesses zu ermöglichen. Dies bedeutet, dass grundsätzlich auch diejenigen Schülerinnen und Schüler mit Rechenschwierigkeiten, die nicht den Umfang einer Rechenstörung aufweisen, in den diagnostischen Phasen berücksichtigt werden.

3.4.1 Hinweise und Symptome in Vorschule und Schule

Mögliche Ursachen von Rechenschwierigkeiten oder einer Rechenstörung sind im vorherigen Kapitel 3.3 in ihrer Komplexität und Vernetzung dargestellt. In diesem Kapitel werden spezifischen Hürden im Erwerb mathematischer Kompetenzen aufgeführt, an denen sich Symptome für Rechenschwierigkeiten und einer Rechenstörung zeigen können. Einen umfangreichen Überblick über Früherkennungshinweise in Vorschule und Schule bietet Barth (2006). Im Regelfall werden Schwierigkeiten im Erwerb mathematischer Kompetenzen erst zu Beginn des Schuleintritts deutlich, da die Kinder erst dann mit mathematischen Anforderungen konfrontiert werden (Fritz & Ricken, 2008; Tröster, 2008). Im mathematischen Anfangsunterricht werden jedoch nicht immer direkt kritische Phasen erkannt, da eine Kompensation von Schwierigkeiten durch Auswendiglernen oder Abzählstrategien erfolgen kann (Schneider et al., 2013). Wie bereits erläutert, können sich Rechenschwierigkeiten bei einer zu späten Erfassung bereits manifestiert haben, sodass eine frühe Erkennung von großer Bedeutung ist. Bereits im Vorschulalter können spezifische Aspekte Hinweise auf Rechenschwierigkeiten geben (siehe Tabelle 11).

TABELLE 11 ANZEICHEN EINER RECHENSCHWÄCHE IN DER VORSCHULZEIT (VGL. TRÖSTER, 2008, S. 353)

Bereich	Hürden und Symptome
Zeitliche Orientierung	• Schwierigkeit im Verständnis von Zeitangaben • Verwechslung von vorher-nachher oder früher-später
Räumliche Beziehungen	• Verwechslung von rechts-links oder oben-unten oder hinten-vorn
Vergleichen und Sortieren von Objekten	• Schwierigkeiten beim Sortieren von Objekten nach räumlichen Kategorien (ist größer, kleiner; liegt vor, über, hinter, neben) • Verwechslung von Angaben zu quantitativen Beziehungen (kürzer-länger, größer-kleiner, breiter-schmaler, höher-niedriger)
Zahlverständnis	• Schwierigkeiten, kleinere Mengen von Objekten als Ganzes auf einen Blick zu erfassen
Gedächtnis-leistungen	• Schwierigkeiten darin, bildliche Darstellungen im Gedächtnis zu behalten oder abgeschriebene Aufgaben auf der Heftseite wiederzufinden
Vermeiden bestimmter Spielaktivitäten	• Vermeiden von Spielen, die Gedächtnisleistungen oder visuell-räumliche Verarbeitung erfordern, wie etwa Memory, Puzzle oder Spielen mit Legosteinen

Im Regelfall werden rechenschwache Kinder jedoch erst in der Grundschule erkannt. Auch in dieser Phase lassen sich spezifische Hürden und Symptome ausmachen, die auf eine Rechenschwäche hinweisen. In der Literatur werden oftmals Schwerpunkte auf *spezifische* Entwicklungsstufen (zum Beispiel Mengenwissen, T-T-G-Konzept o. ä.) gesetzt, sodass die Hürden und Symptome selten umfassend dargestellt werden. Für den schulischen Erwerb mathematischer Kompetenzen ist bereits das Kompetenzstufenmodell nach Reiss (2004) bzw. Reiss und Winkelmann (2008) vorgestellt. Dieses Kompetenzstufenmodell wird hier hinzugezogen und literaturbasiert an den spezifischen Kompetenzstufen für Klasse eins bis vier (für Klasse fünf und sechs siehe Zimmermann (2014)) besondere Hürden für rechenschwache Schülerinnen und Schüler zugeordnet. Die konkreten Literaturhinweise finden sich als Verweise unter der Tabelle.

TABELLE 12 KOMPETENZSTUFEN NACH REISS (2004) BZW. NACH REISS UND WINKELMANN (2008) MIT HÜRDEN UND SYMPTOMEN BEI RECHENSCHWIERIGKEITEN/-STÖRUNGEN

Kompetenzstufe		Hürden und Symptome
I	Numerisches und begriffliches Grundlagenwissen (Routineprozeduren)	• Fehlendes Mengen- und Größenverständnis (b, f, g) • Unzureichender Aufbau des Zahlensystems und der Dezimalstruktur (a, f, g) • Einsatz zählender Rechenstrategien (a, f, g) • Umständliche Zählstrategien (b, f) • Fehlendes Verständnis des Teile-Ganze-Prinzips (d)
II	Einfache Anwendung von Grundlagenwissen	• Zahlen und Rechenoperationen erscheinen sinnentleert (a) • Aufgabenstellungen werden nicht wiedererkannt und Zwischenergebnisse vergessen (b, g) • Übersetzungsfehler beim Lesen und Schreiben (b) • Fehler beim Einsatz der Rechengesetze (c) • Geringe Automatisierung (d, f) • Rechenzeichen werden vertauscht und ein Wechsel nicht beachtet (g) • Schwierigkeiten beim Benennen von Zeichen, Figuren & Formen (g)
III	Sicheres Rechnen in curricularem Umfang und einfaches Modellieren	• Zwischen Darstellungsformen (Bild, Text, Symbol) kann nicht gewechselt werden (a) • In Textaufgaben werden falsche Rechenoperationen angewendet (g) • Rechenfehler (b)
IV	Beherrschung der Grundrechenarten unter Nutzung der Dezimalstruktur und begriffliche Modellierung	• Fehlende Erkenntnis von Zahlbeziehungen und Strukturen (c) • geringes Faktenwissen (d, e, f) • erschwerte Reduzierung der dargestellten Situation auf quantitative Größen (d) • Fehlende Einsicht in das dezimale Stellenwertsystem (f)
V	Anspruchsvolles Problemlösen im mathematischen Kontext	• Unzureichendes Text- und Begriffsverständnis (c) • Erschwerte Entschlüsselung der Oberflächenstruktur (d, g) • Unzureichende Bildung einer Situationsvorstellung (d, f) • Verbleib beim zählenden Rechnen (d) • Schwierigkeiten beim Lesen der Uhr, beim Abschätzen von Längen und Maßen, beim Rechnen mit Geld (g)

Anmerkung: (a) Lorenz (2014); (b) Jacobs & Petermann (2007), (c) Zimmermann (2014), (d) Fritz und Ricken (2008), (e) Busch, Oranu, Schmidt & Grube (2013), (f) Walter (2016), (g) Lambert (2015)

Forschungsergebnisse zeigen, dass sich in der Schulzeit Probleme bei rechen*schwachen* Schülerinnen und Schülern auch schon in den frühen grundlegenden Vorläuferfertigkeiten zeigen: Gaupp und Kollegen (2004) bestätigen für rechenschwache Dritt- und Viertklässler Defizite in den Bereichen Vergleichen von Zahlen, Lokalisieren von Zahlen auf einem Zahlenstrahl, Schätzen von Objektmengen sowie in komplexen Zählaufgaben. Vergleichbar schnitten die Schülerinnen und Schüler im Vergleich zu der Kontrollgruppe beim einfachen Zählen ab.

3.4.2 Anamnese

Weisen erste Symptome auf das Vorliegen einer Störung im Kompetenzerwerb hin, tritt für die diagnostizierende Fachkraft (zum Beispiel Sonderpädagoge/in, Psychologe/in) nach Bundschuh (2014, S. 138 f.) die Phase der Vorinformation ein. Dazu gehört grundsätzlich, dass Gespräche mit den Eltern, den beteiligten Akteuren aus dem (vor-)schulischen Bereich, Ärzten und Lehrkräften und weiteren relevanten Bezugspersonen geführt werden. Ziel ist es, erste Informationen zum bisherigen Entwicklungsverlauf des Kindes zu erhalten und eine mögliche Hypothese für die nächste Phase im diagnostischen Prozess zu entwickeln. Nach Jacobs und Petermann (2012, S. 95 f.) soll in diesem Schritt ebenso eine qualitative Beschreibung der derzeitigen Rechenleistungen erfolgen. Die Eltern sollen dabei die spezifischen Fehler beim Rechnen ihres Kindes beschreiben und dabei, sofern das Kind bereits die Schule besucht, einen besonderen Fokus auf die Grundrechenarten legen. Zur Unterstützung haben die Autoren für Eltern und Lehrkräfte eine Checkliste mit möglichen Fehlern entwickelt. Zudem halten sie für das Elterngespräch einen Gesprächsleitfaden mit 33 Fragen bereit. Beispielfragen sind u. a. „Zu welchen Problemen kommt es beim Rechnen?“, „Wie sind die Leistungen beim Lesen und in der Rechtschreibung?“, „Wie sind die sonstigen schulischen Leistungen?“, „Meidet Ihr Kind Malen, Puzzlen, Basteln und/oder Legospielen?“, „Wie ist das Verhältnis Ihres Kindes zum/zur Mathematiklehrer/in?“ oder „Gibt es in Ihrer Familie noch jemanden, der Rechenprobleme oder andere Schulschwierigkeiten hatte?“.

3.4.3 Analyse der Mathematikleistungen

Auf die Phase der Vorinformation folgt die psychologische/pädagogische/medizinische Informationsphase (Bundschuh, 2014). Diese sollte stets hypothesengeleitet erfolgen, was impliziert, dass die diagnostizierende Person (zum Beispiel ein/e Sonderpädagoge/in) auf Grundlage der ersten Hinweise und der Anamnese Vermutungen über Ursachen der Rechenschwierigkeiten und die dafür einzusetzenden diagnostischen Verfahren hat. Für die Diagnostik kommen dabei sowohl formelle, sog. standardisierte Verfahren, als auch informelle Verfahren in Frage.

3.4.3.1 Formelle Verfahren

Formelle Verfahren als standardisierte Testverfahren zeichnen sich dadurch aus, dass „der Test den sogenannten ‚Gütekriterien’ entspricht und das Verhalten des Testleiters, die Untersuchungssituation und die Kriterien der Auswertung genau festgelegt, ‚standardisiert’ sind. Der Sonderpädagoge oder der Psychologe muss über die Standardisierung eines Verfahrens Bescheid wissen, um den Grad der Wahrscheinlichkeit einer subjektiven Verfälschung der Ergebnisse zumindest einschätzen zu können“ (Bundschuh, 2014, S. 76).

Die Diagnose einer Rechenstörung macht den Einsatz standardisierter Verfahren notwendig, da die Diagnose formal an spezifische Kriterien gekoppelt ist, die sich von den Vorgaben der ICD-10 (Weltgesundheitsorganisation, 2013) ableiten:

- Die Testleistungen in einem standardisierten Mathematiktest (siehe Tabelle 13) müssen unterdurchschnittlich sein (mindestens eine Standardabweichung unter dem Vergleich mit der entsprechenden Alters- bzw. Klassenstufe; oftmals Prozentrang < 10 (Jacobs & Petermann, 2012, S. 96).
- Es muss eine bedeutsame Diskrepanz zwischen der Mathematikleistung und Intelligenz bestehen (ca. 1 bis 1,5 Standardabweichungen oder 12 T-Wert-Punkte) (Jacobs & Petermann, 2012; Mähler, 2014).
- Die Lese- und Rechtschreibleistungen sollten im Normalbereich liegen (Lorenz, 2014, S. 44).

Weitere zu berücksichtigende Aspekte sind:

- Die schulischen Leistungen im Fach Mathematik sind unzureichend (Jacobs & Petermann, 2003).
- Die schulischen Leistungsstörungen sollten vor Eintritt in die sechste Klasse auftauchen (Jacobs & Petermann, 2003).
- Ausschluss von Hörschädigung und Krankheit in der Vergangenheit (Ricken, 2014, S. 69).

Fischbach und Kollegen (2013, S. 66) bezeichnen die Voraussetzung „unterdurchschnittliche Mathematikleistung bei einer durchschnittlichen Intelligenz“ als „doppeltes Diskrepanzkriterium“. Diese Bezeichnung erfolgt in Abgrenzung zu allgemein unterdurchschnittlichen Rechenleistungen, die unterhalb einer bestimmten Norm liegen (sog. „einfaches Diskrepanzkriterium“). So etabliert das doppelte Diskrepanzkriterium derzeit ist, so groß ist auch die Kritik daran (Ehlert et al., 2012). Kritisiert wird u. a.,

- dass essentielle Forschungsergebnisse fehlen, die die Diskrimination des Kriteriums Intelligenz bestätigen (Gold, 2015; Ricken, 2014; Schipper, 2010, S. 107),
- dass Diskrepanzen bezüglich der Güte zwischen standardisierten Intelligenzverfahren bestehen (Bundschuh, 2014),
- dass standardisierte Rechentests oftmals sogar mit Intelligenztests korrelieren, da spezifische Subtests etwas Vergleichbares testen (Bundschuh, 2014) und
- dass die empirische Grundlage zur Komorbidität mit Lese-Rechtschreibleistungen unzureichend ist (Grube, 2008; Landerl & Kaufmannn, 2013).

Uneinigkeit besteht auch darüber, wie hoch der IQ zur Diagnose Rechenstörung tatsächlich ausfallen sollte (verschiedentlich > 70 oder auch > 85) (Fischbach et al., 2013; Schipper, 2010, S. 107). In einigen aktuellen Studien zur Prävalenz und Förderung von Rechenstörungen – wie auch im DSM-V (American Psychiatric Association, 2013) – wird schon auf die Berücksichtigung des Diskrepanzkriteriums verzichtet (siehe zum Beispiel Von Aster et al., 2007; Ehlert et al., 2012; Gold, 2015; Ise, Dolle, Pixner & Schulte-Körne, 2012).

Trotz Kritik muss die Diagnostik der Rechenfähigkeiten über standardisierte Instrumente erfolgen, da die Zuweisung von Ressourcen (sog. „Eingliederungshilfe für seelisch behinderte Kinder und Jugendliche“) nach §35 Sozialgesetzbuch VIII Absatz 1a nur auf diagnostischer Grundlage der ICD erfolgt.

Seit der Jahrtausendwende sind in Deutschland zahlreiche standardisierte Verfahren zur Diagnostik einer Rechenstörung entwickelt worden. Einen umfassenden Einblick ermöglicht die Testzentrale des Hogrefe-Verlags, die Auflistungen und Erläuterungen bei Tröster (2008,

S. 347) oder bei Jacobs und Petermann (2012, S. 185 ff.). Folgende Abbildung 23 gibt einen Überblick über standardisierte Verfahren, die im Sekundarstufen-I-Bereich Anwendung finden können. In der Abbildung 23 werden die Verfahren bezüglich ihres Grades an Fokussierung und Individualisierung aufsteigend abgebildet. Dabei ist die Einordnung bestimmter Verfahren auf *eine spezifische* Ebene zu diskutieren, die Abbildung 23 versucht deswegen lediglich einer Annäherung.

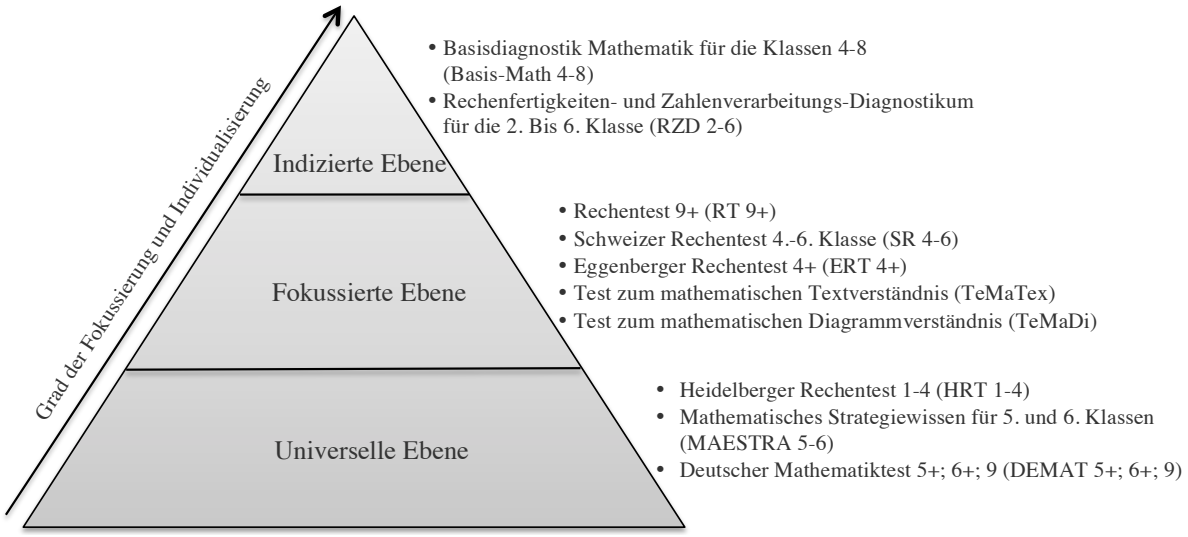


ABBILDUNG 23 STANDARDISIERTE TESTVERFAHREN IM BEREICH MATHEMATIK FÜR DIE SEKUNDARSTUFE I

Die angegebenen Verfahren auf der universellen Ebene können als Gruppentestungen durchgeführt werden und geben einen eher allgemeinen und breiten Überblick über den mathematischen Kompetenzstand. Die DEMAT-Reihe ist zum Beispiel curricular ausgerichtet und testet allgemein die zu erreichenden Kompetenzen in den angegebenen Klassenstufen. „[...] Eine Auswertung individueller Stärken und Schwächen anhand der Aufgabentypen ist aber nicht mit ausreichender Sicherheit gegeben“ (Fritz & Ricken, 2008, S. 70). Aber auch Testverfahren ohne expliziten Lehrplanbezug, wie der HRT 1–4, der auch in Klasse 5 Anwendung finden kann, sind auf dieser Ebene zu verorten, da sie ein breites Spektrum an Untertests bieten. Die Testverfahren auf der fokussierten Ebene sind z. T. als Gruppentests, z. T. aber auch in Form von Einzeltestungen durchführbar. Die Testverfahren auf dieser Ebene differenzieren speziell im unteren Leistungsbereich, sind inhaltlich spezialisierter und geben Hinweise für systematische Förderungsmaßnahmen. Die Testverfahren auf der dritten Ebene sind ausschließlich als Einzeltestungen durchzuführen. Sie dienen speziell der Diagnostik einer Rechenstörung. Bei Auffälligkeiten schließt sich die Testung auf der universellen Ebene im Regelfall eine Diagnostik auf der fokussierten oder indizierten Ebene an, um einen Verdacht auf das Vorliegen einer Rechenstörung bestätigen oder verwerfen zu können (Born & Oehler, 2013, S. 6).

Bei den Verfahren kann zudem unterschieden werden, ob ein entwicklungspsychologisches (zum Beispiel beim HRT 1–4) oder neuropsychologisches (zum Beispiel RZD 2–6) Paradigma mathematischer Kompetenzentwicklung zugrunde liegt.

Folgende Tabelle 13 gibt einen detaillierten Überblick über die angegebenen Testverfahren.

TABELLE 13 DETAILANGABEN ZU DEUTSCHSPRACHIGEN STANDARDISIERTEN TESTVERFAHREN IM BEREICH MATHEMATIK FÜR DIE SEKUNDARSTUFE I

Name, Autor	Kürzel	Einsatzbereich	Durchführung	Inhaltsbereiche	Güte nachgewiesen für...*
Deutscher Mathematiktest 5+ (Götz, Lingel & Schneider, 2013)	DEMAT 5+	Ende der 5. Klassenstufe bis zum Ende des ersten Halbjahres der 6. Klassenstufe	Gruppen- oder Einzeltestung 35 Minuten	Arithmetik (14 Items) Geometrie (4 Items) Sachrechnen (15 Items)	☒ Objektiviät ☒ Reliabiliät ☒ Validität
Mathematisches Strategiewissen für 5. und 6. Klassen (Hasselhorn, Schneider & Trautwein, 2014)	MAESTRA 5–6 +	Vom 2. Halbjahr der 5. bis zum 1. Halbjahr der 7. Klassenstufe	Gruppen- oder Einzeltestung 10 – 15 Minuten	Fünf mathematikspezifische Lern- und Leistungsszenarien mit fünf bis sechs Handlungsalternativen	☒ Objektiviät ☒ Reliabiliät ☒ Validität
Heidelberger Rechentest 1–4 (Haffner, Baro, Parzer & Resch, 2005)	HRT 1–4	Ab Ende der 1. Klassenstufe bis Anfang der 5. Klassenstufe	Gruppen- oder Einzeltestung 50-60 Minuten	Rechenoperationen (6 Subtests) Numerisch-logische und räumlich-visuelle Fähigkeiten (5 Subtests)	☒ Objektiviät ☒ Reliabiliät ☒ Validität
Test zum mathematischen Diagrammverständnis (Stecken & Stein, 2015)	TeMaDi	Von der 3. bis zur 5. Klassenstufe	Gruppen- oder Einzeltestung 45 Minuten	25 Aufgaben zum Diagrammverständnis	☒ Objektiviät ☒ Reliabiliät ☒ Validität
Test zum mathematischen Textverständnis (Jordan & Stein, 2011)	TeMaTex	Klassenstufen 9 bis 10 und alle berufsbildenden Schulen	Gruppen- oder Einzeltestung 45 Minuten	Sieben Aufgabentexte mit kontinuierlichen und nicht kontinuierlichen Elementen	☒ Objektiviät ☒ Reliabiliät ☒ Validität

Name, Autor	Kürzel	Einsatzbereich	Durchführung	Inhaltsbereiche	Güte nachgewiesen für...*
Eggenberger Rechentest 4+ (Schaupp, Holzer & Lenart, 2010)	ERT 4+	Ende der 4. Klassenstufe bis zum Halbjahr der 5. Klassenstufe (eher unteres Leistungsspektrum)	Gruppen- oder Einzeltestung 3 x 45 Minuten	Mathematische Ordnungsstrukturen Algebraische Strukturen Größenbeziehungen Angewandte Mathematik	☒ Objektiviät ☒ Reliabiliät ☒ Validität
Schweizer Rechentest 4.–6. Klasse (Lobeck, Frei & Blöchlinger, 1990)	SR 4–6	4. bis 6. Klassenstufe	Gruppen- oder Einzeltestung 30–100 Minuten	Zahlen Operationen Symbole Kopfrechnen Rechnen nach Stellenwert Größen und Messen Mengen und Maße Textaufgaben Geometrische Darstellungen	☒ Objektiviät ☒ Reliabiliät ☒ Validität
Rechentest 9+ (Bremm & Kühn, 1992)	RT 9+	9.–10. Klasse Haupt- und und Realschule, Beginn Berufsschule, Ausbildung oder Beschäftigung	Gruppen- oder Einzeltestung	Bruchrechnen Prozentrechnen Zinsrechnen Gleichungen Potenzen und Wurzeln Rechnen mit Größen	☒ Objektiviät ☒ Reliabiliät ☒ Validität

Name, Autor	Kürzel	Einsatzbereich	Durchführung	Inhaltsbereiche	Güte nachgewiesen für...*
Rechenfertigkeiten- und Zahlen-verarbeitungs-Diagnostikum für die 2. bis 6. Klasse (Jacobs & Petermann, 2014)	RZD 2–6	Ab dem Ende der 2. Klassenstufe bis zur Mitte der 6. (eher unteres Leistungsspektrum)	Individualtest 29–44 Minuten	Zahlen lesen & schreiben Abzählen Zahlenstrahl Mengenschätzen Größenvergleiche Kopfrechnen Schriftliches Rechnen Flexibles Anwenden Regelverständnis Zählrahmen Textaufgaben	☒ Objektiviät ☒ Reliabiliät ☒ Validität
Basisdiagnostik Mathematik für die Klassen 4–8 (Moser Opitz et al., 2010)	Basis-Math 4–8	Einsetzbar ab dem letzten Quartal der 4. Klassenstufe bis zum Ende der 8. Klassenstufe (eher unteres Leistungsspektrum)	Individualtest 20–45 Minuten	Grundoperationen Rechenwege Dezimales Verständnis Zählkompetenz Operationsverständnis Mathematisierungsfähigkeit	☒ Objektiviät ☒ Reliabiliät ☒ Validität

*Anmerkung: Für Detailangaben zu den Gütekriterien wird auf die Manuale der Testverfahren verwiesen.

Neben den hier vorgestellten standardisierten Testverfahren gibt es u. a. Schulleistungstests, die den Aspekt der mathematischen Kompetenzen berücksichtigen, diesen jedoch nicht ins Zentrum stellen. Dazu gehören zum Beispiel der Hamburger Schulleistungstest 4/5 (als Erweiterung des Hamburger Kombinierten Schulleistungstests für vierte und fünfte Klassen (KS HAM 4/5) (Mietzel, Willenberg, Poerschke & Peek, 2001) oder auch die Diagnostischen Einschätzskalen (Barth, 2006) zum Schuleintritt. Diese Verfahren werden hier nicht weiter vorgestellt.

Auch bezüglich der Erhebung der Intelligenz und der Lese-Rechtschreibleistung im Rahmen der Differenzialdiagnostik sollten standardisierte Verfahren zum Einsatz kommen. Simon und Grünke (2010, S. 36) empfehlen für den Bereich der Intelligenz das Adaptive Intelligenzdiagnostikum 3 (AID-3) (Kubinger & Holocher-Ertl, 2014), den Hamburg-Wechsler-Intelligenztest für Kinder (HAWIK-IV) (Petermann & Petermann, 2010) oder den Nonverbalen Intelligenztest (SON-R 5 ½ - 17) (Snijders, Tellegen & Laros, 2005). Für den Lese-Rechtschreib-Bereich bietet sich zum Beispiel die Diagnostischen Rechtschreibtest (DRT 1; 2; 3; 4 oder 5) (Müller, 2003) an.

3.4.3.2 Informelle Verfahren

Neben standardisierten Testverfahren sollten bei der Diagnostik von Rechenschwierigkeiten auch informelle Verfahren zum Einsatz kommen. Dabei wird der Begriff „informelle Verfahren“ in der Literatur nicht einheitlich verwendet. Hesse und Latzko (2011, S. 245) fassen darunter zum Beispiel Tests, die von Lehrkräften selbst erstellt werden. Bundschuh (2014, S. 138) fasst darunter alle Verfahren, die während der gesamten Informationsphase zum Einsatz kommen, wie zum Beispiel auch Beobachtungen. In diesem Kapitel werden unter informellen Verfahren solche verstanden, die sich von standardisierten Verfahren dadurch abgrenzen, dass die Durchführung und Auswertung nicht genau festgelegt ist und diese Verfahren auch keiner Überprüfung im Hinblick auf Gütekriterien unterzogen sind. Dabei wird die Durchführung einer Fehleranalyse als eine Möglichkeit informeller Diagnostik in den Fokus gestellt. Eine Fehleranalyse wird gerade auch den Schülerinnen und Schülern gerecht, bei denen sich zeigt, dass keine explizite Rechenstörung vorliegt, sondern allgemeine Rechenschwierigkeiten festgestellt werden können.

Folgende Abbildung 24 zeigt den diagnostischen Prozess bei der Überprüfung einer Rechenstörung. Das Pfaddiagramm beschreibt den Weg über die Erhebung des Intelligenzquotienten unter Berücksichtigung der Leistungen im Rechnen sowie Lesen und Rechtschreiben. Ebenso wird der Aspekt der emotionalen Entwicklung (in Abbildung 24 rechts) einbezogen. Aufgrund der aktuellen Umstrittenheit des Diskrepanzkriteriums (siehe Kapitel 3.4.3.1) wird dieses Kriterium bzw. dieser Schritt farblich hinterlegt.

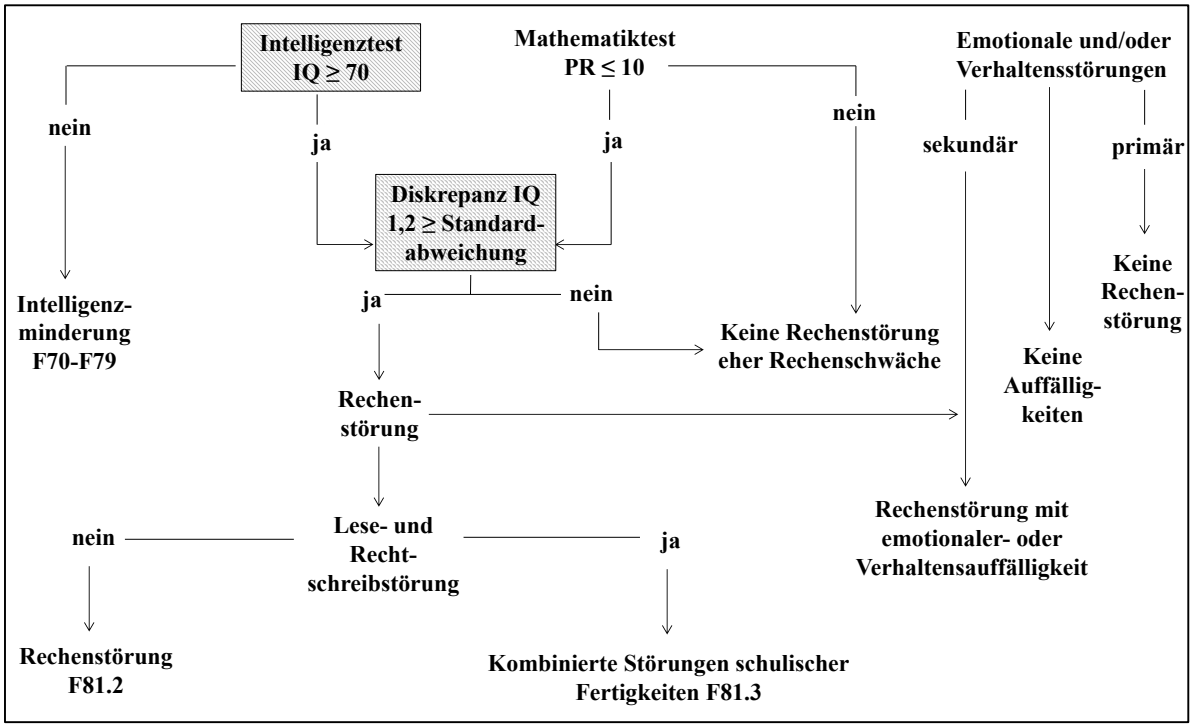


ABBILDUNG 24 DIFFERENZIALDIAGNOSTIK RECHENSTÖRUNG VS. RECHENSCHWIERIGKEIT (JACOBS & PETERMANN, 2012, S. 96; MODIFIZIERT DURCH D. VERF.)

Da vor Einleitung der Differenzialdiagnostik offenbar schon erste Hinweise auf Schwierigkeiten beim Rechnen hingewiesen haben (siehe Kapitel 3.4.1), geht ein Ausschluss einer Rechenstörung nicht damit einher, dass keine Probleme im Rechnen vorliegen sollten. Gerade für diese Schülerinnen und Schüler, aber auch für diejenigen mit einer diagnostizierten Rechenstörung, sollte sich eine differenzierte Fehleranalyse anschließen. „[...] Die mehr medizinische Orientierung der Definition einer Rechenstörung [liefert] keine Förderhinweise“ (Born & Oehler, 2013, S. 5). Eine Fehleranalyse ermöglicht es, den Leistungsstand der Schülerinnen bzw. Schüler explizit beschreiben zu können (Jacobs & Petermann, 2012, S. 108), um eine auf die Rechenprobleme abgestimmte Therapie- bzw. Förderplanung entwickeln und durchführen zu können (Simon & Grünke, 2010, S. 36).

FEHLERANALYSE

Grundsätzlich gilt es in Fragen der Fehleranalyse, auch die Fehlerkultur in den Blick zu nehmen. Wittmann (2007a) fordert für den Mathematikunterricht eine größere Offenheit und Toleranz des Fehlermachens. Die Fehleranalyse ist dabei eine wichtige Voraussetzung für eine positive Fehlerkultur, die letztlich auch im Sinne einer Vertrauenskultur verstanden werden muss.

Uneinigkeit besteht darüber, zu welchem Zeitpunkt eine Fehleranalyse im diagnostischen Prozess durchzuführen ist. Jacobs und Petermann (2012), Fritz und Ricken (2008) sowie Simon und Grünke (2010) verorten diese im Anschluss an die standardisierte Diagnostik (siehe Abbildung 24). Zimmermann (2014) hingegen ersetzt die standardisierte Diagnostik gänzlich durch ein Beratungsgespräch inklusive Fehleranalyse (Frankfurter integratives psycho- und lerntherapeutisches Konzept; FIT), und bei Lorenz (2014) wird die Fehleranalyse sogar vor dem Anamnesegespräch gesehen.

In jedem Fall werden bei einer Fehleranalyse die Probleme konkreter und detaillierter erfasst, als in der quantifizierten standardisierten Diagnostik. Das Ziel ist es herauszufinden, welche Schemata oder Strategien die betroffenen Schülerinnen und Schüler anwenden. Oftmals sind diese Strategien für die Lehrkraft nur schwer zu erkennen, sodass eine Fehleranalyse im Dialog durchgeführt wird, in dem der Schüler oder die Schülerin seine bzw. ihre Rechenwege verbalisiert.

Winter (2010, S. 35) nennt zwei zentrale Ziele einer Fehleranalyse:

1. die Kategorisierung von Fehlern bzw. Fehlermustern und
2. die Ursachenermittlung zu den einzelnen Fehlern und Fehlermustern

In der Definition einer Rechenstörung in der ICD-10 der Weltgesundheitsorganisation (2013) wird hervorgehoben, dass sich die Symptome dieser Störung hauptsächlich in den vier Grundrechenarten und eher weniger in Verfahren der höheren Mathematik zeigen. Bezüglich der inhaltlichen Relevanz und gleichzeitigen Fehleranfälligkeit der Grundrechenarten schreibt Winter (2010, S. 40):

> „Der Erwerb von Kenntnissen und Fähigkeiten in der Arithmetik bildet die Basis für alle weiteren mathematischen Themengebiete und Zusammenhänge. Damit ist die Arithmetik der Ausgangspunkt, der in allen Bereichen der Mathematik zu typischen Fehlern führen kann und auf den sich dementsprechend viele Fehlerphänomene zurückführen lassen."

Zur Unterstützung der Durchführung einer Fehleranalyse stellt Winter (2010) in ihrer Metaanalyse einen umfassenden Katalog mit typischen Fehlermustern und -phänomenen auf. Der Katalog reicht von den halbschriftlichen und schriftlichen Rechenverfahren in den vier Grundrechenarten über Fehlertypen beim Rechnen mit der Null bis hin zu Fehlertypen beim Bruchrechnen in der Addition, Subtraktion, Multiplikation und Division (siehe dazu Wittmann, 2007a). Tabelle 14 zeigt einen Ausschnitt aus dem Katalog für den Bereich der halbschriftlichen und mündlichen Rechenverfahren in den vier Grundrechenarten.

TABELLE 14 OPERATIONSÜBERGREIFENDE FEHLER IN (HALB-)SCHRIFTLICHEN UND MÜNDLICHEN ARITHMETISCHEN RECHENVERFAHREN (WINTER, 2010, S. 43)

Fehlertyp	**Fehlermuster/Fehlerbeschreibung**	**Fehlerphänomen**
E1	Zählfehler, Einsundeinsfehler der Nähe: Ergebnis weicht um eins vom korrekten Ergebnis ab.	$33 + 5 \rightarrow 37$ $33 + 5 \rightarrow 39$ $38 - 5 \rightarrow 32$ $38 - 5 \rightarrow 34$
E2	Zählfehler, Einmaleinsfehler der Nähe: Ergebnis weicht um ein Vielfaches vom korrekten Ergebnis ab.	$7 * 4 \rightarrow 24$ $4 * 7 \rightarrow 35$
E3	Zählfehler, Einsdurcheinsfehler der Nähe: Ergebnis weicht um ein Vielfaches vom korrekten Ergebnis ab.	$42 : 7 \rightarrow 7$ $42 : 7 \rightarrow 5$
IO	Fehler der inversen Operation, bspw.: $a + b \rightarrow a - b$ oder $a * b \rightarrow a : b \mid a : b = c \wedge c \in \mathbb{N}$	$37 + 12 \rightarrow 25$ $25 * 5 \rightarrow 5$

Weitere frei zugängliche informelle Verfahren, die qualitative Hinweise auf die Rechenschwierigkeiten ermöglichen, sind zum Beispiel der „Dortmunder Rechentest für die Eingangsstufe“ (DORT-E) (Moog & Schulz, 2005) oder die „Qualitative Erfassung von Lernschwierigkeiten in Mathematik“ (Gerster & Schulz, 2004). Simon und Grünke (2010, S. 41) empfehlen, die sich anschließende Förderung formativ zu evaluieren. Dazu bieten sich zum Beispiel curriculumbasierte Messverfahren an, die den Lernfortschritt der Schülerinnen und Schüler abbilden (Gold, 2015). Curriculumbasierte Messverfahren ergeben nicht nur quantitative, sondern auch qualitative Hinweise auf die domänenspezifischen Stärken und Schwächen der Schülerinnen und Schüler.

3.5 Grundsätze der Förderung

Je nach Ausweitung des diagnostischen Prozesses wird die Phase der Förderung mit in diesen einbezogen. Dabei sollte sich die Gestaltung der Förderung bzw. der Förderphase eng an diagnostischen Erkenntnissen orientieren. Sobald (standardisierte) Testverfahren und die Fehleranalyse zeigen, welche Hürden für den weiteren Kompetenzerwerb bestehen, ist es möglich, eine passgenaue Förderung zu gestalten (Fritz & Ricken, 2008, S. 78). So stellen Ennemoser und Kollegen (2011, S. 229) fest, dass sich zum Beispiel in der Sekundarstufe I Fördermaßnahmen oftmals auf die dort vermittelten Inhalte beziehen, ungeachtet dessen, dass die Schwierigkeiten auch im Grundschulstoff liegen können. In diesem Kapitel wird es um Grundsätze der Förderung bei Rechenschwierigkeiten gehen. Konkret wird die übergeordnete Frage thematisiert: „Welche Bereiche müssen gefördert werden?“. Hier wird der Schwerpunkt nicht auf eine Förderung bei einer Rechenstörung nach WHO-Kriterien gelegt, sondern es werden alle Schülerinnen und Schüler mit Rechenschwierigkeiten einbezogen. Schipper (2010, S. 106) schreibt, „der Versuch der WHO, Dyskalkulie zu definieren, ist für eine wissenschaftliche Begriffsklärung unbrauchbar und für die Förderung der Kinder eher kontraproduktiv“. Auch Simon (2005, S. 20 f.) schreibt:

> „Im praktischen Umgang mit Kindern ist es sowieso unwichtig, ob ein Kind eine Rechenschwäche hat oder nicht. Der praktische Umgang mit Kindern muss sich individuell an Stärken und Schwächen orientieren. Dafür ist eine differenziertere Aussage nötig als ‚rechenschwach’ oder ‚nicht rechenschwach’. [...] Die Unterscheidung in rechenschwach oder nicht rechenschwach ist nur dann von Bedeutung, wenn es darum geht, ob die Kosten von einer Therapie übernommen werden.“

In der Frage der Förderung bei Rechenschwierigkeiten stellt Lorenz (2014, S. 50) drei Bereiche auf, die parallel unterstützt werden sollten: Die curricularen Inhalte, die kognitiven Kompetenzen und der emotional-motivationale Bereich. Born und Oehler (2013, S. 10) beschreiben, dass nach Misserfolgserlebnissen oftmals ein „Teufelskreis“ in Gang gerät, der sich fortsetzen kann durch Lernabneigungen, Selbstabwertungen, Misserfolgserwartungen und letztlich durch eine Entmutigung zum Lernen. Dieser Teufelskreis spielt besonders bei Schülerinnen und Schülern, die massive Misserfolgserlebnisse im Rechnen erfahren haben, eine nicht zu unterschätzende Rolle (Warnke, 2005, S. 151).

Die folgende Tabelle 15 gibt einen Überblick über die drei zentralen Aspekte mit ihren Inhalten, die bei der Förderung von Rechenschwierigkeiten beachtet werden sollten. Die aufgelisteten Inhalte stellen sich zusammen aus den Beiträgen von Lorenz (2014), Gaidoschik (2010), Simon

und Grünke (2010), Jacobs und Petermann (2007), Schneider und Kollegen (2013, S. 233 f.) und Lambert (2015). Einzelne Aspekte lassen sich dabei auch mehreren Bereichen zuordnen.

TABELLE 15 ASPEKTE DER FÖRDERUNG BEI RECHENSCHWIERIGKEITEN (KATEGORISIERUNG NACH LORENZ, 2014, S. 50)

Förderung bei Rechenschwierigkeiten	
Curricular	• Systematischer Aufbau der Förderung • Fokussierung/Spezifikation auf mathematisches Basiswissen • Verbesserung der mathematischen Vorkonzepte/Vorläuferfertigkeiten (Klassifikationen, Zählen, Vergleichen, Mengen, Seriation, Teil-Ganzes-Konzept u. a.) • Ausbildung eines vertieften Verständnisses für den Zahlenraum bis 20 • Strategien im Zahlenraum bis 100 erarbeiten • Erarbeitung der Rechenverfahren Addition, Subtraktion, Multiplikation, Division • Förderung des Verständnisses des Stellenwertsystems • Einbezug von Sach-/Textaufgaben • Fördern im Bereich Geometrie und Gleichungen; später auch im Sek.-I-Stoff
Kognitiv	• Automatisierung üben und anwenden • Vorstellungsfähigkeit üben • Gedächtnis fördern, Arbeitsgedächtnis entlasten • Verwendung abstrakter Darstellungsmittel (DIENES-Material, Cuisenaire-Stäbchen etc.) • Auf Sprachwahl achten
Emotional-motivational	• Motivation anregen • Verbindung mit Alltagshandeln herstellen • Mathematikängstlichkeit abbauen • An realistischen Zielsetzungen/Attributionsmustern arbeiten • Formatives Feedback zur Leistungsentwicklung geben • Selbstwirksamkeitserwartung fördern • Einbezug der Eltern

Neben der Frage, *was* bei einer Förderung gefördert bzw. berücksichtigt werden sollte, spielt auch das *Wie* eine wichtige Rolle. Grundlegend für Unterricht und Förderung sind die beiden Prinzipen „Strukturorientierung“ und „Anwendungsorientierung“ (Krauthausen & Scherer, 2008, S. 300). Dabei gilt es weniger, eine der beiden Orientierungen in Extremform zu verfolgen, sondern im wechselseitigen Bezug zueinander umzusetzen.

Fuchs und Kollegen (2008, S. 85) stellen sieben zentrale Prinzipien auf:

1. Instructional explicitness
2. Instructional design to minimize the learning challenge
3. Strong conceptual basis
4. Drill and practice
5. Cumulative review
6. Motivators to help students regulate their attention and behavior to work hard
7. Ongoing progress monitoring

Im Zuge der Einführung des inklusiven Bildungssystems ist davon auszugehen, dass sich das Leistungsspektrum innerhalb einer Klasse weiter vergrößern wird (Schels et al., 2015). So wird es in den Klassen zunehmend stark rechenschwache Schülerinnen und Schüler zum Beispiel des Förderschwerpunkts Lernen oder Geistige Entwicklung geben, die spezifische Hilfe bedürfen, die nicht durch den regulären differenzierten Unterricht erbracht werden können. Hinzu kommen diejenigen Schülerinnen und Schüler, die temporär Schwierigkeiten im Fach Mathematik haben, was zum Beispiel durch belastende schulische oder außerschulische Kontextbedingungen verursacht werden kann. Neben den Schülerinnen und Schülern mit allgemeinen Rechenschwierigkeiten und Entwicklungsrückständen sind zudem diejenigen mit einer Rechenstörung (Dyskalkulie) zu erwähnen, die professionelle Unterstützung erwarten.

Die Umsetzung eines inklusiven Schulsystems erfordert, die im Regelfall auf langjährige Erfahrung getroffene Annahme über den Grad des Leistungsspektrums innerhalb einer Klasse dahingehend zu überdenken, dass er sich weiter ausfächern kann. Diese Grundüberlegung stellt einen wichtigen Aspekt bezüglich der Förderung der Schülerinnen und Schüler mit Rechenschwierigkeiten dar.

3.6 Zusammenfassung

Ca. 15–20 % der Schülerinnen und Schüler zeigen in ihrer schulischen Laufbahn *allgemeine* Rechenschwierigkeiten (Fischbach et al., 2013; Lorenz, 1993; Mazzocco & Myers, 2003), wobei die Prävalenzraten in verschiedenen Studien variieren (siehe Kapitel 3.2). Die Zahl der Schülerinnen und Schüler mit einer massiven Rechenstörung nach den Kriterien der Weltgesundheitsorganisation (2013) beläuft sich nach nationalen und internationalen Schätzungen auf ca. 6 % (Hein et al., 2000; Koumoula et al., 2004; Ramaa & Gowramma, 2002). Dabei können die Ursachen für Rechenschwierigkeiten im familiären oder schulischen Kontext oder auch in Bedingungen des Umfelds liegen. Eine spezifische Rechenstörung hingegen kann oftmals auf neurologische Gründe zurückgeführt werden, was jedoch weitergehende Begründungszusammenhänge nicht ausschließt (siehe Kapitel 3.3).

Bezüglich der pädagogischen Folgen besteht Konsens darüber, dass *alle* Schülerinnen und Schüler mit Schwierigkeiten im Rechnen einer Förderung bedürfen und sich diese nicht auf diejenigen mit einer Rechenstörung beschränken (Gold, 2015). In der Entwicklung bestehen oftmals „Lücken im grundlegenden mathematischen Handwerkszeug“ (Gaupp et al., 2004, S. 39) (siehe Kapitel 3.4).

Einer adäquaten und zielführenden Förderphase geht eine intensive Diagnostik voraus. Bereits in der Vorschule kann es erste Hinweise auf das Vorliegen von Entwicklungsschwierigkeiten geben (siehe Kapitel 3.4.1). Diese zeigen sich u. a. durch Probleme in der zeitlichen Orientierung, beim Vergleichen und Sortieren von Objekten oder bereits im Zahlverständnis (Tröster, 2008). Im Regelfall zeigen sich die Symptome von Rechenschwierigkeiten jedoch in der frühen Schulzeit, in der Schülerinnen und Schüler auf allen Kompetenzstufen Defizite erfahren, die den Prozess des mathematischen Kompetenzerwerbs erschweren. Standardisierte Rechentests wie zum Beispiel der Heidelberger Rechentest 1-4 (Haffner et al., 2005) und informelle Verfahren wie eine Fehleranalyse (Winter, 2010) ermöglichen eine genaue Beschreibung und Identifizierung der kritischen Stellen beim Kompetenzerwerb (siehe Kapitel 3.4.3). Diese mehrschrittige Diagnostik ermöglicht die Entwicklung einer Förderung, die am individuellen

Kompetenzstand der rechenschwachen Schülerin oder des rechenschwachen Schülers anknüpft. Dabei sollte die Unterstützung grundsätzlich curriculare Inhalte fördern sowie die Entwicklung kognitiver und emotional-sozialer Kompetenzen berücksichtigen (Lorenz, 2014). Die Frage des „Was ist zu fördern" schließt sich die Frage des „Wie und womit ist zu fördern" an (siehe Kapitel 3.5). Diese Frage wird Thema des folgenden Kapitels sein.

4 Evidenzbasierte Förderung mathematischer Basiskompetenzen

Nationale und internationale Querschnitt- und Längsschnittsudien zeigen, dass ein nicht kleiner Teil der Schülerinnen und Schüler in der Sekundarstufe I massive Defizite in der Beherrschung mathematischer Basiskompetenzen aufweist (siehe Kapitel 2.4). Dabei gehen diese Defizite nur bei ca. 2–6 % der Schülerinnen und Schüler mit einer gravierenden Rechenstörung einher (siehe Tabelle 10). Bei einem Großteil der Schülerinnen und Schüler handelt es sich vielmehr um allgemeine Rechenschwierigkeiten, die multifaktoriell bedingt sein können. Es stellt sich die Frage, wie diesen allgemeinen Schwierigkeiten und Kompetenzrückständen in den mathematischen Basiskompetenzen begegnet werden kann, um zu vermeiden, dass sich Defizite in der Sekundarstufe I manifestieren und damit der Anschluss an die weiteren curricularen Inhalte verpasst wird.

Eine evidenzbasierte Förderung mathematischer Basiskompetenzen, die dem oftmals vorzufindenden „wait-to-fail"-Problem (Vaughn & Fuchs, 2003) vorbeugt, stellt dabei einen möglichen Lösungsansatz dar. Dass das Thema evidenzbasierter Praxis auch in der Sonderpädagogik in Deutschland Einzug gehalten hat, lässt sich an den zahlreichen neueren Publikationen zu diesem Thema erkennen (Blumenthal & Mahlau, 2015; Casale et al., 2015; Hillenbrand, 2014, 2015; Križan & Vossen, 2016; Schad, 2015). Beim Einsatz evidenzbasierter Verfahren wird jedoch auch darauf hingewiesen, den Blick für die „Individuumszentrierung" (Böttinger, 2015) nicht zu verlieren.

4.1 Das Konzept Evidenzbasierung

Der Begriff der Evidenzbasierung hat in Deutschland im Zuge der Ratifizierung der UN-Konvention über die Rechte von Menschen mit Behinderungen im Jahr 2007 (UN-Behindertenrechtskonvention) (Vereinte Nationen, 2006) Einzug gehalten (Nußbeck, 2007). Der englischsprachige Gesetzestext besagt: „Persons with disabilities receive the support required, within the general education system, to facilitate their *effective* education" (Vereinte Nationen, 2006, Artikel 24, 2d; Herv. d. Verf.). Folglich wird durch die UN-Konvention über die Rechte von Menschen mit Behinderungen nicht nur ein *inklusiver*, sondern zudem ein *effektiver* Unterricht gesetzlich verankert. In der Frage der „Effektivität" von Unterricht ist der Begriff der Evidenzbasierung derjenige, der die Forderung inhaltlich greifbar macht. *Evidenzbasiert* bedeutet hier, dass in Unterrichts- und Förder-Settings Maßnahmen und Methoden zum Einsatz kommen, deren Wirksamkeit durch wissenschaftliche Forschung und durch nachvollziehbare Studien bestätigt sind und deren Einsatz sich im Rahmen der praktisch vorliegenden Problemstellung aus Sicht der beteiligten Akteure (z. B. der Lehrkraft) eignet (Hillenbrand, 2015, S. 314). Eine längere Tradition hat der Begriff bereits in der Medizin, wo er bedeutet: „auf der Basis empirisch zusammengetragener und bewerteter wissenschaftlicher Erkenntnisse erfolgend (von diagnostischen oder therapeutischen Maßnahmen)" (DUDEN, 2016c).

Im pädagogischen Kontext findet der Begriff seine Ursprünge in den USA im „No Child Left Behind Act" (NCLB-Act) aus dem Jahr 2002. In dem NCLB-Act wird festgelegt, dass der Unterricht und die Förderung benachteiligter Schülerinnen und Schüler „effective [and] scientifically based" (Senate and House of Representatives of the USA, 2002, Sec. 1001, Absatz 9)

gestaltet werden soll. „The most recent authorizations of the Individuals With Disabilities Education Act [...] and the No Child Left Behind Act of 2001 [...] provide a mandate for the use of evidence-based practices (EBPs) in the areas of academic and behavioral education" (Freeman & Sugai, 2013, S. 6). Durch den No Child Left Behind Act sind Schulen dazu verpflichtet, dass ein großer Anteil der Schülerinnen und Schüler (95 %) die Vorgaben des Bundesstaates im Lesen und im Fach Mathematik erreicht. Gelingt es einer Schule nicht, dieses Ziel umzusetzen, muss sie mit Sanktionen und ggf. mit einer Schließung rechnen (Senate and House of Representatives of the USA, 2002, 1111, Absatz i).

In den Vereinigten Staaten besteht, im Vergleich zu Deutschland, eine eindeutige Rechtslage darüber, dass in Unterricht und Förderung evidenzbasierte Methoden Einsatz finden müssen. Dieser rechtlichen Grundlage schließt sich die Frage an, welche Maßstäbe und Kriterien Methoden erfüllen müssen, um als evidenzbasiert gelten zu können. Weitgehende Einigkeit besteht darüber, dass zur Identifikation evidenzbasierter Methoden im pädagogischen Kontext, in Abgrenzung zum medizinischen Paradigma, *ausschließlich* wissenschaftliche Erkenntnisse und Evidenzen nicht genügen (Ennemoser, 2014; Kuhl & Euker, 2016).

> „Im Bereich der (Sonder-)Pädagogik greift jedoch zu kurz, nur den wissenschaftlichen Nachweis über die Wirksamkeit einer pädagogischen Maßnahme [...] zu berücksichtigen. Im pädagogischen Setting spielen auch das individuelle Wissen und die Erfahrung der Lehrkräfte sowie die Bedürfnisse der Schülerinnen und Schüler zu berücksichtigende Rollen." (Blumenthal & Mahlau, 2015, S. 408)

Whitehurst (2002) hat im Auftrag des Departments of Education der Vereinigten Staaten eine Übersicht zur Evidence-based Education gestaltet (siehe Abbildung 25). Er verdeutlicht darin, dass „Evidence based Education" zu gleichen Teilen sowohl auf empirischen Erkenntnissen (Empirical Evidence) als auch auf Expertenmeinungen (Professional Wisdom) gründet. Abbildung 25 zeigt sein Modell, ergänzt durch die Klassifikation nach Blumenthal und Mahlau (2015, S. 409) (grau hinterlegt). „Professional Wisdom" teilt der Autor in zwei Bausteine auf. „Individual Experience" meint dabei Expertenwissen und die Erfahrungen, die Professionelle (sprich: Lehrkräfte) selbst gemacht und als bewährt wahrgenommen haben. „Consensus" ist dabei eher auf das Kollektiv der Experten bezogen, in dem konsensual bestimmte Methoden und Praktiken als wirksam gelten. Blumenthal und Mahlau (2015, S. 409) bezeichnen den Einbezug des Wissens und der Erfahrung von Experten als „interne Evidenz".

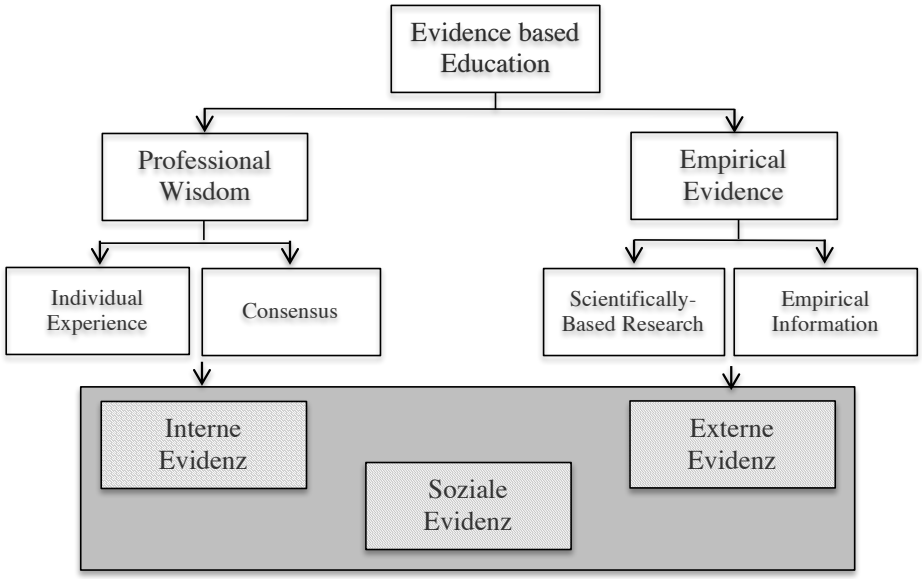


ABBILDUNG 25 ASPEKTE DER EVIDENZBASIERUNG (GRAU: NACH BLUMENTHAL & MAHLAU, 2015, S. 409; WEIß: NACH WHITEHURST, 2002)

Den zweiten großen Aspekt von Evidenzbasierung bezeichnet Whitehurst (2002) übergeordnet als „Empirical Evidence". Der durch den NCLB-Act betitelte und hier untergeordnete Aspekt lautet „Scientifically-Based Research". Im NCLB-Act wird darunter verstanden „research that involves the application of rigorous, systematic and objective procedures to obtain reliable and valid knowledge relevant to education activities and programs" (Senate and House of Representatives of the USA, 2002, 7801, 37A). Der Aspekt „Empirical Evidence" wird bei Blumenthal und Mahlau (2015, S. 409) „externe Evidenz" genannt.

Blumenthal und Mahlau (2015, S. 409) führen in ihrem Konzept der Evidenzbasierung neben der internen und externen Evidenz auch die soziale Evidenz auf. Dieser Aspekt wird in dem Modell nach Whitehurst (2002) nicht explizit benannt und umfasst den Einbezug der spezifischen Bedürfnisse der Schülerinnen und Schüler sowie ihrer Eltern.

Der NCLB-Act hat in der Folge zahlreiche Diskussionen zur Konkretisierung des Begriffs „scientifically-based" angestoßen (Freeman & Sugai, 2013; Zucker, 2004). In der wissenschaftlichen Literatur besteht noch keine inhaltliche Einigkeit über die definitorische Ausgestaltung. „Despite extensive discussion, universal agreement about the exact meaning of the definition of scientifically-based research in NCLB remains elusive" (Zucker, 2004, S. 3). In Deutschland liegen derzeit Vorschläge von Hartke (2005, S. 470), Fingerle und Ellinger (2008, S. 9) und Koch (2008, S. 89) vor. International stammt, neben Vorschlägen des U. S. Departments for Education (What Works Clearinghouse), dessen Kriterien jedoch nicht für den sonderpädagogischen Gebrauch konzipiert sind (Freeman & Sugai, 2013), oder der American Psychological Association (APA) bzw. des Promising Practices Network (PPN), ein Vorschlag für den Bereich der Sonderpädagogik vom Council for Exceptional Children (2014).

Im Folgenden wird die Klassifikation evidenzbasierter Förderung durch das Council for Exceptional Children (2014) detailliert vorgestellt, da sie Hinweise gibt, wie eine Förderung mathematischer Kompetenzen im inklusiven Setting realisiert werden kann, von der auch – bzw. gerade – die schwächsten Schülerinnen und Schüler profitieren.

Das Council for Exceptional Children ist die „größte[n] und einflussreichste[n] internationale[n] Vereinigung für Sonderpädagogik“ (Hillenbrand, 2015, S. 315) und hat die „Standards for Evidence-based Practices in Special Education“ (Council for Exceptional Children, 2014, S. 8 ff.) veröffentlicht. In Abhängigkeit von u. a. der Qualität des Studiendesigns, der Anzahl der Studien, der Anzahl der Probanden, der Effekte oder auch der Randomisierung wird eine Klassifikation von Methoden oder Maßnahmen in eine von fünf Evidenzstufen vorgenommen (Übersetzung der Kriterien in Klammern durch Hillenbrand (2015, S. 316)):

1. Evidence-based Practice (Evidenzbasierte Praxis)
2. Potentially evidence-based practice (Potentiell evidenzbasierte Praxis)
3. Mixed Evidence (Gemischte Evidenz)
4. Insufficient Evidence (Ungenügende Evidenz)
5. Negative effects (Negative Effekte)

Zur Einordnung in die entsprechende Stufe müssen dabei folgende Kriterien erfüllt sein.

TABELLE 16 KLASSIFIKATION EVIDENZBASIERTER PRAXIS (COUNCIL FOR EXCEPTIONAL CHILDREN, 2014, S. 8 FF; ÜBERS. D. VERF.)

1. Evidenzbasierte Praxis	
a	*Eines* der folgenden Kriterien muss mindestens erfüllt sein: • Zwei gut fundierte randomisierte Gruppenvergleichsstudien mit positiven Effekten und jeweils mindestens 60 Probanden • Vier nicht randomisierte Gruppenvergleichsstudien mit positiven Effekten und jeweils mindestens 120 Probanden • Fünf Einzelfallstudien mit positiven Effekten und jeweils mindestens 20 Probanden **ODER**
b	Bei mindestens zwei der in a genannten Möglichkeiten liegen die Kriterien zu 50 % vor. Zum Beispiel: • Eine gut fundierte randomisierte Gruppenvergleichsstudie mit positiven Effekten und mindestens 30 Probanden **und** drei Einzelfallstudien mit positiven Effekten und jeweils mindestens 10 Probanden **ODER** • drei Einzelfallstudien mit positiven Effekten und jeweils mindestens 10 Probanden **und** zwei fundierte nicht-randomisierte Gruppenvergleichsstudien mit positiven Effekten und jeweils mindestens 60 Probanden **UND**
c	• **Keine** gut fundierte Gruppenvergleichsstudie mit negativen Effekten **UND** mindestens einem 3:1 Verhältnis von Studien mit positivem Effekt zu denen mit einem neutralen/uneindeutigen Effekt. Bei c) werden Gruppenvergleichsstudien und Einzelfallstudien gleichwertig behandelt

2. Potentiell evidenzbasierte Praxis	
a	*Eines* der folgenden Kriterien muss mindestens erfüllt sein: • Eine gut fundierte randomisierte Gruppenvergleichsstudie mit positiven Effekten • Zwei oder drei fundierte nicht-randomisierte Gruppenvergleichsstudien mit positiven Effekten • Zwei bis vier Einzelfallstudien mit positiven Effekten
b	Bei mindestens zwei der in a genannten Möglichkeiten liegen die Kriterien zu 50 % vor. Zum Beispiel: • Eine gut fundierte nicht randomisierte Gruppenvergleichsstudie mit positiven Effekten **und** eine Einzelfallstudie mit positiven Effekten **UND**
c	**Keine** gut fundierte Gruppenvergleichsstudie mit negativen Effekten **UND** mindestens einem 2:1 Verhältnis von Studien mit positivem Effekt zu denen mit einem neutralen/uneindeutigen Effekt. Bei c) werden Gruppenvergleichsstudien und Einzelfallstudien gleichwertig behandelt

3.	**Gemischte Evidenz**
a	• Die Kriterien a **ODER** b im Falle Evidenzbasierte Praxis (1) oder Potenziell evidenzbasierte Praxis (2) sind erfüllt **UND**
b	• Das Verhältnis von Studien mit positivem Effekt zu denen mit einem neutralen/uneindeutigen Effekt ist mindestens 2:1 **ODER** eine oder mehr fundierte Gruppenvergleichsstudie zeigt einen negativen Effekt solange mehr fundierte Gruppenvergleichsstudien mit einem positiven Effekt vorliegen.

4.	**Ungenügende Evidenz**
	Eine ungenügende Evidenz liegt vor, wenn keine der drei Fälle Evidenzbasierter Praxis (1, 2, 3) erfüllt wird.

5.	**Negative Effekte**
a	Mindestens eine fundierte Studie jeglichen Designs, die negative Effekte zeigt **UND**
b	Die Anzahl der gut fundierten Studien mit negativen Effekten übersteigt diejenige mit positiven Effekten

Letztlich unabhängig von der konkreten Ausgestaltung des Begriffs ist es für Lehrkräfte im alltäglichen Geschehen oftmals sehr schwierig und aufgrund verschiedener Belastungen nicht möglich, sich einen detaillierten Einblick über die wissenschaftliche Evaluation von Förderkonzepten zu verschaffen (Blumenthal & Mahlau, 2015; Križan & Vossen, 2016). Forschungsergebnisse zeigen zudem, dass sich Lehrkräfte unsicher darin sind, welche Methoden und Maßnahmen im inklusiven Setting wirksam sind und kein gefestigtes Wissen darüber haben (DeSimone & Parmar, 2006; Hintz, Melzer, Scheer, Paal & Hillenbrand, 2015). Verwiesen wird an dieser Stelle auf Blumenthal und Mahlau (2015), die eine Schrittfolge vorschlagen, anhand derer Lehrkräfte die wissenschaftliche Fundierung eines Konzepts in vier Schritten beurteilen können. Die Beurteilung der Praxistauglichkeit dieser Schrittfolge steht noch aus.

Lehrkräften würde die Arbeit insgesamt erleichtert werden, wenn die Wissenschaft ihre Erkenntnisse zur Überprüfung von Unterrichtskonzepten verständlich und aufbereitet zur Verfügung stellt (Casale et al., 2015). Oftmals „[taugen] Ergebnisse für die ein oder andere Veröffentlichung und verschwinden anschließend schnell wieder in der Schublade“ (Križan &

Vossen, 2016, S. 83). Dahingehend ist die US-amerikanische Praxis weit fortgeschritten. Plattformen im Internet ermöglichen es Lehrkräften, evidenzbasierte Methoden und Konzepte zu spezifischen Themen zu finden, die sie dann im Unterricht oder in Fördersettings einsetzen können. Die folgende Tabelle gibt einen Überblick über verschiedene Internetplattformen. Die in der Fachliteratur an verschiedenen Stellen rezipierte Website „The Access Center" des American Institutes for Research ist darin nicht aufgeführt, weil diese derzeit inaktiv ist. Die Grüne Liste Prävention des Landespräventionsrats Niedersachsen ist die einzige deutschsprachige Homepage.

TABELLE 17 INTERNET-PLATTFORMEN UND WEBSEITEN ZUR AUSWAHL EVIDENZBASIERTER METHODEN UND MAßNAHMEN

	What Works Clearinghouse	**Promising Practices Network**	**Council for Exceptional Children**	**Blueprints for Healthy Youth Development**	**Technical Assistance Alliance for Parent Centers**
URL	http://ies.ed.gov/ncee/wwc/	http://www.promising-practices.net/	http://www.cec.sped.org	http://www.blueprints-programs.com/	http://www.pacer.org
Betreiber	Institute of Education Sciences	RAND Cooperation	Council for Exceptional Children	University of Colorado Boulder	U. S. Department of Education's Office of Special Education Programs
Zielgruppe	Fachkundige, die Präventions- und Interventionsprogramme durchführen wollen	Fachkundige, die Präventions- und Interventionsprogramme durchführen wollen	Sonderpädagogen/innen	Fachkundige, die Präventions- und Interventionsprogramme durchführen wollen	Familien mit Kindern mit einer Beeinträchtigung/ Sonderpädagogen
Auswahl nach ...	☒ Zielgruppe ☒ Förderbereich ☒ Fördersetting ☒ Evidenzstufe	☒ Zielgruppe ☒ Förderbereich ☒ Fördersetting ☒ Evidenzstufe	☐ Zielgruppe ☐ Förderbereich ☐ Fördersetting ☒ Evidenzstufe	☒ Zielgruppe ☒ Förderbereich ☒ Fördersetting ☒ Evidenzstufe	☒ Zielgruppe ☐ Förderbereich ☒ Fördersetting ☐ Evidenzstufe
Weiteres	☒ wissenschaftliche Transparenz ☒ nutzerfreundliche Oberfläche ☒ kostenfrei	☒ wissenschaftliche Transparenz ☐ nutzerfreundliche Oberfläche ☒ kostenfrei	☒ wissenschaftliche Transparenz ☐ nutzerfreundliche Oberfläche ☐ kostenfrei	☒ wissenschaftliche Transparenz ☒ nutzerfreundliche Oberfläche ☒ kostenfrei	☐ wissenschaftliche Transparenz ☒ nutzerfreundliche Oberfläche ☒ kostenfrei
Anmerkungen	Besonders benutzerfreundlich und differenziert; ermöglicht direkten Vergleich verschiedener Programme	Betrieb derzeit eingestellt	kostenpflichtig	Selektion nach vorhandenen individuellen Risiko- und Schutz-faktoren möglich	Stark elternzentriert

	Social Programs That Work	**Best Evidence Encyclopedia (BEE)**	**Grüne Liste Prävention**	**The Center for Comprehensive School Reform and Improvement**
URL	www.evidencebasedprograms.org	www.bestevidence.org	www.ctc-info.de	http://www.sedl.org
Betreiber	Coalition for Evidence-Based Policy	John Hopkins University, School of Education	Landespräventionsrat Niedersachsen	U.S. Department of Education
Zielgruppe	Fachkundige, die Präventions- und Interventionsprogramme durchführen wollen	Lehrkräfte	Kommunale Akteure und Netzwerke	Fachkundige aus dem Bereich emotional-soziale Entwicklung
Auswahl nach ...	☒ Zielgruppe ☐ Förderbereich ☒ Fördersetting ☒ Evidenzstufe	☒ Zielgruppe ☐ Förderbereich ☒ Fördersetting ☒ Evidenzstufe	☒ Zielgruppe ☐ Förderbereich ☐ Fördersetting ☒ Evidenzstufe	☒ Zielgruppe ☐ Förderbereich ☒ Fördersetting ☐ Evidenzstufe
Weiteres	☒ wissenschaftliche Transparenz ☒ nutzerfreundliche Oberfläche ☐ kostenfrei	☒ wissenschaftliche Transparenz ☐ nutzerfreundliche Oberfläche ☒ kostenfrei	☒ wissenschaftliche Transparenz ☒ nutzerfreundliche Oberfläche ☒ kostenfrei	☒ wissenschaftliche Transparenz ☒ nutzerfreundliche Oberfläche ☒ kostenfrei
Anmerkungen	kostenpflichtig	–	elternzentriert	–

Abschließend wird an dieser Stelle erläutert, wie der Ablauf eines evidenzbasierten Entscheidungsprozesses bei Lehrkräften aussehen kann. Dabei sollte jedoch nicht die Verantwortlichkeit der Wissenschaft und der Politik außer Acht gelassen werden (Casale et al., 2015; Križan & Vossen, 2016). Es liegen verschiedene Vorschläge zum Ablauf eines solchen Prozesses vor (siehe Blumenthal & Mahlau, 2015; Brown-Chidsey & Steege, 2010; Freeman & Sugai, 2013), die unterschiedlich konkretisiert und umfassend sind. Es wird hier das Vorgehen nach Freeman und Sugai (2013) vorgestellt, da es die verschiedenen Vorschläge gut zusammenführt und verständlich veranschaulicht (Übersetzungen nach Hillenbrand (2015, S. 319). Der siebte Schritt ist von Hillenbrand (2015, S. 319) ergänzt worden.

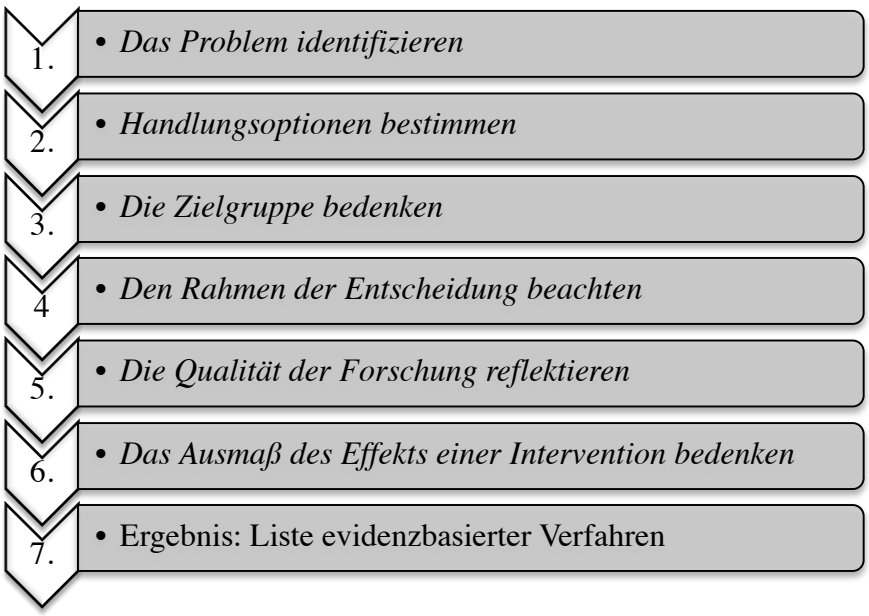


ABBILDUNG 26 SCHRITTE IN DER AUSWAHL EVIDENZBASIERTER VERFAHREN (FREEMAN & SUGAI, 2013, S. 10)

In einem **ersten Schritt** sollte die Lehrkraft die Problemlage genau wahrnehmen und das Problem konkret benennen können (vgl. Kapitel 3.4 Diagnostik). In welchem Bereich liegt das Problem des Schülers bzw. der Schülerin oder der Lerngruppe (Lesen, Schreiben, Rechnen, emotionale-soziale Entwicklung, Verschiedenes) und in welchem Kontext tritt es in welcher Intensität auf? Auch müssen individuelle Aspekte, die Einfluss auf eine mögliche Intervention nehmen könnten, berücksichtigt werden. Dazu gehören u. a. die allgemeine schulische Entwicklung, Ressourcen, schulische und außerschulische Risikofaktoren und auch die individuellen Bedürfnisse des Schülers bzw. der Schülerin. Hier sollte sich die Lehrkraft auch darüber bewusst werden, was das gewünschte „Ziel" ist.

In einem **zweiten Schritt** ist es die Aufgabe der Lehrkraft, mögliche Interventionen für die spezifische Fördersituation herauszufinden. Dies erfolgt über verschiedene Recherche-Kanäle wie zum Beispiel die o. a. Webseiten oder auch über Fachliteratur. Die Recherche im Internet mag aufgrund der geringen Anzahl zuverlässiger Webseiten erschwert sein. Hillenbrand (2015, S. 319) empfiehlt deswegen auch die Nutzung allgemeiner Recherchestrategien und auch das Hinzuziehen von Literaturdatenbanken wie www.fis-bildung.de. In diesem zweiten Schritt entsteht eine Liste mit potenziell passenden Fördermaßnahmen, die auch in Kooperation mit einer weiteren Lehrkraft erstellt werden kann.

In einem **dritten Schritt** werden die potenziellen Maßnahmen auf die Kompatibilität mit der „Zielgruppe“ geprüft. Sind die Bedingungen der Versuchsgruppe mit denen der eigenen Zielgruppe bzw. des einzelnen Schülers oder der Schülerin (Alter, Sprache etc.) vergleichbar? Lässt sich die Maßnahme so modifizieren, dass sie noch dem ursprünglichen Konzept entspricht? Können bei dem eigenen vorliegenden Fall ähnliche Effekte wie in der Versuchsgruppe erwartet werden?

Der **vierte Schritt** dient dazu, sich über die Reichweite einer möglichen Entscheidung bewusst zu werden. Soll zum Beispiel auf Schulebene ein neues Beratungskonzept oder für einen spezifischen Jahrgang eine verpflichtende Unterrichtsmethode eingeführt werden, bedarf diese Entscheidung einer gründlicheren Recherche sowie stärkeren Absprachen als die Bearbeitung eines kleinen Problembereichs mit einem einzelnen Schüler oder einer Kleingruppe (zum Beispiel Zehnerübergang). Im Einzelsetting ist es Lehrkräften schneller und einfacher möglich, Modifikationen im Förderprozess durchzuführen. Dies ist bei weitreichenden Implementationen eher schwierig, sodass bei der Einführung einer solchen umfassenden Maßnahme gründlich geprüft werden sollte, ob in der Evaluationsstudie ein klarer Bezug zwischen Maßnahme und Effekt zu erkennen ist und ob diese verschiedentlich rezipiert ist, sodass die Forschungslage ausreichend für eine fundierte Entscheidung ist.

Im **Schritt 5** wird die Qualität der Evaluationsstudien untersucht. Wird ein hochwertiges Design verwendet? Welche Effekte ergeben sich im Zusammenhang mit der Intervention (Kausalität)? Zudem ist zu prüfen, ob die Intervention, die Stichprobe und die Untersuchungsbedingungen konkret beschrieben werden. Dieser Schritt 5 ist sehr komplex, sodass dafür PfadDiagramme zum Beispiel von Freeman und Sugai (2013) oder von Blumenthal und Mahlau (2015) genutzt werden können, die die zentralen Aspekte berücksichtigen.

Im **Schritt 6** wird ein kritischer Blick auf die Effekte in der Maßnahme in den Evaluationsstudien geworfen. Untersucht wird, ob der Effekt der Maßnahme tatsächlich auf der Maßnahme beruht, ob die Ergebnisse eindeutig und klar verständlich präsentiert werden und ob die Beurteilung des Effekts nicht nur visuell sondern ebenso anhand von statistischen Verfahren erfolgt.

Abschließend ist das Rechercheergebnis eine Liste mit potentiellen evidenzbasierten Verfahren, die jedoch nur bei Hillenbrand (2015) im **siebten Schritt** aufgeführt wird.

Im „Individuals with Disabilities Education Improvement Act“ (Senate and House of Representatives of the USA, 2004) wird gesetzlich ein schulisches Rahmenkonzept verankert, das präventiv ausgerichtet ist und in der Evidenzbasierung einen essentiellen Bestandteil sieht. Dieses Rahmenkonzept, das Response to Intervention-Modell, wird im Folgenden dargestellt.

4.2 Rahmenkonzept Responsives Handlungsmodell (RTI)

Der weit verbreitete englischsprachige Begriff für das „Responsive Handlungsmodell“ lautet „Response to Intervention“ oder oftmals auch synonym „Response to Instruction“ (kurz: RTI) (Brown-Chidsey & Steege, 2010). „Response“ heißt ins Deutsche übersetzt „Antwort“ oder „Reaktion“, „Response to Intervention“ folglich etwa „Reaktion auf Intervention“.

Das Responsive Handlungsmodell stammt aus den USA, wo es jedoch keinen „Erfinder“ oder „Entdecker“ im engeren Sinne hat. Verbreitete Anwendung findet das RTI-Modell seit 2004, das Jahr des „Individuals with Disabilities Education Improvement Act“ (siehe 4.2.1).

Im Rahmen des Responsiven Handlungsmodells spielen zwei Aspekte eine besondere Rolle: Der Einsatz evidenzbasierter Verfahren in Unterricht und Förderung sowie eine entwicklungsbezogene, engmaschige Lernverlaufsdiagnostik (Hillenbrand, 2014) (siehe Abbildung 27). Das Responsive Handlungsmodell ermöglicht es Lehrkräften, frühzeitig zu erkennen, ob die Schülerinnen und Schüler im Unterrichtssetting oder in Förderphasen gewünschte Lernfortschritte zeigen und somit positiv auf den Unterricht oder die Förderung „reagieren". Das RTI-Modell stellt eine Lösung für das „Wait-to-fail"-Problem (Vaughn & Fuchs, 2003) dar, bei dem Lern- bzw. Entwicklungsprobleme bei Schülerinnen und Schüler häufig erst dann erkannt werden, sobald sie sich manifestieren (Kuhl & Euker, 2016; Voß, 2013).

Die implementierte Lernverlaufsdiagnostik (in Abbildung 27: Data) ermöglicht es, diejenigen Schülerinnen und Schüler zu identifizieren, für die der reguläre Unterricht (Ebene 1: Universelle Ebene) nicht ausreicht, da sie keine ausreichenden Lernfortschritte machen. Dies sind ca. 20 % der Schülerinnen und Schüler (siehe Abbildung 27). Sie benötigen eine intensive Hilfe auf der Ebene 2 (Selektive Ebene). Durch eine fortlaufende enge diagnostische Begleitung können auf der selektiven Ebene diejenigen Schülerinnen und Schüler erkannt werden, für die auch diese zusätzliche Unterstützung nicht ausreicht. Sie erhalten, oftmals im Einzelsetting, weitere Unterstützung auf der dritten Ebene (Ebene 3: Indizierte Ebene). Dies sind etwa 5 % der Schülerinnen und Schüler.

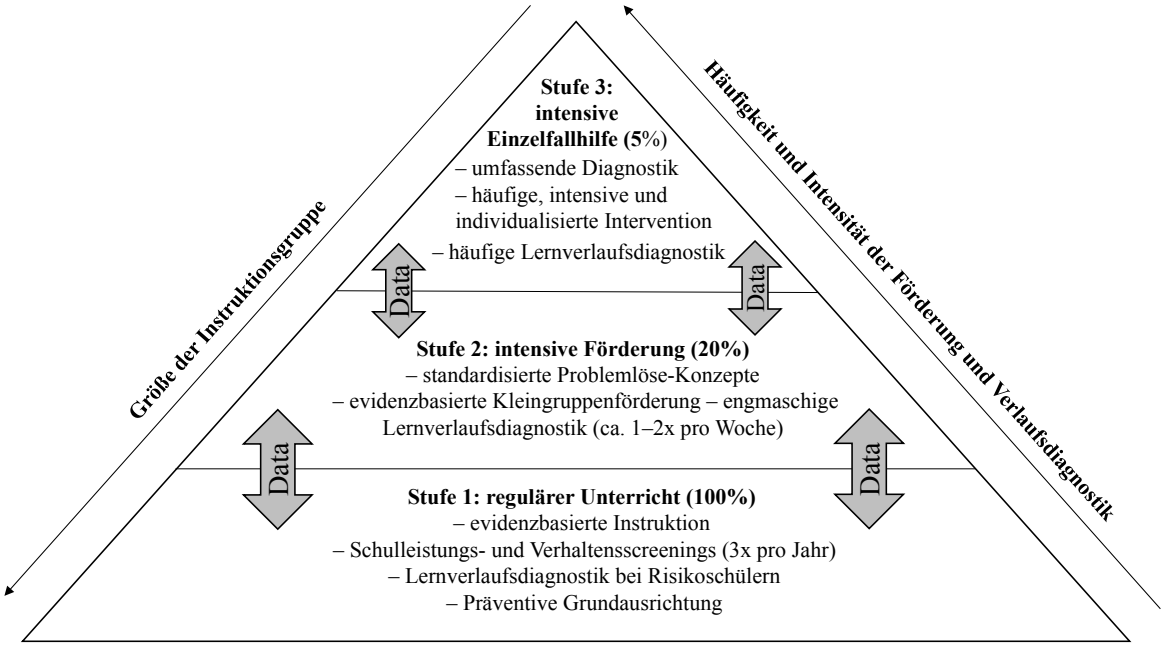


ABBILDUNG 27 RESPOSE TO INTERVENTION-MODELL (HUBER & GROSCHE, 2012, S. 314 ERWEITERT NACH BROWN-CHIDSEY & STEEGE, 2010, S. 5; D. VERF.)

Im Sinne einer Mehrebenenprävention werden alle Schülerinnen und Schüler, je nach Bedarf, unterschiedlich intensiv gefördert. Auch der reguläre Unterricht versteht sich dabei bereits als Förderstufe. Das Responsive Handlungsmodell „encompasses a framework within which teachers may evaluate whether students perform as expected when provided with evidence-based instructional practices" (Lembke, McMaster & Stecker, 2010, S. 22).

Die Abbildung 27 zeigt das Responsive Handlungsmodell in der häufig vorzufindenden Form einer Pyramide. Der Begriff „Data", der als Verbindungselement für die einzelnen Stufen steht, zeigt, dass Entscheidungen über die Intensivierung oder Extensivierung der Unterstützung stets

datenbasiert erfolgt. Das Responsive Handlungsmodell ist als durchlässiges Modell zu verstehen, in dem anhand diagnostischer Erkenntnisse ein steter, flexibler Wechsel zwischen den Stufen möglich ist (Hillenbrand, 2014; Huber & Grosche, 2012).

4.2.1 Entstehung und Entwicklung

Mit dem „Individuals with Disabilities Education Improvement Act“ (IDEA, 2004) wird im Jahr 2004 auf eine unzureichende diagnostische Praxis reagiert, die in der Vergangenheit oftmals dazu führte, dass zu viele Kinder zu unrecht als lernschwach eingestuft wurden (Voß, 2013, S. 115). Dieser Anstieg spiegelte sich auch im Anstieg des Risikoindexes für Lernbehinderungen (Risk Index for Learning Disabilities) wider (zum Beispiel 1974: 1,21 %; 1998: 6,02 %) (Donovan & Cross, 2002, S. 50).

Im IDEA wird festgelegt, wie beeinträchtigte oder von einer Beeinträchtigung bedrohte Kinder und Jugendliche frühzeitig erkannt und gefördert werden können. Das IDEA bildet dabei die rechtliche Grundlage für die schulische Umsetzung des Response to Intervention-Konzepts. Es wird darin deutlich, dass a) schulischer Unterricht und Förderung evidenzbasiert erfolgen, b) dass Unterstützungsbedarfe auf Grundlage der Responsitivät auf den schulischen Unterricht oder Förderung festgestellt werden, c) Lernentwicklungen formativ erhoben werden und d) Entscheidungen über weitere Fördermaßnahmen datenbasiert erfolgen.

Da die deutschen und englischsprachigen Begrifflichkeiten nicht deckungsgleich verwendet werden können, wird ergänzend angemerkt, dass im IDEA unter „specific learning disabilities“ gefasst wird: „a disorder in 1 or more of the basic psychological processes involved in understanding or in using language, spoken or written, which disorder may manifest itself in the imperfect ability to listen, think, speak, read, write, spell, or do mathematical calculations“ (IDEA, 2004, 602(30)).

Das Responseto Intervention-Konzept ist in den USA mittlerweile weit verbreitet. Im Jahr 2009 sind bereits 47 von den 50 US-Amerikanischen Staaten auf dem Weg, das Konzept in ihren Schulen zu implementieren. Einige Staaten sind dahingehend schon fortgeschrittener, andere stecken noch in der Vorbereitungsphase (Berkeley, Bender, Peaster & Saunders, 2009). Zahlreiche Handlungsleitfäden und nutzerfreundliche Internetplattformen unterstützen dabei die Umsetzung. Auch auf wissenschaftlicher Seite wird auf die Entwicklung reagiert. Im Jahr 2005 veröffentlicht zum Beispiel das Journal of Psychoeducational Assessment ein Themenheft (Nr. 23) ausschließlich zum RTI-Konzept.

Die Umsetzung und Evaluation des Response to Intervention-Konzepts bezieht sich national (Diehl, Mahlau, Voß & Hartke, 2010; Huber et al., 2013) und international (Fitzell, 2011; Lenz & Deshler, 2008) bisher im Regelfall auf die Primarstufe. Der Einsatz in der „besondere[n] Dynamik des Sekundarstufensetting“ (Lenz & Deshler, 2008, S. 445) wird bisher oftmals außer Acht gelassen. Hier besteht eine Forschungslücke sowohl in Wissenschaft als auch in Praxis.

4.2.2 Grundprinzipien

Nach Diehl, Mahlau, Voß und Hartke (2010) zählen die Mehrebenenprävention, eine evidenzbasierte Praxis, datenbasierte Entscheidungen und die Lernfortschrittsdiagnostik zu den vier

Grundprinzipien des RTI-Ansatzes. In der Literatur lassen sich jedoch unterschiedliche Akzentsetzungen bzw. „Key Features“ (Brown-Chidsey & Steege, 2010) feststellen, die als die zentralen Aspekte des Modells gelten (Forbringer & Fuchs, 2014; Kuhl & Euker, 2016). Die vier genannten Grundprinzipien nach Diehl, Mahlau, Voß und Hartke (2010) stellen zusammenfassend die wichtigsten Aspekte dar und werden auch den verschiedenen Akzentsetzungen anderer Autoren gerecht, sodass sie im Folgenden erläutert werden.

4.2.2.1 Mehrebenenprävention

Das RTI-Modell versteht sich als Konzept der Mehrebenenprävention. Schülerinnen und Schüler, für die die Angebote auf Stufe 1 oder 2 nicht ausreichen, um die gewünschte Lernentwicklung zu zeigen, werden als sog. „nonresponder“ (Fuchs & Fuchs, 2006, S. 98) bezeichnet. Sie erhalten differenzierte und gestufte Unterstützung. Die Förderung auf Stufe 2 kann in einer Kleingruppe von drei bis fünf Schülerinnen und Schüler erfolgen. Empfohlen wird jedoch auch, dass auf der Stufe 2 jeden Tag für 30 Minuten eine zusätzliche Förderung durch die Klassenlehrkraft erfolgt (Forbringer & Fuchs, 2014). Die Förderung auf Stufe 3 wird im Regelfall im Einzelsetting durch eine Sonderpädagogin bzw. einen Sonderpädagogen durchgeführt. Die Schülerinnen und Schüler auf dieser Stufe benötigen ca. 60 Minuten am Tag zusätzliche Förderung (Forbringer & Fuchs, 2014). Zeigt sich, dass ein Kind die anvisierten Lernfortschritte aufweist, können die zusätzlichen Fördermaßnahmen reduziert werden, sodass das Kind wieder auf Stufe 1 (oder 2) unterrichtet wird. Unter welchen Bedingungen ein Schüler oder eine Schülerin als „nonresponder“ gilt, ist dabei sehr unterschiedlich (Tran, Sanchez, Arellano & Swanson, 2011), was zu divergierenden Prävalenzraten von 2,7 % bis zu 44 % führen kann (Burns, Appleton & Stehouwer, 2005)

4.2.2.2 Evidenzbasierte Praxis

Der „No Child left Behind Act“ und der „Individuals with Disabilities Education Improvement Act“ aus den Jahren 2002 und 2004 fordern einen Unterricht, in dem auf allen Ebenen evidenzbasierte Maßnahmen und Methoden eingesetzt werden (siehe Kapitel 4.2.1). Somit bezieht sich die Forderung auch auf die universelle Ebene (Stufe 1), in der Unterricht qualitativ hochwertig erfolgen soll (Forbringer & Fuchs, 2014, S. 26; Voß, 2013, S. 121). Evidenzbasierte schulische Praxis ist jedoch kein Garant für schulischen Erfolg, sondern kann lediglich die Wahrscheinlichkeit für das Erreichen eines bestimmten Lernziels erhöhen (Brown-Chidsey & Steege, 2010, S. 42). Im Kapitel 4.1 ist bereits dargelegt, wie Lehrkräfte entscheiden können, ob es sich bei einer spezifischen Methode um eine evidenzbasierte handelt und welche Webseiten es derzeit gibt, die einen Überblick über evidenzbasierte Maßnahmen anbieten. Auf weitere Ausführungen wird deswegen an dieser Stelle verzichtet.

4.2.2.3 Datenbasierte Entscheidung

Die Informationen über die Lernentwicklungen der Schülerinnen und Schüler werden im RTI-Modell über drei diagnostische Wege erhoben (vgl. auch Kapitel 3.4 zum Thema Diagnostik) (Forbringer & Fuchs, 2014, S. 2 f.):

1. Universelle Screenings
2. Förderdiagnostik
3. Lernverlaufsdiagnostik (CBM)

Alle Schülerinnen und Schüler durchlaufen ca. zweimal im Jahr Screeningverfahren in spezifischen schulischen Domänen (zum Beispiel lesen, schreiben, rechnen, Verhalten). Diese Screeningverfahren ermöglichen es, auf einem ökonomischen Weg diejenigen Schülerinnen und Schüler zu identifizieren, die sich in ihrer Entwicklung im risikobehafteten Bereich befinden und intensivere Unterstützung benötigen. Zur Einschätzung der Ergebnisse halten die Screeningverfahren Normwerte bereit (sog. Cutoff-Werte), die eine Verortung der individuellen Leistungen im Verhältnis zur Referenzstichprobe ermöglichen. Bedeutsam ist dabei, wissenschaftlich geprüfte und fundierte Verfahren zu verwenden, die den hohen Ansprüchen an die Güte diagnostischer Verfahren Stand halten (Hillenbrand, 2015). Für die durch die Screenings identifizierten „Risikokinder“ schließt sich eine Förderdiagnostik an. Diese gibt Aufschlüsse über die spezifischen Problembereiche und ist dahingehend von großer Bedeutung, dass die Fördermaßnahmen, die in Stufe 2 und 3 angeboten werden, genau zu den Bedürfnissen der einzelnen Schülerinnen und Schüler passen. Einen gestuften Überblick über standardisierte Verfahren im deutschsprachigen Bereich für die Sekundarstufe I für die drei Stufen des RTI-Konzepts bietet Abbildung 23.

Alle Schülerinnen und Schüler werden letztlich durch eine Lernverlaufsdiagnostik begleitet, die die individuelle Lernentwicklung zeitnah abbildet und Aufschlüsse über Fördererfolge bietet.

Ein Beispiel für datenbasierte Entscheidungen auf Grundlage von Lernverlaufsdiagnostik zeigt Abbildung 28.

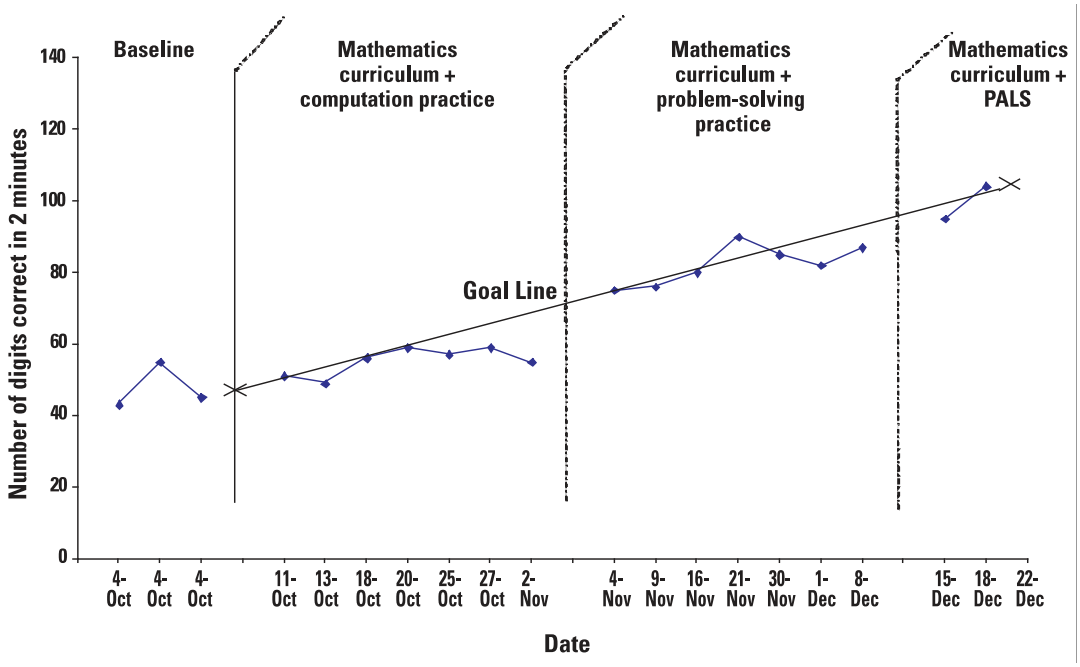


ABBILDUNG 28 FORMATIVE DIAGNOSTIK DER LEISTUNGSENTWICKLUNG EINES KINDES IM BEREICH MATHEMATIK (LEMBKE & STECKER, 2007, S. 2)

Die eingesetzten curriculumbasierten Messverfahren (CBM) sind hier zweimüntige Kurztests, die als Parallelformen zur Anwendung kommen. Abgebildet ist die Anzahl richtig gelöster Aufgaben. Bleibt der Schüler in seiner Entwicklung unterhalb der Goal Line, wird die Förderung modifiziert (siehe 2. November oder 8. Dezember).

Diehl und Hartke (2007, S. 20) bezeichnen das curriculumbasierte Messen lediglich als einen „Impulsgeber in einem mehr oder minder qualifiziert stattfindenden Vermittlungsprozess". Um zu vermeiden, dass Daten seitens der Lehrkräfte fehlinterpretiert oder auch gar keine Beachtung finden, intendieren Stecker und Kollegen (2005): „Teachers will need systematic support in interpreting and using CBM data for instructional planning". In der Literatur liegen verschiedene Vorschläge zur Interpretation von CBM-Daten vor. Walter (2014a) schlägt vor, in einer Baseline drei Messzeitpunkte zu erheben, davon den Median zu bilden und daran orientiert das Entwicklungsziel abzuleiten. Liegen drei aufeinanderfolgende Punkte unterhalb der Entwicklungslinie, sollte eine Modifikation der Förderung stattfinden (siehe Abbildung 28).

Hosp und Hosp (2007) haben die wöchentlichen Wachstumsraten zahlreicher Studien und Untersuchungen synthetisiert und davon eine Orientierung abgeleitet, wie groß der Lernzuwachs pro Woche sein kann (siehe Tabelle 18). Die Werte beziehen sich dabei auf den US-amerikanischen Durchschnitt; sie sind also nicht speziell für Schülerinnen und Schüler mit Lernschwierigkeiten normiert. Neben den Werten im Bereich Mathematik, haben die Autoren auch welche im Bereich Lesen und Rechtschreiben veröffentlicht. Keine Normwerte gibt es bisher zum Kompetenzbereich Schreiben (i. S. von Texte verfassen) (Hosp et al., 2007, S. 93).

TABELLE 18 „GROWTH RATES“ IM BEREICH MATHEMATIK (HOSP ET AL., 2007, S. 47; S. 79; S. 109)

	Math	
Grade	Realistic	Ambitious
1	0,3	0,5
2	0,3	0,5
3	0,3	0,5
4	0,7	1,15
5	0,75	1,20
6	0,45	1,0

4.2.2.4 Lernverlaufsdiagnostik

Lernverlaufsdiagnostik stellt ein zentrales Element des RTI-Konzepts dar und konkretisiert den zuvor erläuterten Aspekt „datenbasierte Entscheidung“. In regelmäßigen Abständen finden auf allen drei Ebenen (siehe Abbildung 27) sowohl im Gruppen- als auch im Einzelsetting curriculumbasierte Messungen (CBM) als Möglichkeit der Lernverlaufsdiagnostik statt (siehe Abbildung 28).

Der im RTI-Konzept verankerte Begriff des „Progress Monitoring“ (Riccomini & Witzel, 2010) wird in Deutschland mit zahlreichen Begriffen wie „Lernfortschrittsmessung“ (Diehl & Hartke, 2007; Klauer, 2006; Walter, 2008a), „Lernfortschrittsdiagnostik“ (Walter, 2009) oder „Lernverlaufsdiagnostik“ (Klauer, 2011a) übersetzt. „Da die Forschung aber zeigte, dass keinesfalls immer von Lern*fortschritten* die Rede sein kann, scheint der Begriff der ‚Lernverlaufsdiagnostik’ (Strathmann & Klauer, 2010) angemessener zu sein, der inzwischen auch verbreitet Anwendung findet (Klauer, 2014, S. 8; Herv. i. Orig.). Grundsätzlich aber können die Begriffe „synonym behandelt werden“ (Voß, 2013, S. 165).

Zu beachten ist zudem, dass der englischsprachige Begriff „Progress Monitoring“ nicht zwingend nur das Verfahren der Lernverlaufsdiagnostik beinhaltet. Etscheidt (2007, S. 58 f.) fasst darunter zum Beispiel auch das „curriculum-based assessment“ (CBA), das *viele* Möglichkeiten der Informationsbeschaffung über die Leistungsentwicklung des Schülers oder der Schülerin zulässt (Beobachtungen, Klassenarbeiten, Analysen von Arbeitsblättern, CBM etc.) (zu CBA siehe auch Deno, 1993; Klauer, 2006; Voß & Hartke, 2014). CBM können dabei lediglich als *eine* Möglichkeit des CBA verstanden werden. Im Rahmen des RTI-Modells ist mit dem Begriff „progress monitoring“ im Regelfall die Lernverlaufsdiagnostik in Form von curriculumbasierten Messungen gemeint. Die unterschiedliche inhaltliche Ausgestaltung des Begriffs muss jedoch berücksichtigt werden.

CBM sind Kurztests, die den Ansprüchen der wissenschaftlichen Güte genügen, nicht länger als 1 bis 3 Minuten dauern und die Lernentwicklung der Schülerinnen und Schüler in spezifischen schulischen Domänen zuverlässig abbilden. Lembke und Stecker (2007, S. 1) formulieren drei zentrale Charakteristika von CBM:

- „The measures need to be reliable and valid for the purpose for which they are used,
- they need to be short and easy to administer, and, most importantly
- they need to be highly related to other measures of proficiency in academic areas.“

CBM geben Auskunft darüber, ob die Schülerinnen und Schüler in ihrer Entwicklung Lernfort- oder -rückschritte machen oder stagnieren. Die Angaben darüber, wie regelmäßig die CBM-Messungen im RTI-Modell durchgeführt werden, variieren in der Literatur, wie Tabelle 19 zeigt.

TABELLE 19 HÄUFIGKEIT CURRICULUMBASIERTER MESSUNGEN IM RTI-MODELL (KÄTER, KÄTER, MARTENSTEIN & HILLENBRAND, 2016, S. 169)

Stufe	**Häufigkeit der Messung**			
1	3x im Jahr	mind. monatlich	3–4x im Jahr	monatlich
2	wöchentlich bis zweiwöchentlich	wöchentlich	1–4x im Monat	wöchentlich
3	wöchentlich	mind. wöchentlich	wöchentlich	wöchentlich
	Forbringer & Fuchs (2014)	National Center on Response to Intervention (2010)	Riccomini & Witzel (2010)	Mahlau, Diehl, Voß & Hartke (2010)

Einigkeit bzgl. CBM besteht darüber, dass die Häufigkeit der Messungen davon abhängt, zu welchem Zweck sie angewendet werden. Genießt ein Schüler eine gesonderte Förderung auf Stufe 2 oder 3, wird seine Lernentwicklung in geringeren Zeitabständen erhoben und abgebildet als in der gesamten Lerngruppe auf universeller Ebene (Voß & Hartke, 2014, S. 87). Die Dokumentation der Lernentwicklung bietet die Möglichkeit zu erkennen, ob ein Schüler oder eine Schülerin eine spezifische Förderung benötigt oder ob eine eingesetzte Intervention prognostisch zum gewünschten Ziel führt. Zudem erhält die Lehrkraft ein Feedback zur Wirksamkeit ihres Unterrichts oder ihrer Förderung. Eine Lernverlaufsdiagnostik dient,

> „to assess students' performance over time, to quantify student rates of improvement or responsiveness to instruction, to evaluate instructional effectiveness, and for students who are least responsive to effective instruction, to formulate effective individualized programs." (National Center on Response to Intervention, 2010, S. 6)

In den letzten Jahren hat die Forschung zu verschiedenen Themen im Bereich der Lernverlaufsdiagnostik stark zugenommen. Eine rege Diskussion besteht darüber, welche Gütekriterien für die Instrumente der Lernverlaufsdiagnostik gelten, da die Klassische Testtheorie mit ihren Grundsätzen an ihre Grenzen stößt (Strathmann & Klauer, 2010; Voß, 2013; Wilbert & Linnemann, 2011), oder wie dieser diagnostische Ansatz Einzug in die Schule halten kann (Gold, 2015; Križan & Vossen, 2016; Voß, Marten, et al., 2016).

Die Ausführungen zeigen, dass curriculumbasiertes Messen in den Vereinigten Staaten bereits weit verbreitet ist. In den letzten Jahren sind auch in Deutschland einige standardisierte Verfahren der Lernverlaufsdiagnostik entwickelt worden (Inventar zur Erfassung der Lesekompetenz im 1. Schuljahr: Diehl & Hartke, 2012; Lernverlaufsdiagnostik Mathematik: Strathmann & Klauer, 2012; Lernfortschrittsdiagnostik Lesen: Walter, 2009, Verlaufsdiagnostik sinnerfassenden Lesens: 2014b). Ein Leitfaden für Lehrkräfte zur Entwicklung von CBM im Bereich Lesen stammt dabei von Käter und Kollegen (2016).

Lernverlaufsdiagnostik in Form standardisierter Verfahren per se führt nicht zu einer Leistungsverbesserung der Schülerinnen und Schüler. „Spekuliert der Fragesteller darauf, dass sich möglicherweise allein durch die Applikation des Messverfahrens Leistungsverbesserungen einstellen, so muss die Frage verneint werden“ (Walter, 2009, S. 163). Im RTI-Konzept geht es nicht nur um das reine Sammeln von Informationen über Lernverläufe (siehe 4.2.2.4), sondern vielmehr darum, Lernverläufe auszuwerten und für jedes einzelne Kind Schlüsse daraus zu ziehen und systematisch zu reagieren und individuelles Feedback zu geben.

4.2.3 Wirksamkeit

Befunde zeigen, dass die *einzelnen* Bestandteile des Response to Intervention-Konzepts eine hohe Relevanz für die Lernentwicklung von Schülerinnen und Schülern haben. Folgende Ausführungen verdeutlichen dies:

- *Curriculumbasiertes Messen* kann schon alleine zu Leistungsverbesserungen der Schülerinnen und Schüler führen, sofern die Daten grafisch aufgearbeitet, nach einem bestimmten System interpretiert, als Rückmeldung an die Schüler gegeben und bestimmte Fördermaßnahmen von negativen Entwicklungstrends abgeleitet werden (Diehl & Hartke, 2007; Diehl, Hartke & Knopp, 2009; Fuchs & Fuchs, 1990; Graney & Shinn, 2005). „Frequent progress monitoring alone (i.e. without instructional modifications) did not appear to boost student achievement“ (Stecker et al., 2005, S. 815).

- Auch die Bedeutung *früher Prävention* nimmt einen hohen Stellenwert in der pädagogischen Praxis ein (Barth, 2010; Breuer & Weuffen, 2000; Tröster, 2008). Längsschnittstudien zeigen, wie gravierend frühe Lernschwierigkeiten für die spätere Entwicklung sein können (Krajewski & Schneider, 2006; Moser Opitz, 2005; Stern, 1997, 2008). So gibt es zahlreiche Forschungsbemühungen zur Entwicklung von präventiven Förderkonzepten für verschiedene (vor-)schulische Domänen (Mit Baldur Ordnen, Zählen, Messen: Clausen-Suhr, 2009; Mengen, zählen, Zahlen: Krajewski et al., 2013). Es stellt sich jedoch die Frage, ob das *Gesamte* (hier: das Response to Intervention-Modell) mit der Zusammenführung der einzelnen Bestandteile letztlich mehr sein kann als die Summe dieser Teile. Brown-Chidsey und Steege (2010, S. 4) schreiben: „RTI did not just ‚fall out of the sky’. Rather, it represents the culmination of many years of effort to provide effective education for all students.“

Es gibt bisher nur wenige Untersuchungen zur Wirksamkeit des RTI-Gesamtkonzepts. In seiner Megaanalyse ermittelt Hattie (2014, S. 276) eine große Effektstärke von $d = 1{,}07$ (zur Klassifikation von Cohens d siehe Kapitel 4.3.1). Burns, Appleton und Stehouwer (2005) vergleichen in ihrer Metaanalyse 31 Studien, in denen das RTI-Konzept umgesetzt wird. 24 Studien geben dabei eine Effektstärke an. Insgesamt ergibt die Implementierung des RTI-Modells eine durchschnittliche Effektstärke von $d = 1.49$ (SD: 1.43) mit einem Range von $d = .18$ bis $d = 6.71$. In acht Studien wird dabei das Modell von universitärer Seite implementiert, was eine Effektstärke von $d = 1.04$ (SD: .88) ergibt. Die Implementierung von schulischer Seite (16 Studien) zeigt dabei einen höheren Effekt von $d = 1.38$ (SD: .98). Bezogen auf die Schülerleistungen (student outcome; 11 Studien) liegt die Effektstärke bei $d = .96$ (SD: .77) und bezogen auf systematische Effekte (u. a. Anzahl an Unterstützungsbedarfen) (systematic outcome; 13 Studien) bei $d = 1.53$

(SD: 1.28). Insgesamt zeigt die Metaanalyse, dass sich alle Effektstärken in einem hohen Bereich ($d \geq 0{,}8$) befinden (Cohen, 1988).

Die Befunde zeigen insgesamt, dass die Umsetzung des Response to Intervention-Konzepts mit hohen Effektstärken einhergehen kann, was grundsätzlich für eine positive Wirksamkeit der Maßnahme spricht. Auch in Deutschland gibt es Bemühungen, das Response to Intervention-Modell zu implementieren und die Wirksamkeit zu untersuchen.

In Deutschland findet erstmals im Jahr 2010 eine großflächige Umsetzung des Response to Intervention-Modells im Rahmen des „Rügener Inklusionsmodells“ (RIM; www.rim.uni-rostock.de) unter der Leitung von Prof. Dr. Bodo Hartke und Prof. Dr. Kirsten Diehl statt. Ungefähr seit 2010 gehört der Ansatz „in der Wissenschaft, aber auch in der Kultusbürokratie und in Praktikerkreisen [...] zu den meist diskutierten“ (Kuhl & Hecht, 2014, S. 409). Insbesondere das RIM rückt das Response to Intervention-Modell in das Zentrum der deutschsprachigen Diskussion. Die Laufzeit des Projekts ist von 2010 bis 2014. Im RIM wird das RTI-Modell in allen staatlichen Grundschulen auf der Insel Rügen implementiert. Dies geht u. a. mit umfassenden Fortbildungsmaßnahmen für die Lehrkräfte einher. Als Kontrollgruppe fungieren Schülerinnen und Schüler aus Stralsund ($n_{Rügen} = 441$; $n_{Stralsund} = 385$). Im Jahr 2016 veröffentlicht die Projektgruppe ein Buch zu den abschließenden Befunden (Voß, Blumenthal, et al., 2016), was auch im März 2016 mit einem eigenen Themenheft in der Zeitschrift für Heilpädagogik einhergeht.

Zusammenfassend lässt sich festhalten, dass die Schülerinnen und Schüler auf Rügen nach Abschluss des Projekts insgesamt eine geringere Quote an formal festgestellten Unterstützungsbedarfen aufweisen ($Ø_{Rügen} = 3{,}8\ \%$; $Ø_{Stralsund} = 12{,}3\ \%$), über eine deutlich bessere emotional-soziale Schulsituation verfügen und dort diejenigen mit einem Unterstützungsbedarf im Schwerpunkt Lernen nach drei Jahren schulische Leistungen (Lesen, Schreiben, Rechnen) erzielen, die diejenigen in der Kontrollgruppe erst nach vier Jahren erreichen. Im Rechnen und Rechtschreiben erzielen hingegen Schülerinnen und Schüler ohne Unterstützungsbedarfe in Stralsund bessere Werte als diejenigen der Insel Rügen. Im Lesevermögen gibt es keine Unterschiede. Auch im Bereich Sprache zeigen beide Gruppen vergleichbare Fördererfolge (Voß, Marten, et al., 2016). Die Autorengruppe schreibt zusammenfassend:

> „Betrachtet man die Schulleistungen als den entscheidenden Indikator für die Bewertung eines Schulsystems, geht der Vergleich zwischen dem RIM und der traditionellen Arbeitsweise bezogen auf die Gesamtgruppenergebnisse unentschieden aus. Leichte Vorteile bei den Zwillingsgruppenvergleichen des traditionellen Systems erklären sich durch schulstrukturelle und sozioökonomische regionale Unterschiede und nicht durch die differierenden Beschulungskonzepte. Im Hinblick auf die Förderung von Schülern mit sehr ungünstigen Voraussetzungen für schulisches Lernen überzeugt das Rügener Inklusionsmodell durch deutlich bessere Lernerfolge betroffener Schüler.“ (Voß, Marten, et al., 2016, S. 147)

Neben dem RIM-Projekt auf der Insel Rügen gibt es in Deutschland weitere Projekte, in denen das RTI-Konzept als Gesamtkonzept erprobt wird. In Begleitung durch die Universität Wuppertal (Prof. Dr. Christian Huber) und der Universität Köln (Prof. Dr. Christian Rietz) wird im Jahr 2012 in einer Schule in Leverkusen und im Jahr 2013 in einer Schule im Kreis Viersen das RTI-Konzept implementiert. Dabei wird das Projekt auf den Bereich Lesen beginnend mit der Klasse eins beschränkt. Das Umsetzungskonzept und der Untersuchungsplan sind bereits veröffentlicht (Huber et al., 2013); Ergebnisse stehen noch aus.

Eine weitere Umsetzung des RTI-Konzepts erfolgt 2013 in Kooperation der Justus-Liebig-Universität Gießen (Prof. Dr. Jan Kuhl) mit zwei Beratungs- und Förderzentren sowie mehreren Grundschulen aus dem Marburger Umland (Kuhl & Hecht, 2014). Zu Beginn des Projekts werden Teams aus einer Förderschullehrkraft und mehreren Grundschullehrkräften gebildet. Sie absolvieren Weiterbildungen zum Thema Diagnostik (nur die Förderschullehrkräfte) und Förderung (sowohl Förderschul- als auch Grundschullehrkräfte). Zu Beginn und in der Mitte des ersten Schuljahres durchlaufen alle Schülerinnen und Schüler der ersten Klasse Testungen im Bereich Mathematik (DEMAT-1 & MBK-1) und im Bereich Schriftsprache (HSP & Elfe 1–6). Auf Grundlage der Ergebnisse werden dann die schwachen Schülerinnen und Schüler auf der selektiven Ebene (II) für eine Förderung ausgewählt und mit einer Verlaufsdiagnostik (IEL-1) begleitet. Zur Förderung werden möglichst evidenzbasierte Verfahren eingesetzt (u. a. Hören, lauschen, lernen I & II; Kieler Leseaufbau; Mengen, zählen, Zahlen; Lesen und schreiben lernen mit der Hexe Susi). Auch für dieses Projekt sind bisher lediglich die Planungen veröffentlicht; Ergebnisse stehen auch hier noch aus (Kuhl & Hecht, 2014).

4.2.4 Kritik

Grundsätzlich sind die Meinungen zum RTI-Konzept in Deutschland sehr divergierend, wobei das Konzept „in letzter Zeit verstärkt befürwortet [wird]“ (Križan & Vossen, 2016, S. 85). So schreibt Walter (2008b, S. 213), dass der Ansatz „Merkmale [besitzt], die für eine gute Förderpädagogik unverzichtbar und ohne echte Alternative sind. Die deutsche Sonderpädagogik wäre gut beraten, sich auf diesem Feld stärker zu engagieren“. Hingegen gibt es auch kritische Stimmen gegenüber dem RTI-Ansatz. Hinz (2013) kritisiert, dass der RTI-Ansatz zu technokratisch sei. Der Autor schreibt:

> „Bei dem RTI-Modell scheint ein deutlich behavioristischer Zugang durch, der von einer linearen Vorstellung des Lernens ausgeht, eine massive Verstärkung der Kontrolle von Leistungsentwicklungen anstrebt und offenbar dem Glauben folgt, sie fern jeglicher Bezüge zur Lebenswelt durch Intervention zur Normalentwicklung bringen zu können.“

Weitere Kritik richtet sich auch gegen die schwierige Messbarkeit bestimmter Outcome-Variablen (wie zum Beispiel moralische Urteilsfähigkeit, politische Orientierung), deren Vermittlung jedoch in den Bildungsplänen fest verankert ist. Auch die Uneinheitlichkeit der Evaluationsstandards gerät in Kritik (Hinz, 2013). Dieser Aspekt der Programmevaluation ist in den bisherigen Ausführungen bereits angesprochen – derzeit besteht de facto noch Uneinigkeit darüber, welche Evaluationsschritte gegangen werden müssen, um eine Maßnahme oder ein Konzept „evidenzbasiert“ nennen zu dürfen (Blumenthal & Mahlau, 2015; Brown-Chidsey & Steege, 2010; Freeman & Sugai, 2013; Kuhl & Euker, 2016). Jedoch handelt es sich bei dem Konzept der Evidenzbasierung und somit auch beim RTI-Modell um noch junge Konzepte, für die – nach intensiver wissenschaftlicher Diskussion – Evaluations- und Implementationsmöglichkeiten im deutschen Schulsystem erarbeitet werden sollten. Als Vorbild könnten auch hier die USA dienen, in der verschiedene Bundesstaaten schon derartige Implementationsleitfäden entwickelt haben (z. B. South Dakota: South Dakota Department of Education, 2012; Tennessee: Tennessee Department of Education, 2013). „Minimalkonsens“ ist, dass Evidenzbasierung stets ein Zusammenspiel verschiedener Expertisen ist und nicht rein auf wissen-

schaftlichen Erkenntnissen beruhen darf (Blumenthal & Mahlau, 2015; Whitehurst, 2002). Dieses Zusammenspiel garantiert die pädagogische Handlungsfähigkeit. Fatal wäre es, aus den genannten Gründen zu fordern, Evidenzbasierung gänzlich aus der Diskussion um Gestaltung inklusiver Bildung auszuschließen. Die Frage nach effektiver inklusiver Bildung sollte im Kern keine reine Systemfrage sein. „Statt der Strukturfragen sind [...] für die Qualität von Schule offenbar die Formen des Lehrens und Lernens, der Stil pädagogischer Arbeit, konkrete Förderprogramme, zum Beispiel in der Beherrschung der deutschen Sprache [...] für die ‚Risikogruppe' am ehesten geeignet, Abhilfe zu schaffen, die Qualität zu steigern und Bildungskarrieren offener und gerechter zu machen" (Ahrbeck, 2014, S. 107). Die gesellschaftliche Relevanz hochwertiger inklusiver Bildung wird mittlerweile auch von staatlicher Seite erkannt, was zum Beispiel in einer umfassenden Ausschreibung zur Unterstützung von Forschungsvorhaben zur Qualifizierung pädagogischer Fachkräfte für die Inklusion sichtbar wird (Bundesministerium für Bildung und Forschung, 2016). Der Forderung nach hoher Qualität in inklusiver Bildung kann durch das Response to Intervention-Modell begegnet werden.

Das folgende Kapitel 4.3 fokussiert in Fragen evidenzbasierter Förderung den Bereich mathematischer Basiskompetenzen in der Sekundarstufe I.

4.3 Evidenzbasierte Förderung mathematischer Basiskompetenzen in der Sek. I

> „We have all seen students with their heads on their desks, rubbing their foreheads, or avoiding eye contact to avoid being called on – or even acting out in other ways when they can't access the mathematical ideas. We have also seen the common mistakes that recur year after year as students try, but often fail, to remember the arithmetic rules they have been taught."
> (Humphreys & Parker, 2015, S. 5)

In Fragen schulischer Förderung zeigt die aktuelle Diskussion, dass der Einsatz evidenzbasierter Fördermethoden und -maßnahmen erfolgsversprechender ist als der Einsatz derjenigen Verfahren, die ausschließlich auf anekdotischer Evidenz oder subjektivem Erfolgsversprechen beruhen (Brown-Chidsey & Steege, 2010, S. 43; Gold, 2011; Lembke & Stormont, 2005) (siehe Kapitel 4.1 und 4.2). In diesem Kapitel wird der aktuelle Forschungsstand zu evidenzbasierten Maßnahmen und Methoden zur Förderung mathematischer Kompetenzen, die sich für den Einsatz in der Sekundarstufe I eignen, aufgearbeitet. Es wird dabei zwischen standardisierten Förderkonzepten und Fördermethoden (Koch, 2008; Petermann, 2015a) unterschieden. Vorerst wird ein Überblick über die Strategie zur Auswertung des Forschungsstandes gegeben.

4.3.1 Hinweise zur Auswertung des Forschungsstandes

Welche Methoden und Maßnahmen bieten sich zur Förderung mathematischer Basiskompetenzen in der Sekundarstufe I an? Die Auswertung des Forschungsstandes zu eben dieser Frage erfolgt eng gekoppelt an die empfohlene Schrittfolge zur Identifikation und Analyse evidenzbasierter Verfahren nach Freeman und Sugai (2013), die in Kapitel 4.1 bereits ausführlich vorgestellt ist.

Das Analyseverfahren wird getrennt für standardisierte deutschsprachige Förder*programme* und Förder*methoden* (Koch, 2008) durchgeführt. Es wird bei den Methoden sowohl der deutschsprachige als auch der internationale Forschungskontext berücksichtigt.

Um die Schritte nach Freeman und Sugai (2013) gründlich zu durchlaufen, werden die Unterkapitel 4.3.2 (Deutschsprachige standardisierte Förderprogramme) und Unterkapitel 4.3.3 (Fördermethoden) wie folgt gegliedert und aufgeschlüsselt:

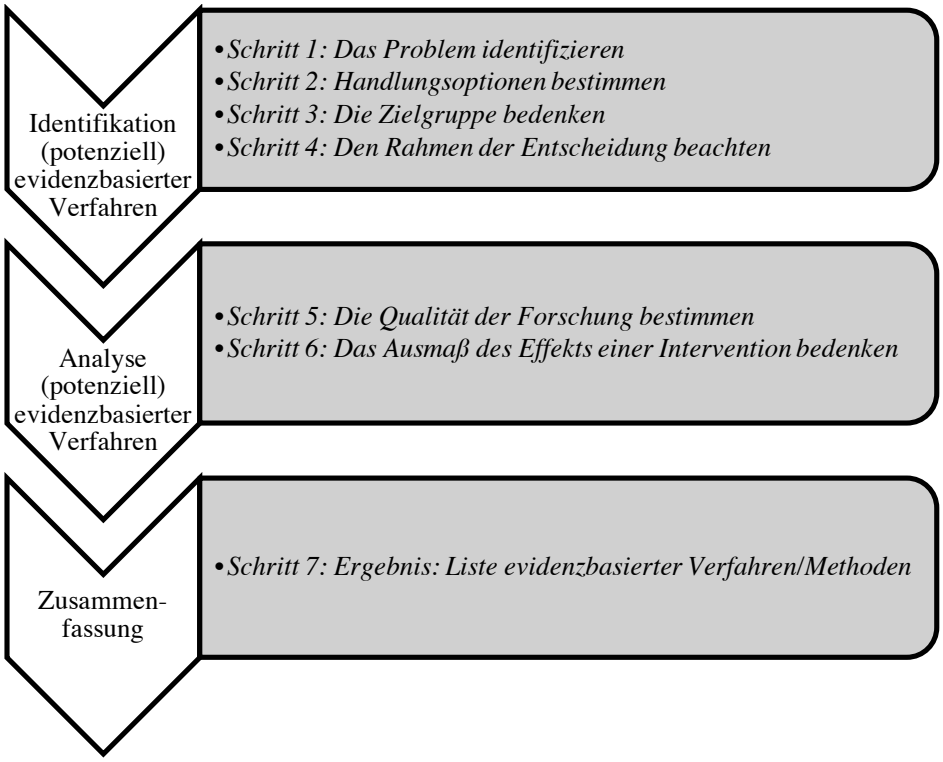


ABBILDUNG 29 VORGEHEN UND GLIEDERUNG DER KAPITEL 4.3.2 UND 4.3.3 ZUR AUSWERTUNG DES FORSCHUNGSSTANDES

Bei der *Identifikation potenziell evidenzbasierter Verfahren* werden nach der Problemdarstellung (Schritt 1) sowohl die Recherchestrategie als auch deren Ergebnisse dargestellt (Schritt 2). Weiterhin wird berücksichtigt, ob sich die identifizierten Fördermaßnahmen bzw. -methoden für den Einsatz im Kontext der Problemdarstellung eignen (Schritt 3 und 4). Da sich ein Großteil der bisherigen Forschung auf den Primarbereich bezieht, wird berücksichtigt, ob auch der Einsatz in der Sekundarstufe I möglich und sinnvoll ist. Im Schritt der *Analyse potenziell evidenzbasierter Verfahren* wird weitere fachrelevante Literatur hinzugezogen, durch die die Qualität des Förderprogramms oder der Methode eingeschätzt werden kann (Schritt 5). In Schritt 6 wird abschließend eine möglichst konkrete Einordnung in die Klassifikation des Councils for Exceptional Children (2014, S. 8 ff.) angestrebt (siehe Kapitel 4.1): 1. Evidence-based Practice (Evidenzbasierte Praxis), 2. Potentially evidence-based practice (Potentielle evidenzbasierte Praxis), 3. Mixed Evidence (Gemischte Evidenz), 4. Insufficient Evidence (Ungenügende Evidenz), 5. Negative effects (Negative Effekte). Bei Schritt 5 und Schritt 6 ist die Effektstärke *d* nach Cohen (1988) ein zentrales Maß. Die Effektstärke *d* gibt an, wie groß ein beobachteter Effekt zwischen zwei untersuchten Gruppen ausfällt. Dabei gilt nach Cohen (Cohen, 1988, S. 24 ff.):

$d \approx .20$ = kleine Effektstärke,

$d \approx .50$ = mittelgroße Effektstärke,

$d \approx .80$ = große Effektstärke.

Die in Metaanalysen angegeben Effektstärke ist oftmals „korrigiert“ (oder synonym: gewichtet, gemittelt), sodass dabei unterschiedliche Gruppengrößen und Vortestunterschiede herausgerechnet werden. In den Schritten „Identifikation“ und „Analyse“ wird von *potenziell* evidenzbasierten Verfahren gesprochen, da der Grad an Evidenzbasierung erst am Ende der Analyse zuverlässig eingeschätzt werden kann.

Im Punkt *Zusammenfassung* (Schritt 7) erfolgen eine Synthese und die Auswertung der Rechercheergebnisse unter Aufführung von Forschungsdesiderata.

Die Schrittfolge nach Freeman und Sugai (2013), die in den Kapitel 4.3.2 und 4.3.3 durchlaufen wird, ist für praktizierende Lehrkräfte entwickelt worden, sodass sich die „Identifizierung des Problems“ (Schritt 1) im Regelfall auf einen Schüler oder eine Schülerin im alltäglichen Unterrichtsgeschehen bezieht. Im Folgenden wird diese Schrittfolge jedoch aus wissenschaftlicher Perspektive durchlaufen in der Frage nach der Identifikation evidenzbasierter Maßnahmen oder Methoden zur Förderung mathematischer Kompetenzen in der Sekundarstufe I. Die Orientierung an dem schrittweisen Vorgehen ermöglicht eine gründliche Recherche, an deren Ende zuverlässige Hinweise auf evidenzbasierte Methoden und Maßnahmen stehen, die konzeptionell verwendet und verarbeitet werden können.

4.3.2 Deutschsprachige standardisierte Förderprogramme

In Deutschland sind bereits standardisierte Programme zur Förderung mathematischer Kompetenzen veröffentlicht. Standardisierte Förderprogramme zeichnen sich, im Gegensatz zu sog. alltagsintegrierten Methoden, durch eine explizite entwicklungs- und lernpsychologische theoretische Fundierung und ein differenziert ausgearbeitetes Manual mit detaillierten Handlungsanweisungen aus (Petermann, 2015a, S. 161). Petermann (2015a, S. 162 f.) führt als Merkmale standardisierter Förderprogrammen auf:

- Ein hoher Strukturierungsgrad der Manuale,
- eine große Spezifität hinsichtlich der Förderinhalte,
- eine große Transparenz der Fördermaßnahmen,
- Einbeziehung diagnostischer Prozeduren im Vorfeld,
- Schritte der Qualitätsprüfung (Erfolgskontrolle),
- Möglichkeit der Kombination mit anderen (therapeutischen) Maßnahmen und
- (ergänzt durch d. Verf.): Einhaltung der Manualtreue.

Da die Entwickler in der Umsetzung eines Förderprogramms durch Lehrkräfte eine sog. Manualtreue erwarten, ist die Kreativität und Flexibilität in der Umsetzung in einem gewissen Maße eingeschränkt. Doch kann sich nur bei einer gewissenhaften und genauen Umsetzung des Programms die Wahrscheinlichkeit für die Wirksamkeit einstellen, die in den Evaluationsstudien ermittelt wird.

Der Begriff „Förderprogramm“ wird in der Literatur nicht konsequent für Konzepte nach den o. a. Kriterien nach Petermann (2015b) verwendet. So fassen Fingerle und Ellinger (2008) in ihrem Buch „Sonderpädagogische Förderprogramme im Vergleich“ auch den Offenen Unterricht oder die Portfolio-Arbeit unter diesen Begriff, womit sie *nicht* den o. a. Kriterien

i. e. S. entsprechen. Im Folgenden werden unter Förderprogrammen diejenigen verstanden, die den Kriterien nach Petermann (2015a) entsprechen.

4.3.2.1 Identifikation (potentiell) evidenzbasierter standardisierter Programme

SCHRITT 1: DAS PROBLEM IDENTIFIZIEREN

Die umfassende Analyse der Mathematikleistungen von Sekundarstufenschülerinnen und -schülern zeigt, dass ein nicht zu missachtender Teil der Schülerinnen und Schüler in der weiterführenden Schule massive Defizite im Beherrschen der mathematischen Basiskompetenzen, die basalen arithmetischen Operationen, aufweist (vgl. u. a. Brown et al., 2003; Ehlert et al., 2013; Galton et al., 2002; Humbach, 2008) (siehe Kapitel 2.4.2 und 2.4.3). Dabei zeigen u. a. die Bildungsstandards für die Primarstufe, dass diese basalen Kompetenzen von „fundamentaler Bedeutung" (Kultusministerkonferenz, 2004, S. 8) für den weitere Kompetenzerwerb in der Sekundarstufe I sind (siehe Kapitel 2.3.3). Diese Schülerinnen und Schüler benötigen eine Förderung in eben diesem Bereich, um den inhaltlichen Anschluss an die weiterführende Mathematik zu erlangen. Eine Möglichkeit der Förderung stellen fundierte standardisierte Förderprogramme dar (Schneider et al., 2013, S. 233). Im folgenden Kapitel wird erarbeitet, welche deutschsprachigen Förderprogramme derzeit zur Förderung der hier definierten mathematischer Basiskompetenzen in der Sekundarstufe I vorliegen, in welcher Evidenzstufe des Councils for Exceptional Children (2014) sie eingeordnet werden können und ob der Einsatz im Kontext der identifizierten Problemlage realisierbar ist.

SCHRITT 2: HANDLUNGSOPTIONEN BESTIMMEN

Die Recherche standardisierter Förderprogramme erfolgt kriteriengeleitet. Gesucht werden:

- Standardisierte Programmme zur Förderung der hier definierten mathematischen Basiskompetenzen, die
- für den Einsatz ab Mitte/Ende der Primar- und/oder in der Sekundarstufe I geeignet sind,
- in deutscher Sprache vorliegen und
- nicht vor dem Jahr 2000 veröffentlicht sind.

Zur Identifikation potenziell evidenzbasierter Förderprogramme werden fachrelevante Überblicksarbeiten recherchiert (Fingerle & Ellinger, 2008; Fritz & Ricken, 2008; Schneider & Hasselhorn, 2016; Schneider et al., 2013) und die Testzentrale des Hogrefe-Verlags (www.testzentrale.de; Filter: „Kinder und Jugendliche") genutzt. Zudem werden Hinweise auf Tagungen und Kongressen verfolgt und eine allgemeine Recherche nach dem Schneeballprinzip im Internet durchgeführt (z. B. www.fis-bildung.de; www.plus.orbis.uni-oldenburg.de; www.pubpsych.zpid.de; www.eric.ed.gov).

Ausgeschlossen werden in der Recherche die zahlreichen Programme für das Vorschulalter (z. B. Mit Baldur ordnen, zählen, messen: Clausen-Suhr, 2009; Mengen, zählen, Zahlen: Krajewski et al., 2013) oder auch Programme für spezifische Zielgruppen wie zum Beispiel Schülerinnen und Schüler im Förderschwerpunkt Geistige Entwicklung (z. B. DIFMaB: De Vries, 2013) (vgl. auch Hellmich, 2007). Hingegen können sich Programme für den mittleren und späten Primarbereich aufgrund der mathematischen Kompetenzdefizite der Schülerinnen

und Schüler in der Sekundarstufe I für einen Einsatz anbieten, sodass sie in der Recherche Berücksichtigung finden.

Tabelle 20 gibt einen Überblick über die Rechercheergebnisse.

TABELLE 20 POTENZIELL EVIDENZBASIERTE DEUTSCHSPRACHIGE STANDARDISIERTE FÖRDERPROGRAMME IM BEREICH MATHEMATIK FÜR DIE SEKUNDARSTUFE I

Programm		Autoren (Jahr)	Einsatz gemäß RTI	Zielgruppe	Inhalte	Umfang
Mathe sicher können	Band 1: Natürliche Zahlen	Selter, Prediger, Nührenbörger & Hußmann (2014)	Stufe 2 fokussiert & Stufe 3 indiziert	Schülerinnen und Schüler mit Rechenschwierigkeiten in der Jahrgangsstufe 5 und 6	Zahlverständnis, Operationsverständnis, Zahlenrechnen, Ziffernrechnen	k. A. (je nach individuellem Bedarf und Fortschritt)
	Band 2: Brüche, Prozente, Dezimalzahlen	Prediger, Selter, Hußmann & Nührenbörger (2014)	Stufe 2 fokussiert & Stufe 3 indiziert	Schülerinnen und Schüler mit Rechenschwierigkeiten in der Jahrgangsstufe 5 bis 7	Bruchverständnis, Rechnen mit Brüchen, Dezimalverständnis, Rechnen mit Dezimalzahlen, Zusammenhang von Dezimalzahlen und Brüchen	k. A. (je nach individuellem Bedarf und Fortschritt)
Kalkulie I, II, III		Fritz, Ricken, Gerlach & Schmidt (2007)	Stufe 2 fokussiert & Stufe 3 indiziert	Schülerinnen und Schüler mit Rechenschwierigkeiten in der Klasse 1 (Teil I), der Klasse 2 (Teil II) und der Klasse 3 (Teil III)	Kardinalität, Teil-Teil-Ganzes-Beziehungen (Teil 1), der Zwanzigerraum & Übergang vom Auszählen zum Rechnen (Teil 2), Nicht-zählende Rechenstrategien (Teil 3)	k. A. (je nach individuellem Fortschritt)
Rechenspiele mit Elfe und Mathis II		Lenhard, Lenhard & Lingel (2010)	Stufe 2 fokussiert & Stufe 3 indiziert	Schülerinnen und Schüler mit Rechenschwierigkeiten in der Jahrgangsstufe 3 bis 5	Geometrie, Rechnen, Sachaufgaben	15–20 Trainingseinheiten à 20–30 Minuten
Calcularis		Gross, Jäncke, Von Aster, Meyer (2015)	Stufe 3 indiziert	Schülerinnen und Schüler mit Rechenschwierigkeiten oder Dyskalkulie	Zahlendarstellung und -verständnis und arithmetische Operationen im Zahlenraum bis 1000	3x wöchentlich über mindestens 4 Monate

Programm	Autoren (Jahr)	Einsatz gemäß RTI	Zielgruppe	Inhalte	Umfang
Dortmunder Zahlbegriffstraining	Moog & Schulz (2005)	Stufe 3 indiziert	Förderschülerinnen und –schüler oder Schülerinnen und Schüler der Klassen 1 bis 3 mit Rechenschwierigkeiten oder einer Dyskalkulie	Addition und Subtraktion im Zahlenraum bis 10 (Zählen, Mengen- und Zahlrelationen)	19 Sitzungen
Mathematik- und Rechenkonzepte im Vor- und Grundschulalter – Training (Marko-T)	Gerlach, Fritz & Leutner (2013)	Stufe 3 indiziert	Kinder mit Rechenschwierigkeiten und/oder Entwicklungsverzögerungen im Alter von 5 bis 8 Jahren	Zählzahl, oridnaler Zahlenstrahl Kardinalität und Zerlegbarkeit, Enthaltensein und Klasseninklusion, Relationalität	57 Einheiten
Meister Cody	Kuhn, Holling, Raddatz & Dobel (2016)	Stufe 3 indiziert	Schülerinnen und Schüler der Klassen 2 bis 4 mit Rechenschwierigkeiten oder einer Dyskalkulie	Zahlen-Mengen-Verknüpfung, Teil-Teil-Ganzes-Verhältnis, Zahlenraumvorstellung, Relationalzahlbegriff, Faktenabruf, (Quasi-) Simultanerfassung, Verständnis des Stellenwertsystems, Transkodieren	3x wöchentlich über 30 Minuten, mindestens 10 Wochen

Schritt 3: Die Zielgruppe bedenken und

Schritt 4: Den Rahmen der Entscheidung beachten

Bevor die identifizierten Programme bezüglich ihrer Wirksamkeit analysiert werden, wird ihre Passung im Kontext der Problemstellung begutachtet. Gemäß ihren Konzeptionen eignen sich alle identifizierten Programme für den Einsatz ab der Mitte bzw. dem Ende der Grundschule oder in der Sekundarstufe I und berücksichtigen, zumindest in Aspekten, den Bereich der mathematischen Basiskompetenzen.

Ein Blick ins Detail zeigt, dass das „Dortmunder Zahlbegriffstraining“ (Moog & Schulz, 2005) nur den Zahlenraum bis 10 umfasst und sich das „Marko-Training“ (Gerlach et al., 2013) sowie das Programm „Meister Cody“ (Kuhn et al., 2016) im Schwerpunkt mit den sog. Vorläuferfertigkeiten befassen (siehe Kapitel 2.1). Diese drei Programme berücksichtigen folglich nicht die hier definierten mathematischen Basiskompetenzen und passen somit nicht zur hier diskutierten Problemstellung. Sie werden in der Analyse im Folgenden nicht weiter berücksichtigt.

Erwähnenswert sind darüber hinaus weitere recherchierte Interventions- und Förderkonzepte, die jedoch nicht den Kriterien nach Petermann (2015a) genügen, da sie zum Beispiel keine exakte Manualtreue erwarten. Dazu zählen zum Beispiel die Reihe „Produktives Lernen für Kinder mit Lernschwächen“ (Scherer, 2002) und die beiden Bände „Handbuch produktiver Rechenübungen“ (Wittmann & Müller, 1990). Diese sind zwar keine Förderprogramme i. e. S., können jedoch als unterstützender Aufgaben- und Ideenpool dienen.

4.3.2.2 Analyse (potentiell) evidenzbasierter standardisierter Programme

Schritt 5: Die Qualität der Forschung bestimmen und

Schritt 6: Das Ausmaß des Effekts einer Intervention bedenken

Die Tabelle 20 zeigt die in Schritt 2 potenziell evidenzbasierten standardisierten Förderprogramme. Zur Analyse der Programme werden Evaluationsstudien, die aus den Handreichungen der Programme oder aus einer weiteren Datenbankrecherche stammen, hinzugezogen.

Mathe sicher können

Das Förderprogramm „Mathe sicher können“ besteht aus zwei Bestandteilen. Im ersten Baustein „Natürliche Zahlen“ (Selter et al., 2014) für die Jahrgänge fünf und sechs sind die Inhalte Zahlverständnis, Operationsverständnis, Zahlenrechnen und Ziffernrechnen. Im zweiten Teil „Brüche, Prozente, Dezimalzahlen“ (Prediger, Selter, Hußmann & Nührenbörger, 2014) für die Jahrgänge fünf bis sieben werden die Themen Bruchverständnis, Rechnen mit Brüchen, Dezimalverständnis, Rechnen mit Dezimalzahlen sowie der Zusammenhang von Dezimalzahlen und Brüchen thematisiert. Das Förderprogramm ermöglicht durch Diagnoseaufgaben die Erhebung der Lernausgangslage, um passende Fördereinheiten anschließen zu können. Es können auch vereinzelt Bausteine thematisiert werden. Dementsprechend gibt es keine Angaben zum Umfang und Dauer des Förderkonzepts. Es wird empfohlen, die Förderung begleitet von einer

Lehrkraft in Kleingruppen durchzuführen. „Mathe sicher können“ verfolgt die drei didaktischen Leitlinien „Diagnosegeleitetheit“, „Verstehensorientierung“ und „Kommunikationsförderung“. Eine Materialbox zu jedem Förderbaustein ermöglicht den handlungsorientierten Aufbau der mathematischen Verstehensgrundlage. Das Programm setzt mit dem Baustein 1 bei den Inhalten der dritten und vierten Klasse der Grundschule an. Falls die Rückstände massiver sind, empfehlen die Autoren den Einsatz anderer Förderkonzepte (Selter et al., 2014, S. 7). Die konzeptionelle Vorarbeit des Förderprogramms leistet Okka Freesemann (2014) im Projekt SimBA (Sicher im mathematischen Basisstoff). Freesemann (2014) fördert dabei über 14 Wochen mit 90 Minuten wöchentlich 539 Schülerinnen und Schüler am Anfang der Sekundarstufe I in Kleingruppen in den Bereichen Operationsverständnis, Dezimalsystem und Zahlenstrahl. Ihre Arbeit gilt als Vorarbeit zum Programm „Mathe sicher können“.

WIRKSAMKEIT

Zur Überprüfung der Wirksamkeit des Förderprogramms „Mathe sicher können“ wird aktuell eine Gruppenvergleichsstudie umgesetzt. Die Untersuchung wird extern vom IPN (Leibniz-Institut für die Pädagogik der Naturwissenschaft und Mathematik) unter Leitung von Prof. Köller durchgeführt. Die Interventionsgruppe besteht aus 4.400 und die Vergleichsgruppe aus 2.100 Schülerinnen und Schülern. Die erste Testung erfolgt am Anfang des Schuljahrs 2015/2016, die zweite Testung am Anfang des Schuljahres 2016/2017 und die letzte Testung am Ende des Schuljahres 2016/2017. Ergebnisse bezüglich der Wirksamkeit stehen derzeit folglich noch aus.

FAZIT

Da für das Programm „Mathe sicher können“ bisher keine Ergebnisse veröffentlicht sind, kann keine konkrete Einordnung in die Stufen der Evidenzbasierung des dem Council for Exceptional Children (2014) vorgenommen werden.

Das Fazit der Wirksamkeitsanalyse nach dem Council for Exceptional Children (2014) für das Programm „Mathe sicher können“ lautet: **Keine Einstufung möglich**

KALKULIE I, II, III

„Kalkulie“ (Gerlach et al., 2007) ist ein Diagnose- und Trainingsprogramm für rechenschwache Schülerinnen und Schüler der ersten bis dritten Klasse, bei denen eine Wahrscheinlichkeit für die Ausbildung einer Rechenstörung vorliegt. Theoretische Grundlage ist dabei das Entwicklungsmodell nach Fritz und Ricken (2008) (siehe Kapitel 2.2.1). Auf Basis der eingebauten Diagnostik wird ersichtlich, welche Förderung bei dem Kind durch welchen Baustein (I, II oder III) angezeigt ist. Der Baustein I orientiert sich an den Inhalten der ersten Klasse und umfasst die Aspekte Reihen bilden, Mengenaspekte und Zahlenwissen. Baustein II orientiert sich an dem Curriculum der zweiten Klasse und umfasst die Bereiche Strukturen erkennen, Strukturen geschickt nutzen und flexibilisieren. Der Baustein III, der auf den Inhalten der dritten Klasse aufbaut, umfasst nicht-zählende Rechenstrategien. „Kalkulie“ kann sowohl in Kleingruppen als auch in Einzelarbeit angewendet werden. Die Dauer der Förderung richtet sich nach den individuellen Fortschritten der Schülerinnen und Schüler.

WIRKSAMKEIT

Im Jahr 2009 liegt laut Autorengruppe noch keine kontrollgruppengestützte Evaluation des Programms vor (Fritz & Ricken, 2009, S. 394). Eine Publikation unter dem Titel „Implementation des Trainingsprogramms ‚Kalkulie' in der Grundschule" in der Zeitschrift Lernen & Lernstörungen ist derzeit im Internet angekündigt.

FAZIT

Da für das Programm „Kalkulie I, II, III" bisher keine Kontrollgruppenvergleichsstudien vorliegen, kann keine konkrete Einordnung in die Stufen der Evidenzbasierung des Council for Exceptional Children (2014) vorgenommen werden.

Das Fazit der Wirksamkeitsanalyse nach dem Council for Exceptional Children (2014) für das Programm „Kalkulie" lautet: **Keine Einstufung möglich**

RECHENSPIELE MIT ELFE UND MATHIS II

Das Programm „Rechenspiele mit Elfe und Mathis II" (Lenhard et al., 2010) ist eine Fortführung des Programms „Rechenspiele mit Elfe und Mathis I". Das Computerprogramm besteht aus 15 Spielen zu den Bereichen Geometrie, Rechnen und Sachaufgaben. Jedes Spiel umfasst drei Niveaustufen. Sobald ein Kind 80 % der Aufgaben eines Spiels auf einem bestimmten Niveau erfolgreich löst, erhält es ein Puzzleteil. Hat es alle Spiele eines Niveaus erfolgreich gelöst, gelangt es auf die nächste Niveaustufe. Die Aufgaben (insgesamt 9.161) entsprechen der Konzeption der Bildungsstandards für den Primarbereich (Kultusministerkonferenz, 2004). Das Förderprogramm bietet die Möglichkeit, diagnostische Erkenntnisse, hier mittels der DEMAT-Reihe, in die Förderung zu implementieren, um das Niveau und die Schwierigkeit der Inhalte passend zu wählen. Das Programm eignet sich für den Einsatz in Kleingruppen oder im Einzelsetting bei Schülerinnen und Schülern zwischen dem zweiten Halbjahr der dritten bis zur Mitte der fünften Jahrgangsstufe. Die Begleitung durch eine Lehrkraft wird empfohlen.

WIRKSAMKEIT

Es liegt eine Studie zur Wirksamkeit der Rechenspiele-Programme vor (Lenhard & Lenhard, 2011). 395 Schülerinnen und Schüler der Jahrgänge 1 bis 4 durchlaufen im Jahr 2009 in einem Zeitraum von acht bis zehn Wochen das für ihre Jahrgangsstufe entsprechende Förderprogramm (I oder II; nicht näher durch die Autoren aufgeschlüsselt). Die Förderung findet einmal in der Woche über den Zeitraum einer Schulstunde statt. Die Größe der Kontrollgruppe wird von den Autoren nicht genannt. Die Schülerinnen und Schüler der Kontrollgruppe nehmen lediglich am regulären Mathematikunterricht teil. Im Rahmen einer Prä- und Posttestung werden die mathematischen Kompetenzen mittels der jahrgangskonformen DEMAT-Reihe erhoben. Es können folgende kleine positiven Effektstärken ermittelt werden: $d = .32$ (Jahrgangsstufe 1), $d = .39$ (Jahrgangsstufe 2), $d = .33$ (Jahrgangsstufe 3) und $d = .19$ (Jahrgangsstufe 4).

Eine weitere Evaluationsstudie liegt ausschließlich für das Programm „Rechenspiele mit Elfe und Mathis I" vor (Lenhard, Lenhard, Schug & Kowalski, 2011). 87 Schülerinnen und Schüler der ersten und zweiten Klasse durchlaufen das Trainingsprogramm über zehn Wochen. Eine

Kontrollgruppe von 116 Schülerinnen und Schüler erhält in der Zeit den regulären Mathematikunterricht. Nach Abschluss der Förderung kann eine Effektstärke von $d = .59$ (1. Klasse) bzw. $d = .62$ (2. Klasse) ermittelt werden.

Das Programm erhält im Jahr 2011 das „Comenius EduMedia-Siegel für exemplarische Bildungsmedien" der Gesellschaft für Pädagogik und Information e. V.

Fazit

Für das Förderprogramm „Rechenspiele mit Elfe und Mathis II" liegt aktuell nur eine Gruppenvergleichsstudie vor, die von positiven Effekten berichtet (Lenhard & Lenhard, 2016a). Kritisch angemerkt wird an dieser Stelle, dass nicht alle zentralen Kennzahlen angegeben sind, wie zum Beispiel die konkrete Größe der Kontrollgruppe. Es wird keine Studie identifiziert, die von negativen Effekten berichtet. Die durch zwei Studien belegte Wirksamkeit des Förderprogramms Elfe und Mathis I spricht für die Wirksamkeit des Konzepts. Lenhard und Lenhard (2011) selbst bezeichnen das Programm als einziges evidenzbasiertes Computerprogramm zur Förderung mathematischer Kompetenzen in Deutschland. Die geringe Anzahl vorliegender Evaluationsstudien für Elfe und Mathis II schwächt jedoch den hier ermittelten Grad der Evidenz.

Das Fazit der Wirksamkeitsanalyse nach dem Council for Exceptional Children (2014) für das Programm „Rechenspiele mit Elfe und Mathis II" lautet: **Potentiell evidenzbasierte Praxis**.

Calcularis

„Calcularis" (Gross et al., 2015) ist ein Computerprogramm für Schülerinnen und Schüler mit Rechenschwierigkeiten oder einer Dyskalkulie. Das Alter der Zielgruppe ist nicht weiter angegeben. Das Programm wird in Zusammenarbeit von Neuropsychologen, Lehrkräften und Informatikern entwickelt. Es soll im Einzelsetting mindestens dreimal wöchentlich über einem Zeitraum von vier Monaten eingesetzt werden. Eine Sitzung dauert ca. 15 bis 20 Minuten. Die geförderten Inhalte sind die Zahlendarstellung und das Zahlenverständnis sowie arithmetische Operationen im Zahlenraum bis 1000. Das Förderprogramm ist dabei adaptiv ausgerichtet, sodass es sich im Niveau und in der Schwierigkeit den Lernständen des Schülers bzw. der Schülerin anpasst. Die Übungen beginnen im Zahlenraum bis zehn, gehen dann weiter in den Zahlenraum bis 20, dann bis 100 und in der höchsten Stufe bis 1000. Das Belohnungssystem erfolgt für jüngere Lerner in Form eines Zoos, für den sie Tiere kaufen können, die dann später gefüttert werden. Für ältere Schülerinnen und Schüler ist das Belohnungssystem durch den virtuellen Erwerb von materiellen Gegenständen angepasst. Zudem ist das Programm mit einer Darstellung des Lernfortschritts versehen, der den Lernenden und Lehrenden (Eltern, Lehrkräfte) zur Verfügung steht.

Wirksamkeit

In einer Pilotstudie (Käser et al., 2013) führen 41 schweizerische Schülerinnen und Schüler mit Rechenschwierigkeiten in den Jahrgängen zwei bis fünf das Programm durch. Die Trainingsgruppe (n = 20) übt fünfmal in der Woche über jeweils 20 Minuten über einen gesamten Zeitraum zwölf Wochen mit „Calcularis". Eine Wartekontrollgruppe (n = 21) startet sechs Wochen

später und führt die Förderung nur noch sechs Wochen durch. Die mathematischen Kompetenzen werden zu Beginn, nach sechs und nach zwölf Wochen u. a. in den Bereichen Addition und Subtraktion mit dem Heidelberger Rechentest und einem Computertest (AC) erhoben.

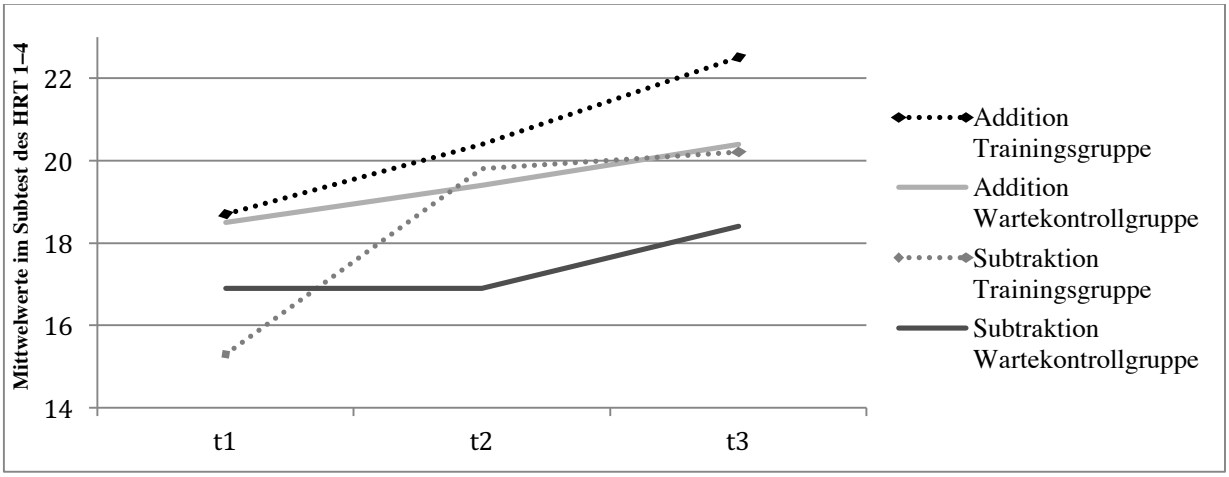


ABBILDUNG 30 ERGEBNISSE DER CALCULARIS-PILOTSTUDIE (KÄSER ET AL., 2013, S. 9)

Die Unterschiede zwischen dem t1- und dem t2-Messzeitpunkt sind in der Trainingsgruppe in beiden Subtests des HRT signifikant, die Effektstärken liegen bei $d = .16$ (Addition) und $d = .52$ (Subtraktion). Zwischen dem t2- und dem t3-Messzeitpunkt kann die Trainingsgruppe nur im Subtest Addition und die Wartekontrollgruppe in keinem Subtest signifikante Verbesserungen erzielen (vgl. Abbildung 30). Im Computertest (AC) sind die Unterschiede zwischen dem t1- und dem t2-Messzeitpunkt in der Trainingsgruppe ebenso in beiden Subtests signifikant, die Effektstärken liegen bei $d = .31$ (Addition) und $d = .39$ (Subtraktion). Zwischen dem t2- und dem t3-Messzeitpunkt kann sowohl die Trainingsgruppe als auch die Wartekontrollgruppe in beiden Subtests signifikante Verbesserungen erzielen (siehe Abbildung 30).

Die Ergebnisse bestätigen die Befunde aus einer vorherigen Parallelstudie (Käser et al., 2012). In Form desselben Designs werden 33 deutsche Schülerinnen und Schüler in einer Trainingsgruppe und 32 in einer Wartekontrollgruppe gefördert. Die Schülerinnen und Schüler in der Trainingsgruppe verbessern sich nach sechs Wochen signifikant, wohingegen die Wartekontrollgruppe bis dahin keine Verbesserung erzielt.

Die vom Bundeministerium für Bildung und Forschung (BMBF) geförderte Hauptstudie läuft insgesamt über sechs Jahre und ist derzeit noch nicht abgeschlossen. Die Studie ist Teil des Projekts „Schulbezogene Umschriebene Entwicklungsstörungen“ (SCHUES). Die Stichprobe besteht aus 135 Schülerinnen und Schülern der ersten bis fünften Klassenstufe. Verglichen wird eine Teilgruppe dieser Schülerinnen und Schüler, die „Calcularis“ an fünf Tagen täglich 20 Minuten über sechs bis zehn Wochen durchführt, mit einer Wartekontrollgruppe und einer dritten Gruppe, die ein computerbasiertes Rechtschreibtraining erhält (Von Aster et al., 2016). Erste Ergebnisse zeigen, dass sich in der Trainingsgruppe sowohl Schülerinnen und Schüler mit und ohne Rechenschwierigkeiten im Kontrast zu beiden Vergleichsgruppen deutlich in ihren Rechenleistungen verbessern. Der größte Effekt liegt auch hier für den Bereich Subtraktion vor. Leider zeigt sich jedoch, dass der Lernzuwachs nach Beendigung des Trainings nicht auf dem erreichten Niveau bleibt und die Leistungen wieder etwas zurückgehen. Ein weiteres Ergebnis ist, dass das Programm offenbar auch zur Reduktion von Mathematikangst führen kann (Von Aster et al., 2016).

FAZIT

Für das Förderprogramm „Calcularis“ liegen aktuell drei Gruppenvergleichsstudien vor, die von positiven Effekten berichten. Zwei Studien verfügen über die benötigte Stichprobengröße zur Einstufung als „Evidenzbasierte Praxis“ (Von Aster et al., 2016; Käser et al., 2012). Es wird keine Studie identifiziert, die von negativen Effekten berichtet. Die durch die zwei Studien belegte Wirksamkeit des Förderprogramms spricht für die Stärke des Programms. Kritisch angemerkt wird an dieser Stelle, dass in den berichteten Studien nicht immer alle Kennzahlen angegeben sind, wie zum Beispiel die konkreten Größen der Teilgruppen.

Das Fazit der Wirksamkeitsanalyse nach dem Council for Exceptional Children (2014) für das Programm „Calcularis“ lautet: **Evidenzbasierte Praxis**.

4.3.2.3 Zusammenfassung und Forschungsdesiderata

SCHRITT 7: ERGEBNIS: LISTE EVIDENZBASIERTER VERFAHREN

Im vorherigen Kapitel sind die – für den Einsatz in der Sekundarstufe I geeigneten – und als zuvor potentiell evidenzbasierten standardisierten Förderprogramme „Mathe sicher können“ (Band 1 und 2), „Kalkulie I, II & III“, „Rechenspiele mit Elfe und Mathis II“ sowie „Calcularis“ eingehend bezüglich ihrer Wirksamkeit entsprechend der Kriterien des Council for Exceptional Children (2014) analysiert worden. Die dafür herangezogenen Studien stammen aus den Manualen der Programme oder werden im Rahmen weiterer Datenbankrecherchen gefunden.

Folgende Einstufungen der Evidenz können zusammenfassend vorgenommen werden:

1. „Mathe sicher können Band 1 und 2“:	keine Einstufung möglich
2. „Kalkulie I, II, III“:	keine Einstufung möglich
3. „Rechenspiele mit Elfe und Mathis II“:	Potentiell evidenzbasierte Praxis
4. „Calcularis“:	Evidenzbasierte Praxis

Insgesamt wird für keines der vier Programme eine ungenügende oder negative Evidenz festgestellt. Aufgrund der unzulänglichen Forschungslage kann für das Programm „Mathe sicher können Band 1 und 2“ und für „Kalkulie I, II, III“ keine Einstufung vorgenommen werden. Hier besteht ein großer Forschungsbedarf, die Wirksamkeit zu überprüfen, um Aussagen über den Grad der Evidenz treffen zu können. Die beiden Computerprogramme „Rechenspiele mit Elfe und Mathis II“ und „Calcularis“ können als potenziell bzw. eindeutig evidenzbasierte Praxis eingeordnet werden. Für die „Rechenspiele mit Elfe und Mathis II“ fehlen weitere randomisierte Gruppenvergleichsstudien, die über die vom Council for Exceptional Children (2014) angeforderte Stichprobengröße (mindestens N = 60) verfügen. Die Überschaubarkeit der Probandinnen und Probanden in den wenigen Evaluationsstudien führt letztlich dazu, dass das Programm nicht in die erste Stufe (Evidenzbasierte Praxis) eingeordnet werden kann. Zudem sollten zukünftige Evaluationsstudien verstärkt lern- und/oder rechenschwache Schülerinnen und Schüler oder diejenigen mit einem Unterstützungsbedarf im Bereich Lernen in den Fokus nehmen, da die Programme schließlich speziell für schwache Rechnerinnen und Rechner konzipiert sind.

Das Fazit dieser Analyse ordnet sich in bisherige Befunde zur Überprüfung der Evidenz sonderpädagogischer Förderprogramme ein. Fingerle und Ellinger (2008, S. 10) schreiben bezüglich ihrer Zusammenschau sonderpädagogischer Förderprogramme: „Nach Lage der Dinge konnten in einzelnen Bereichen nur wenige Ansätze den Kategorien ‚Bewährt' oder ‚Vermutlich effektiv' zugeordnet werden. Dies kann nur konstatiert und als Aufforderung zu weiterer Forschung dargestellt werden". Auch Schneider, Küspert und Krajewski (2013, S. 233, Herv. d. Verf.) formulieren: „Unser Literaturüberblick lässt erkennen, dass bei der Entwicklung von *effektiven vorschulischen* Präventionsprogrammen in neuerer Zeit national wie auch international große Fortschritte verzeichnet werden konnten. Im Vergleich hierzu stellt sich die Situation hinsichtlich *wirksamer schulischer* Interventionsprogramme nicht so günstig dar".

Bezüglich der Förderung mathematischer Kompetenzen in der Sekundarstufe I und der Einordnung der Programme auf den Ebenen des Responsiven Handlungsmodells (Brown-Chidsey & Steege, 2010, S. 5) lässt sich feststellen, dass sich die meisten Programme für den Einsatz auf der fokussierten (2.) und/oder indizierten Stufe (3.) eignen. „Mathe sicher können", „Kalkulie" und die „Rechenspiele mit Elfe und Mathis II" können sowohl in Kleingruppen oder im Einzelsetting durchgeführt werden. „Calcularis", das „Dortmunder Zahlbegriffstraining" und das „Marko-Training" hingegen eignen sich nur für den Einsatz auf der individuellen Ebene. Lediglich das Programm „Mathe sicher können Band 1 und 2" kann in gewissem Maße auf universeller Ebene im Klassenkontext umgesetzt werden. Der Einsatz wird dafür jedoch nur bedingt empfohlen. Alle Programme sollten in enger Begleitung durch die Lehrkraft durchgeführt werden, da die Programmkonzeption und die eher wenig selbsterklärenden Materialien dies einfordern.

Abschließend muss als Forschungsdesiderat festgehalten werden, dass hier keine deutschsprachige Maßnahme zur Förderung mathematischer Basiskompetenzen in der Sekundarstufe I identifiziert wird, die im Anschluss an die Transition auf universeller Ebene im inklusiven Setting mit allen Schülerinnen und Schülern durchgeführt werden kann.

4.3.3 Fördermethoden für den Bereich Mathematik

Es stellt sich die Frage, welche evidenzbasierten *Methoden* sich zur Förderung mathematischer Basiskompetenzen in der Sekundarstufe I anbieten. In Abgrenzung zu standardisierten Förderprogrammen (siehe Kapitel 4.3.2) können Fördermethoden auch als „unspezifische Förderung" bezeichnet werden. Die Fördermethoden sind nämlich „nicht *explizit* in Form von geschlossenen Konzeptionen, als Förderprogramme für den mathematischen Bereich ausgewiesen [...], sondern [werden] zur Leistungsförderung *auch* im mathematischen Bereich empfohlen" (Koch, 2008, S. 100, Herv. i. Orig.). „Unspezifisch" bedeutet folglich, dass der Einsatz auch in anderen Domänen oder Fächern (Leseförderung, Schreibförderung) erfolgen kann.

4.3.3.1 Identifikation (potentiell) evidenzbasierter Fördermethoden

SCHRITT 1: DAS PROBLEM IDENTIFIZIEREN

Bereits in Kapitel 4.3.2 ist die Problemlage im Kontext der standardisierten Förderprogramme identifiziert und konkretisiert. Sie wird hier deswegen nur zusammenfassend vorgestellt: Untersuchungen zeigen, dass ein nicht zu missachtender Teil der Schülerinnen und Schüler in der Sekundarstufe I Schwierigkeiten in der Beherrschung der mathematische Basiskompetenzen aufweist (Brown et al., 2003; Ehlert et al., 2013; Galton et al., 2002; Humbach, 2008). Dies zeigt sich vor allem in Schwierigkeiten in den Operationen der Grundrechenarten.

Recherchiert werden im Folgenden Fördermethoden im Bereich Mathematik, die sich für inklusive Settings – d. h. für Schülerinnen und Schüler sowohl *mit* als auch ohne sonderpädagogischem Unterstützungsbedarf – in der Sekundarstufe I anbieten. Aufgrund der aktuellen Entwicklung bezüglich der inklusiven Beschulung von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf, sind die sog. „weichen Förderbedarfe“ (Preuss-Lausitz, 2016, S. 205) im Bereich Lernen und Emotionale-soziale Entwicklung inklusiv voraussichtlich am häufigsten vertreten und werden in der Recherche deswegen besonders berücksichtigt.

SCHRITT 2: HANDLUNGSOPTIONEN BESTIMMEN

Auf Basis der Problemlage wird zur Identifikation potenzieller Fördermethoden eine Datenbankrecherche durchgeführt (www.fis-bildung.de; www.plus.orbis.uni-oldenburg.de; www.pubpsych.zpid.de; www.eric.ed.gov). Um in diesem Schritt einen ersten Überblick über mögliche Methoden zu bekommen, werden gezielt Übersichtsarbeiten (Metaanalysen, Reviews) gesucht. Verwendete Schlagwörter mit Trunkierung lauten u. a.: Effekt*, Evidenz*, Förder*, Intervention*, Lernschwierigkeit*, Mathe*, Kompetenz*, Metaanalyse, Rechen*. Die Recherche wird mit den englischsprachigen Begriffen wiederholt. Es werden dabei Quellen, die vor dem Jahr 2000 veröffentlicht sind, ausgeschlossen. Neben der Datenbankrecherche werden zudem Metaanalysen nach Sichtung weiterer fachrelevanter Literatur hinzugezogen (z. B. Grünke, 2006). Bei der Recherche konnten die Quellen aus Tabelle 21 identifiziert werden. Sie werden in der tabellarischen Übersicht mit ihren wichtigsten Merkmalen dargestellt.

TABELLE 21 INFORMATIONSQUELLEN ZU (POTENZIELL) EVIDENZBASIERTEN FÖRDERMETHODEN IM BEREICH MATHEMATIK

Bibliografische Angabe	Anzahl Studien	Zielgruppe	Alter	Analysierte Aspekte
Ise, E., Dolle, K., Pixner, S. & Schulte-Körne, G. (2012). Effektive Förderung rechenschwacher *Kinder. Eine Metaanalyse. Kindheit und Entwicklung*, 21(3), S. 181–192.	8	Rechenschwache Schülerinnen und Schüler	Primarstufe	Förderansatz, Setting, Interventionsleiter, Dauer der Förderung, Umfang der Förderung, Testverfahren, Auswertungsmaße
Gersten, R., Chard, D. J., Jayanthi, M., Baker, S. K., Morphy, P. & Flojo, J. (2009). Mathematics Instruction for Students with Learning Disabilities: A Meta-Analysis of Instructional Components. *Review of Educational Research*, 79(3), S. 1202–1242.	42	Lernschwache Schülerinnen und Schüler *(Students with learning disabilities)*	k. A.	Fördermethoden (Explizite Instruktion, Schülerverbalisierung, Visuelle Repräsentation, formative Leistungsevaluation + Feedback, Feedback, Peergestützte Instruktion, altersübegreifende Peer-Tandems, klassenintegrierte Peer-Tandems), Studien-Design
Kroesbergen, E. H. & Van Luit, J. E. H. (2003). Mathematics Interventions for Children with Special Educational Needs: A Meta-Analysis. *Remedial and Special Education*, 24, S. 97–114.	58	Rechenschwache Schülerinnen und Schüler (Students being at risk for disabilities or lagging behind to having mathematical learning disabilities)	5–12 Jahre	Studien-Design, Interventionsleiter, Interventionsmethoden (Strategie/Selbstinstruktion, Direkte Instruktion, Tutorielles Lernen, Computergestützte Förderung, Konstruktivistische Methoden)
Baker, S., Gersten, R. & Lee, D.-S. (2002). A Synthesis of Empirical Research on Teaching Mathematics to Low-Achieving Students. *The Elementary School Journal*, 103(1), S. 51–73.	15	Lernschwache Schülerinnen und Schüler *(Low-achieving students)*	7–16 Jahre	Interventionsmethoden (Tutorielles Lernen, Direkte Instruktion, Direkte Rückmeldung, Elternberatung, Konstruktivistische Unterrichtsformen
Swanson, H. L. & Sachse-Lee, C. (2000). A meta-analysis of single-subject-design intervention research for students with LD. *Journal of Learning Disabilities*, 22, S. 114–136.	85	Lernschwache Schülerinnen und Schüler (Students with learning disabilities)	Ø 10;91 Jahre	Interventionsmethoden (Direkte Instruktion, Strategieinstruktion, weitere unspezifische Methoden)

Schritt 3: Die Zielgruppe bedenken und

Schritt 4: Den Rahmen der Entscheidung beachten

Es werden im Rahmen der Fragestellung die spezifischen *Methoden* zur Förderung mathematischer Basiskompetenzen in der Sekundarstufe I aus den Metaanalysen in Tabelle 21 betrachtet. Ise und Kollegen (2012) thematisieren in ihrer Metaanalyse dabei eher Rahmenbedingungen von mathematischer Förderung und weniger explizite Methoden. Diese Rahmenbedingungen können in der Frage der Entwicklung einer Fördermaßnahme von Relevanz sein (siehe Kapitel 6).

Im Folgenden werden diejenigen Methoden dargestellt, die für das Unterrichtsfach Mathematik konzipiert und ökonomisch einsetzbar sind, im inklusiven Gruppensetting umgesetzt werden können und auch für Schülerinnen und Schüler im Sekundarstufen-I-Bereich ansprechend sind.

Die Tabelle 22 gibt einen Überblick über die – als relevant erachteten – Methoden und vergleicht die errechnete Wirksamkeit in den Metaanalysen.

Tabelle 22 Vergleich der Effektstärken (d) der Methoden aus den recherchierten Metaanalysen

Methode	Kroesbergen et al. (2003)	Gersten et al. (2009)	Baker et al. (2002)	Swanson et al. (2000)
Direkte/Explizite Instruktion	1.67	1.22	0.58	0.89
Selbstinstruktion	1.12			
Nutzung von Heuristiken		1.56		
Verbalisierung (durch Schüler)		1.04		
Visualisierungen		0.46		
Darbietung von Beispielen		0.82		
Feedback an Schüler		0.23	0.57	
Feedback an Lehrkraft inkl. Handlungsoptionen		0.40	0.51	
Peergestütztes/Kooperatives Lernen	0.76	0.14	0.42	

Anmerkung:

$d < 0.2$	$0.2 \leq d < 0.5$	$0.5 \leq d < 0.8$	$0.8 \leq d$
kein Effekt	kleiner Effekt	mittlerer Effekt	großer Effekt

Die nähere Betrachtung der Metaanalysen zeigt, dass einige Methoden von mehreren Autorengruppen betrachtet werden (zum Beispiel Direkte/Explizite Instruktion), andere hingegen nur von einer (zum Beispiel Darbietung bei Beispielen). Die dargestellten Methoden werden in der errechneten Wirksamkeit (gewichtete Effektstärke) verglichen (siehe Tabelle 22): Die Methode der direkten bzw. expliziten Instruktion erzielt durchweg mittlere bis hohe Effektstärken. Ebenso hohe Effektstärken verzeichnen die Selbstinstruktion, die Nutzung von Heuristiken, Verbalisierungen durch Schüler und die Darbietung von Beispielen. Mittlere bis kleine Effektstärken erzielen Visualisierungen, Feedback an Schüler (durch Lehrkräfte), Feedback an Lehrkräfte und peergestützte bzw. kooperative Lernformen.

Im folgenden Kapitel 4.3.3.2 werden einzelne Methoden aus der Tabelle 22 begründet ausgewählt, beschrieben und die Wirksamkeit analysiert (Schritt 5 und Schritt 6).

4.3.3.2 Analyse (potentiell) evidenzbasierter Fördermethoden

Die Methoden der Direkten Instruktion, des peergestützten bzw. kooperativen Lernens und die Methode des Feedbacks eignen sich im Rahmen strukturierter Unterrichtsplanung als *grundlegende* Methoden an. Die Direkte Instruktion wird im Regelfall in der Hinführungsphase, die peergestützten/kooperativen Lernformen in der Erarbeitungsphase und das Feedback in der Abschlussphase implementiert (siehe Kapitel 6.3.2).

Innerhalb zentraler Unterrichtsphasen eignen sich die Ansätze der Selbstinstruktion, der Nutzung von Heuristiken, der Verbalisierung durch Schüler und die Darbietung von Beispielen, die in Tabelle 22 auch hohe Effektstärken verzeichnen, für die innere Ausgestaltung der Unterrichtsphasen. So beinhaltet zum Beispiel die Phase der Direkten Instruktion auch das Lernen anhand von Beispielen; diese können jedoch auch in kooperativen Lernformen implementiert werden.

Im Folgenden werden die zentralen Methoden der Direkten Instruktion, des peergestützten/kooperativen Lernens und des Feedbacks ausgewählt und detailliert in ihrer Wirksamkeit analysiert.

SCHRITT 5: DIE QUALITÄT DER FORSCHUNG BESTIMMEN UND

SCHRITT 6: DAS AUSMAß DES EFFEKTS EINER INTERVENTION BEDENKEN

Zur Analyse dieser Methoden werden die Verzeichnisse der o. a. Metaanalysen und weitere fachrelevante Literatur hinzugezogen. Schritt 5 und Schritt 6 werden jeweils gemeinsam erarbeitet.

DIREKTE INSTRUKTION

Der Begriff „Direkte Instruktion“ wird oftmals synonym verwendet mit den Begriffen „Explizite“ oder „Systematische Instruktion“ (Lebens & Lauth, 2014). Die Direkte Instruktion ist eine lehrerzentrierte Methode, die Schülerinnen und Schüler darin unterstützt, vor allem Fakten- und Grundlagenwissen zu erlernen und zu festigen (Wiechmann, 2016, S. 40). Sie unterscheidet sich dabei von konstruktivistischen Methoden, bei denen die Verantwortung für das Lernen eher bei den Schülerinnen und Schülern selbst liegt (Przychodzin, Marchand-Martella, Martella & Azim, 2004, S. 57). Dazu gehören zum Beispiel entdeckendes Lernen, selbstgesteuertes Lernen oder Projektunterricht (Lebens & Lauth, 2014; Wiechmann, 2016), die bei lern- und rechenschwachen Schülerinnen und Schülern oftmals nur wenige Erfolge erzielen (Simon & Grünke, 2010, S. 43; Walter, 2002). Die Direkte Instruktion stützt sich auf die Annahme, dass die Arbeitsgedächtniskapazitäten und -ressourcen bei Schülerinnen und Schülern mit Lernschwierigkeiten begrenzt sind (vgl. Cognitive-Load-Theorie bei Ayres & Paas, 2012) und benötigtes Wissen, Strategien oder Strukturen nicht abrufbar sind, wenn neues Wissen gelernt werden soll.

Durch die Direkte Instruktion

> „werden wertvolle, weil begrenzte Ressourcen für die Verarbeitung und die Festigung des Lernstoffs freigemacht. Die jeweiligen Schülerinnen und Schüler sollen den Blick auf das Wesentliche richten und sichere Lerngewinne erreichen. Deshalb werden möglichst eindeutige Lernaufgaben gestellt. Durch die Art der Anleitung werden sie außerdem von Ungewissheiten über das Lernziel und den Lernweg befreit. Zudem geht es um den sicheren Erwerb von Grundkenntnissen und das Erlernen von förderlichen Lernstrategien, die das kognitive System der Schülerin bzw. des Schülers entlasten." (Lebens & Lauth, 2014, S. 420)

Die Umsetzung der Direkten Instruktion erfolgt über mehrere Teilschritte (siehe Tabelle 23). Je nach vorhandenem Vorwissen der Schülerinnen und Schüler erhalten die einzelnen Schritte mehr oder weniger Gewicht. Zum Beispiel erhält die Phase 2 (Präsentation, Modellierung) bei kaum vorhandenen Vorwissen (siehe Tabelle 23) mehr Relevanz, als läge bereits ein umfangreiches Vorwissen vor. In der Literatur sind weitere Schrittfolgen vorzufinden (vgl. Doabler & Fien, 2013; Hattie, Beywl & Zierer, 2013; Przychodzin et al., 2004; Simon & Grünke, 2010), die sich jedoch in bestimmten Aspekten (Anzahl der Schritte, Umfang) unterscheiden. Die Schrittfolge nach Lebens & Lauth (2014) bietet darunter einen konzentrierten und praxisnahen Überblick:

TABELLE 23 SCHRITTWEISE UMSETZUNG DER DIREKTEN INSTRUKTION (NACH LEBENS & LAUTH, 2014, S. 424)

Schritt	Inhalt
1. Rückschau	• Zusammenfassung und Wiederholung bereits bekannter Inhalte durch die Lehrkraft im Dialog mit der Lerngruppe
2. Präsentation	• Benennung des aktuellen Lernziels und der Lernintention • Schrittweise und systematische Präsentation und Modellierung neuer Lerninhalte anhand ausgearbeiteter Lösungsbeispiele und Muster (zum Beispiel schriftliches Addieren) durch die Lehrkraft an der Tafel, Videos, Bilder o. ä. • Verknüpfung an das vorhandene Wissen
3. Angeleitetes Üben	• Üben unter Anleitung (Beispiel: Die Lerngruppe bearbeitet gemeinsam eine neue Aufgabe zum Lernstoff an der Tafel) • Die Lehrkraft leitet an, unterstützt und verstärkt
4. Korrektur und Feedback	• Die Schülerinnen und Schüler erhalten unmittelbar Feedback zu ihren Ergebnissen
5. Selbstständiges Üben	• Die Schülerinnen und Schüler üben selbstständig, die Lehrkraft reduziert die Hilfe schrittweise • Abgleich und Korrektur der Ergebnisse • Transfer/Einordung des neuen Wissens in bisheriges Wissensnetz
6. Wöchentliche und monatliche Rückschau	• Wiederholung der Inhalte in regelmäßigen Abständen zur Festigung

Einigkeit herrscht in der Literatur darüber, dass das Design bzw. die Struktur der Direkten Instruktion, die im Vorfeld der Durchführung intensiv durchdacht werden muss, „the heart of all Direct Instruction curriculum programs" (Kame'Enui, Fien & Korgesaar, 2014, S. 494) darstellt.

Der Einsatz der lehrerzentrierten Direkten Instruktion geht nicht einher mit dem Ausschluss konstruktivistischer Unterrichtsformen (Wember, 2007). Simon und Grünke (2010, S. 45) empfehlen, eine „Lektion" ca. 15 Minuten andauern zu lassen und die restliche Zeit der Unterrichtsstunde zum Beispiel mit entdeckenden Lernformen zu gestalten und die Verortung der neuen Inhalte im „kohärenten Ganzen" (Hattie et al., 2013, S. 244) zu verdeutlichen. „Direktiv vermittelte Inhalte sind mittels offener Methoden zu verfestigen und in einen lebenspraktischen Problemkontext zu stellen" (Simon & Grünke, 2010, S. 45). Hattie, Beywl und Zierer (2013, S. 242) betonen, dass Direkte Instruktion nicht mit Frontalunterricht oder mit einem Vortrag vom Lehrertisch verwechselt werden dürfe, da die Direkte Instruktion durch klare Schrittfolgen und Grundprinzipien als fundierte Lernförderung gekennzeichnet sei. Breits seit den 1970er Jahren werden in den USA standardisierte Förderprogramme mit der Methode der Direkten Instruktion veröffentlicht (DISTAR Arithmetic, Corrective Mathematics, Connecting Maths Concepts) (für einen Überblick siehe Przychodzin et al., 2004).

WIRKSAMKEIT

Die Direkte Instruktion umfasst verschiedene Förderaspekte bzw. -methoden wie u. a. die Darbietung von Beispielen, die Nutzung von Visualisierungen oder die Vergabe von Feedback (siehe Tabelle 22). Doch stellt sich die Frage, welche Wirksamkeit die Methode im *Gesamten* mit sich bringt. In der Megaanalyse von Hattie und Kollegen (2013) erreicht die Direkte Instruktion insgesamt Platz 26 von 138 mit einer Effektstärke von $d \approx .59$. In der Domäne „Unterrichten", der diese Methode dort zugeordnet ist, nimmt sie Platz 12 ein. Grundlage der Analyse der Direkten Instruktion bilden bei Hattie und Kollegen (2013) vier Metaanalysen mit insgesamt 304 Studien.

Grundsätzlich lässt sich festhalten, dass „die Wirksamkeit der direkten Instruktion [...] in der Literatur umfassend belegt [ist]. Nur wenige andere Interventionsverfahren können sich auf eine derart breite Befundlage berufen" (Lebens & Lauth, 2014, S. 426) (vgl. auch Doabler & Fien, 2013). Die Ergebnisse in der Megaanalysen von Hattie und Kollegen (2013) sowie in der Metaanalyse von Swanson (2000) beziehen sich jedoch nicht ausschließlich auf den Bereich Mathematik. Hier ist eine domänenspezifische, differenzierte Analyse der Wirksamkeit angezeigt. Problematisch ist dabei, dass viele Veröffentlichungen aus den Jahren 1970 bis 1990 stammen. So kommen zum Beispiel Adams und Carine (2003) in ihrer Metaanalyse für die Wirksamkeit im Bereich Mathematik zu einer großen Effektstärke von $d = 1.08$, doch stammen die dafür verwendeten Studien bereits aus den Jahren 1986 bis 1990.

Im Folgenden werden jüngere Studien zur Wirksamkeit des Einsatzes der Direkten Instruktion im Bereich Mathematik dargestellt und im Detail analysiert, um anschließend die Einordnung der Methode in eine Evidenzstufe (Council for Exceptional Children, 2014) zu ermöglichen. Die aufgeführten Studien stammen aus den o. a. identifizierten Metaanalysen (Baker et al., 2002; Gersten et al., 2009; Kroesbergen & Van Luit, 2003; Swanson & Sachse-Lee, 2000) oder entstammen weiterer fachrelevanter Literatur (z. B. Grünke, 2006; Lebens & Lauth, 2014; Walter, 2002) und verzeichnen im Regelfall eine Veröffentlichung nach dem Jahr 2000 (ausgenommen die Follow-Through-Studie).

Die erste empirische Auseinandersetzung mit der Methode der Direkten Instruktion erfolgt im Kontext des Follow-Through-Projekts, das im Jahr 1967 beginnt und 1995 endet. Im Rahmen einer Bildungsinitiative zur Armutsbekämpfung des U. S. Departments of Education (DOE)

und des Office of Economic Opportunity zählt das Projekt mit über 200.000 Kindern und Jugendlichen heute zu den weltweit größten pädagogisch-empirischen Studien. Innerhalb des Projekts können Eltern im Rahmen von Werbe- und Informationsveranstaltungen (parent assembly) zu Beginn darüber abstimmen, welche der zahlreichen vorgeschlagenen Methoden (vgl. Abbildung 31) in der Schule ihrer Kinder für mehrere Jahre umgesetzt werden soll. Das Modell des Direkten Unterrichts wird dabei am häufigsten gewählt. Im Zeitraum der Umsetzung werden dann anhand standardisierter Schulleistungstests und psychometrischer Skalen die Leistungen in schulischen Domänen (Lesen, Rechnen, Rechtschreibung, Sprache, schulisch kognitive Fähigkeiten, Selbstwertgefühl) erhoben (Lebens & Lauth, 2014, S. 426) und Schulen mit gleichen Ausgangsniveau untereinander verglichen. Die Methode der Direkten Instruktion gilt dabei als „undisputed winner" (Engelmann, 2007, S. 223). Im Rahmen des Projekts gilt eine Methode dann als lernförderlich, sobald die Schülerinnen und Schüler nach Abschluss im Durchschnitt über dem 20. Perzentil liegen. Die Ergebnisse (siehe Abbildung 31) zeigen, dass die Methode der Direkten Instruktion in den vier schulischen Domänen Lesen, Rechnen, Rechtschreiben und Sprache jeweils über dem 20. Perzentil liegt und auch absolut die höchsten Werte erzielt. Zudem haben diejenigen Schülerinnen und Schüler, die über drei Jahre mit der Methode unterrichtet werden, ein Jahr Vorsprung zu den Schülerinnen und Schülern, die mit anderen Methoden unterrichtet werden (Engelmann, 2007, S. 229).

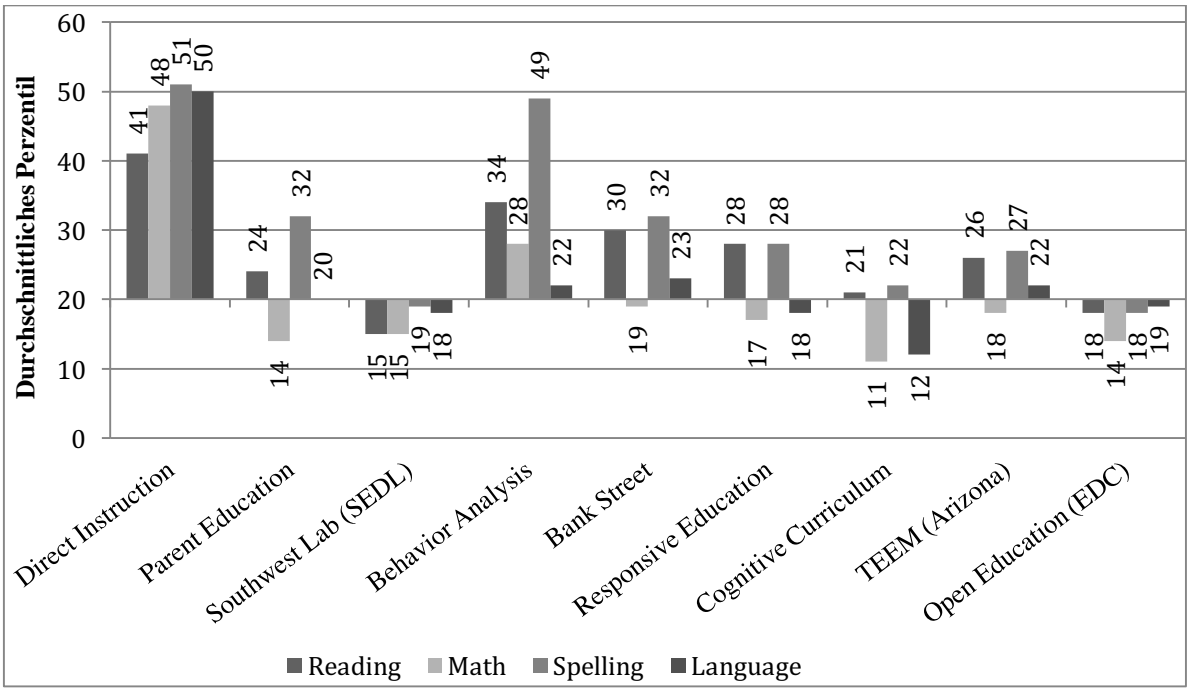


ABBILDUNG 31 ERGEBNISSE DES FOLLOW-THROUGH-PROJEKTS IN DEN DOMÄNEN LESEN, RECHNEN, RECHTSCHREIBEN UND SPRACHE (W. C. BECKER & ENGELMANN, 1996, S. 36 FF; EIGENE DARSTELLUNG)

Auch bezogen auf „mechanisches Wissen" (basic skills), auf höhere kognitive Kompetenzen wie Problemlösung oder logisches Denken (cognitive skills) und auf Selbstvertrauen (affective skills) zeigt die Direkte Instruktion die größten Effekte. Sowohl bezogen auf die Kulturtechniken als auch auf persönliche Kompetenzen zeigt die Follow-Through-Studie, dass es Methoden gibt, die sogar schädlich sind und die Schülerinnen und Schüler sich in bestimmten Bereichen verschlechtern. Dazu zählt in diesem Projekt zum Beispiel das Konzept des „Open Education".

Trotz eindeutiger Forschungsergebnisse des umfangreichen und kostspieligen Projekts unternimmt die amerikanische Bildungspolitik kaum Folgemaßnahmen. Engelmann nennt sein 2007 herausgebrachtes Buch entsprechend „The Outrage of Project Follow Through: 5 Million Failed Kids Later". Watkins (1996) bemängelt, dass in regelmäßigen Abständen schulorganisatorische Änderungen vorgenommen werden (zum Beispiel Uhrzeit des Unterrichtsbeginns), jedoch nicht erkannt wird, dass „our educational problems will not be solved until it is recognized that how well students learn is directly related to how well they are taught".

Neben der etwas älteren Follow-Through-Studie haben sich in jüngerer Vergangenheit Forschergruppen mit der Wirksamkeit der Direkten Instruktion auseinandergesetzt:

Tournaki (2003) vergleicht die Wirksamkeit der Direkten Instruktion mit der sog. Drill-and-Practice-Methode bei 48 Schülerinnen und Schülern mit und 48 Schülerinnen und Schülern ohne Lernbeeinträchtigungen (learning disabilities) in der zweiten Klasse. Bei der Methode des Drill and Practice werden die Schülerinnen und Schüler aufgefordert, 20 Aufgaben zum Plusrechnen (zum Beispiel „6 + 3 =") auf einem Aufgabenblatt so schnell wie möglich zu lösen. Falsch gerechnete Aufgaben werden im Anschluss von der Lehrkraft markiert und von den Schülerinnen und Schülern erneut bearbeitet. Sind sie dann wieder nicht richtig gelöst, wird die richtige Lösung genannt.

Bei der Direkten Instruktion hingegen wird zu Beginn des Unterrichts die Strategie zur Lösung der Plusaufgaben präsentiert. Im Anschluss rechnen die Schülerinnen und Schüler ebenso 20 Aufgaben so schnell sie können. Fehler werden unmittelbar aufgedeckt und die Lösungsstrategie entsprechend wiederholt. Ein Aufgabenblatt gilt als gelöst, sofern es zu 90 % richtig ist und in weniger als 80 Sekunden bearbeitet wird. Die beiden Methoden werden über acht Unterrichtsstunden mit unterschiedlichen Aufgabenblättern jeweils in einem Umfang von ca. 15 Minuten durchgeführt. In beiden Gruppen und einer Kontrollgruppe befinden sich randomisiert Schülerinnen und Schüler mit und ohne Lernbeeinträchtigungen.

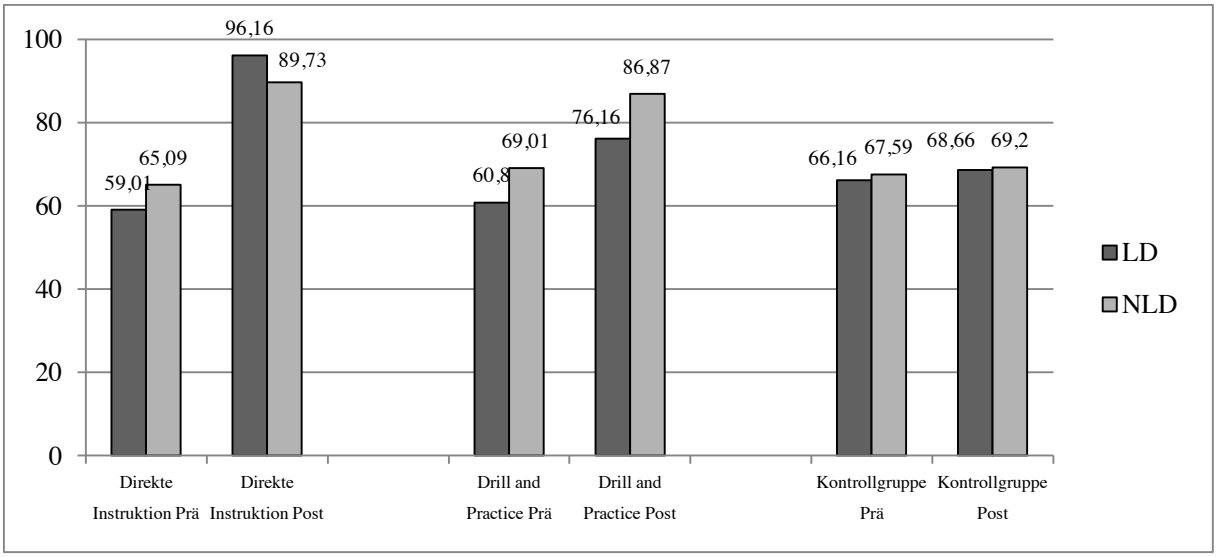


ABBILDUNG 32 VERGLEICH DER DIREKTEN INSTRUKTION UND DEM SOG. DRILL AND PRACTICE (TOURNAKI, 2003)

Die Ergebnisse (siehe Abbildung 32) im Prä-Posttest-Vergleich zeigen, dass die Schülerinnen und Schüler mit Lernbeeinträchtigungen (LD) stärker von der Methode der Direkten Instruktion profitieren, wohingegen bei Schülerinnen und Schüler ohne Lernbeeinträchtigungen (NLD)

beide Methoden lernförderlich sind. Bei der Kontrollgruppe stellt sich erwartungsgemäß keine Veränderung ein. Transferaufgaben können im Anschluss diejenigen Schülerinnen und Schüler besser lösen, die an der Methode der Direkten Instruktion teilnehmen. Beide Methoden führen im Vergleich zur Kontrollgruppe dazu, dass die Schülerinnen und Schüler schneller rechnen können (Rechengeschwindigkeit).

Die Methode der Direkten Instruktion eignet sich nicht nur zur Vermittlung arithmetischen Basiswissens, sondern auch im Kontext des Lösens von Sach- und Problemaufgaben. Xin und Kollegen (2005) ermitteln eine Effektstärke von $d = 1.69$ für das Bearbeiten von Sachaufgaben (zum Beispiel „Vito earned $ 12 from shoveling snow over the weekend. He earned 1/3 as much as his friend Guy did. Guy earned $ 36 from shoveling snow.“) über die Entwicklung von mathematischen Diagrammen/Skizzen, die dann in eine Rechenaufgabe (math sentence) übergehen.

Owen und Fuchs (2002) evaluieren mit 24 lernbeeinträchtigen Drittklässlern das Thema „Halbieren“ mit drei verschiedenen Designs der Direkten Instruktion. Es zeigt sich, dass der größte Effekt entsteht, wenn sich an zwei Einführungsstunden erst eine Wiederholungsphase anschließt und dann Transferaufgaben folgen (jeweils in zwei Unterrichtsstunden) ($d = 1.63$). Geringer ist der Effekt, wenn keine Wiederholung ($d = 1.17$) oder kein Transfer ($d = .69$) durchgeführt werden. Im Vergleich zur Kontrollgruppe zeigt sich, dass alle Designs große Effekte mit sich bringen.

Fuchs und Kollegen (2002) zeigen bei dem Einsatz der Direkten Instruktion bei Sachaufgaben bei Schülerinnen und Schülern mit einer Rechenstörung (mathematics disabilities), dass eine zusätzliche Förderung mittels eines Computers förderlich ist, der alleinige Einsatz eines Computers jedoch nur geringe Effekte mit sich bringt (siehe Tabelle 24).

Dass die lehrerzentrierte Direkte Instruktion jedoch gerade bei Schülerinnen und Schülern mit Rechenschwierigkeiten wirksamer ist als ein schülerzentrierter Zugang oder der Einsatz von Veranschaulichungsmaterial (sog. *„manipulatives“*), Musik oder Bewegung zeigen die Ergebnisse von Morgan und Kollegen (2014), die leider jedoch keine Effektstärken angeben.

FAZIT

Nach der Klassifikation Evidenzbasierter Praxis des Council for Exceptional Children (2014) liegen für die Methode der Direkten Instruktion im Bereich der Förderung mathematischer Basiskompetenzen zwei gut fundierte randomisierte Gruppenvergleichsstudien mit überaus positiven Effekten und jeweils über 60 Probanden vor (Engelmann, 2007; Tournaki, 2003). Zudem werden weitere Studien im Kontext des Lösens von Sach- und Problemaufgaben mit geringerer Probandenzahl identifiziert, die ebenso von positiven Effekten sprechen (Fuchs, Fuchs, Hamlett, et al., 2002; Owen & Fuchs, 2002; Xin et al., 2005). Es wird keine Studie identifiziert, die von einer negativen Wirksamkeit berichtet. Die Methode der Direkten Instruktion kann als evidenzbasiert deklariert werden (vgl. Tabelle 16).

Das Fazit der Wirksamkeitsanalyse nach dem Council for Exceptional Children (2014) für die Methode der Direkten Instruktion lautet: **Evidenzbasierte Praxis**.

Peer-mediierte Instruktion

In der deutschsprachigen und anglo-amerikanischen Literatur ist nicht abschließend geklärt, was konkret unter tutoriellem und kooperativen Lernen verstanden wird und wie die Methoden voneinander abgegrenzt werden (Büttner, Warwas & Adl-Amini, 2012). Beide Methoden zählen zur „peer-mediierten Instruktion" (Büttner et al., 2012; Mastropieri et al., 2008), was grundlegend bedeutet, dass sich Schülerinnen und Schüler beim Lernen gegenseitig unterstützen. Im Folgenden wird versucht, die beiden Methoden voneinander zu unterscheiden und differenziert in ihrer Wirksamkeit zu analysieren.

Peergestütztes Lernen

Peergestütztes Lernen wird hier als Oberbegriff für „peer-assisted learning" (Baker et al., 2002), „peer tutoring" (Kroesbergen & Van Luit, 2003) oder auch „peer-mediated instruction" (Wexler, Reed, Pyle, Mitchell & Barton, 2015) gewählt. Peergestütztes Lernen hat viele Facetten, sodass vorerst eine konkrete Beschreibung der hier betrachteten Methode angezeigt ist.

Peergestütztes Lernen wird im deutschsprachigen Zusammenhang oftmals synonym verwendet mit „tutoriellem Lernen" (Haag & Streber, 2011). Folgende Definition beschreibt die Methode:

> „Beim tutoriellen Lernen arbeiten zwei Schülerinnen bzw. Schüler gemeinsam an der Wiederholung, Vertiefung und Überprüfung ihrer Kenntnisse in einem Unterrichtsfach. Das eine Kind übernimmt die Rolle der Tutorin bzw. des Tutors, das andere die der Tutandin bzw. des Tutanden. Die Tutorin bzw. der Tutor fungiert als Lehrkraft, die Wissen vermittelt, abprüft und die Antworten der Tutandin bzw. des Tutanden korrigiert. Dazu stellt die Tutorin bzw. der Tutor Fragen, entwickelt Testaufgaben und erteilt Rückmeldungen zur Richtigkeit der Antworten. Die Tutandin bzw. der Tutand verhält sich wie eine Schülerin bzw. ein Schüler. Sie bzw. er beantwortet Fragen, bearbeitet Aufgaben und erläutert die Lösungen." (Haag, 2014, S. 463)

Im Regelfall besteht die Konstellation aus zwei Personen (sog. *„dyads"*) (1 : 1) (Büttner et al., 2012), es sind jedoch auch größere Gruppen mit unterschiedlichen Rollenverhältnissen möglich (zum Beispiel 2 : 10) (Topping, 2005, S. 633). Die Zusammensetzung der Tandems kann dabei unterschiedlich sein. Neben der Bildung von Tandems innerhalb einer Klasse („peer-assisted learning within a class") können Tandems auch aus einem älteren und jüngeren Part bestehen („cross-age tutoring") (Gersten u. a., 2009; siehe auch die Untersuchung bei Haag & Streber, 2011). Oftmals sind die Tandems leistungsheterogen, denkbar ist jedoch auch eine leistungshomogene Zusammensetzung (Haag, 2014). Topping (2005, S. 633 f.) ergänzt weitere Aspekte: Er unterscheidet, ob die Tandems über einen längeren Zeitraum fest bestehen oder nur für spezifische Arbeitsphasen gebildet werden. Weiterhin kann die Rollenzuweisung als Tutand oder Tutor dauerhaft bestehen bleiben oder regelmäßig wechseln. Auch kann die Zielgruppe eine gesamte Klasse oder nur eine Lerngruppe sein, die zum Beispiel besondere Schwierigkeiten oder Stärken im Lernen aufweist. Weiterhin unterscheidet Topping (2005), ob das peergestütze Lernen auf Freiwilligkeit oder auf Aufforderung basiert.

Durch das Hineinschlüpfen in eine Rolle kann peergestütztes Lernen die Lernaktivität steigern und vertiefen. Es wird vermieden, dass Schülerinnen und Schüler eine passive Lernhaltung einnehmen, wie es im allgemeinen Unterricht schnell passieren kann (Büttner et al., 2012). Tutand und Tutor nehmen beständig einen aktiven Part mit konkreten Aufgaben ein. Das Geben von Rückmeldungen zum Lernprozess und das gemeinsame Reflektieren und Lernen an einem Ge-

genstand führen zu einem vertieften Erwerb neuen Wissens und fördert die sozialen Kompetenzen (Haag, 2014, S. 463). Damit peergestütztes Lernen erfolgreich ist und Schwierigkeiten vorgebeugt werden, sollte es gründlich vorbereitet sein (Souvignier, 2007). Haag und Streber (2011, S. 359) schlagen folgende Aspekte vor, die in der Vorbereitung und Durchführung berücksichtigt werden sollten:

- Die Zusammensetzung der Lernpartner (Wer soll wem helfen?)
- Die Zielvorgaben (Was soll erreicht werden?)
- Die Zeitplanung (Wann und wie lange soll gelernt werden?)
- Das Lernmaterial (Womit soll gelernt werden?)
- Die Einübung des Tutorenverhaltens (Was sind die jeweiligen Aufgaben?)
- Die Registrierung von Lernfortschritten (Wie werden Lernfortschritte festgehalten?)

Die Bedeutung eines Klimas unterstützenden und hilfsbereiten Lernens innerhalb der Klasse spielt für eine erfolgreiche Umsetzung eine bedeutsame Rolle. Souvignier (2007, S. 144) betont in diesem Zusammenhang, dass „soziale Kompetenzen nicht automatisch eine Folge von Gruppenarbeit ist, sondern dass – umgekehrt – Gruppenarbeit nur dann funktionieren kann, wenn kooperative Kompetenzen explizit vermittelt und eingeübt werden“.

WIRKSAMKEIT

Bei der Analyse der Wirksamkeit der Methode des peergestützten Lernens muss stets berücksichtigt werden, welche Umsetzung die Methode in dem jeweiligen Forschungskontext erfahren hat. Die bisherigen Ausführungen zeigen, dass vielfältige Umsetzungen möglich sind. Der Großteil der Studien betrachtet die Wirksamkeit bei heterogenen Lerngruppen bestehend aus je zwei Personen, sodass auch im Folgenden in der Frage der Wirksamkeit der Schwerpunkt auf dieser Form der Umsetzung liegt, sofern nicht explizit anders benannt.

Hattie und Kollegen (2013, S. 435) verorten die Methode des Peer-Tutorings in ihrer domänenübergreifenden Megaanalyse mit einer mittleren Effektstärke von $d = .55$ auf Platz 36 von 138. In der spezifischen Kategorie „Unterricht“, in der die Methode der Direkten Instruktion zum Beispiel Platz 12 einnimmt, platziert sich das Peer-Tutoring an Stelle 17. Grundlage für die Analyse bilden 14 Metaanalysen mit insgesamt 767 Studien. Rohrbeck und Kollegen (2003) führen in ihrer Metaanalyse 90 Studien zur Wirksamkeit des Peer-Tutoring in den Klassen eins bis sechs zusammen und ermitteln eine mittlere Wirksamkeit von $d_{ungewichtet} = .59$ bzw. $d_{gewichtet} = .33$. Einen ebenso allgemeinen Überblick über die Wirksamkeit des peergestützten Lernens in den Klassen sechs bis 12 bei Schülerinnen und Schülern ohne und mit leichten intellektuellen Beeinträchtigungen geben Stenhoff und Linguaris-Kraft (2007) in ihrem Literaturreview. Die Autoren kommen nach der Zusammenführung von 20 Studien zu dem Schluss: „Peer tutoring in secondary settings results in improved academic performance of students with mild disabilities and could be classified as an evidence-based practice“ (Stenhoff & Linguaris-Kraft, 2007, S. 9). Die 20 Studien beziehen sich jedoch auf verschiedene schulische Domänen und Zielgruppenspezifika, sodass das Review einen eher allgemeinen Überblick zulässt.

Im Folgenden wird auf die Wirksamkeit der Methode im Bereich Mathematik fokussiert und hier Literatur ab dem Jahr 2000 berücksichtigt, um möglichst aktuelle Forschungserkenntnisse

zu gewinnen. Nicht berücksichtigt werden in der Analyse ältere Studien und diejenigen, die das Vorschulalter betreffen (siehe z. B. Fuchs, Fuchs & Karns, 2001).

Die Forschungen bezüglich der Wirksamkeit des peergestützten Lernens im Bereich Mathematik sind, vor allem im Vergleich zum Bereich Lesen und Rechtschreiben, sehr gering. Wexler und Kollegen (2015) identifizieren in ihrer Literatursynthese (2001 bis 2012) unter 13 Studien zum peergestützten Lernen bei lernbeeinträchtigten Schülerinnen und Schülern lediglich eine Studie im Bereich Mathematik. Bei Stenhoff und Linguaris-Kraft (2007) sind es fünf von 20 Studien aus den Jahren 1980 bis 1998. In vier der fünf Studien liegen bei den Schülerinnen und Schülern keine Beeinträchtigungen vor. Zudem zeigen sich die Studienergebnisse insgesamt als wenig einheitlich, wenn nicht gar widersprüchlich. Bezogen auf die o. a. Metaanalysen (siehe Tabelle 21) reichen die Effektstärken von noch nicht einmal kleinen (d = .14 Gersten et al., 2009), über mittlere (d = .66 Baker et al., 2002) bis hin zu großen Werten (d = .87 Kroesbergen & Van Luit, 2003). Kunsch, Jitendra und Sood (2007) fassen in ihrer Metaanalyse speziell zum Einsatz des peergestützten Lernens bei lern- und rechenschwachen Schülerinnen und Schülern 17 Studien zusammen, die von Effektstärken von $d = -0.02$ bis hin zu $d = 1.77$ (Ø $d = 0.47$) berichten. Die Wirksamkeit der Methode sinkt dabei mit dem Grad der Beeinträchtigung ($d_{at\ risk} = 0.66$, $d_{Combined} = 0.53$, $d_{Disabled} = 0.21$). Auch zeigt sich, dass die Förderung von Rechenoperationen und -strategien durch Peer-Tutoring eine höhere Wirksamkeit (d = .63) aufweist als die zeitgleiche Förderung von Anwendungswissen (d = .34). Dass die Förderung von Rechenstrategien durch Peer-Tutoring wirksamer ist als auf die Fokussierung auf anwendungsbezogene mathematische Kompetenz, bestätigen auch Baker und Kollegen (2002) (d = 62 vs. d = .29) sowie Calhoon und Fuchs (2003). Slavin, Lake und Groff (2009) stufen das PALS-Programm (Peer-Assisted Learning Strategies) im Bereich Mathematik für die Klassen 5 bis 12 sogar in der Stufe „Insufficient evidence of effectiveness“ ein, wobei sich die Autoren in der Analyse nach den strengeren Kriterien des „What Works Clearinghouse“ richten.

Eine oftmals zitierte und zentrale Studie zum peergestützten Lernen im Bereich Mathematik stammt von Calhoon und Fuchs (2003). Die Autorinnen untersuchen die Wirksamkeit des peergestützten Lernens (PALS) in Kombination mit curriculumbasierten Messen bei 92 Schülerinnen und Schülern mit verschiedenen Beeinträchtigungsformen in den Klassen neun bis 12. Die Schülerinnen und Schüler werden randomisiert zwei Gruppen ($n_{PALS+CBM} = 45$; $n_{Kontrollgruppe} = 47$) zugewiesen. Die Experimentalgruppe führt zweimal in der Woche peergestütztes Lernen und einmal curriculumbasiertes Messen durch. Der Untersuchungszeitraum umfasst 15 Wochen.

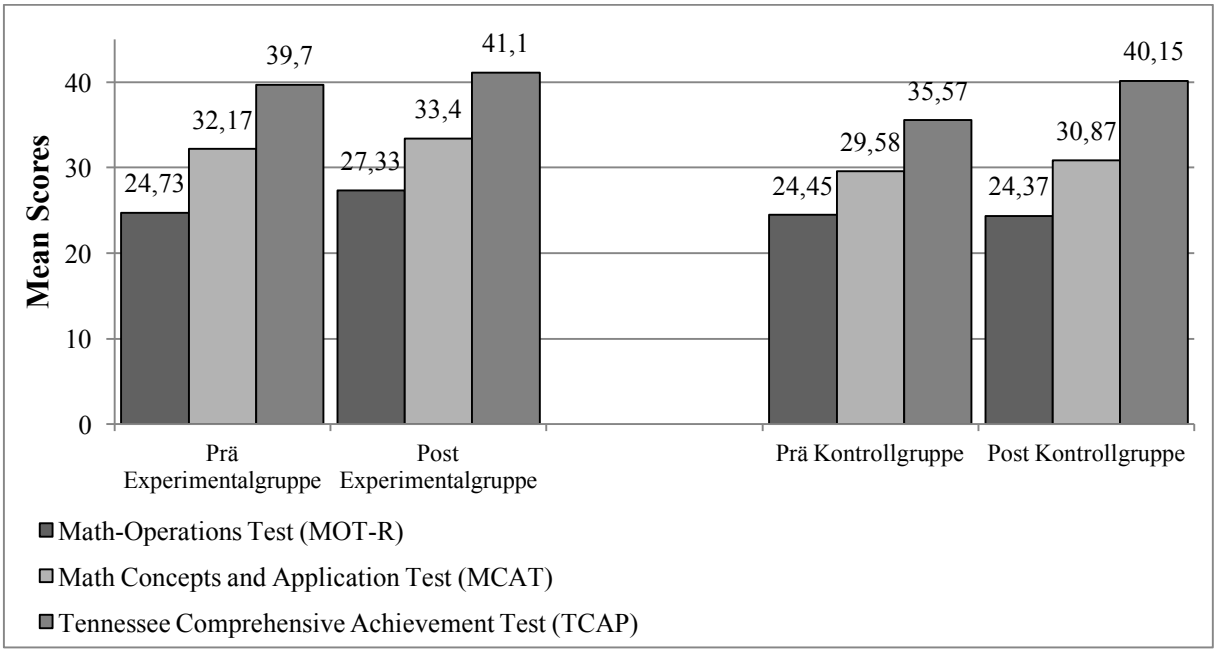


ABBILDUNG 33 ERGEBNISSE DER UNTERSUCHUNG (CALHOON & FUCHS, 2003, S. 241)

Die Ergebnisse (siehe Abbildung 33) zeigen, dass die Experimentalgruppe in allen drei eingesetzten Mathematiktests in der Post-Erhebung höhere Werte erzielt. Aber auch die Kontrollgruppe macht Fortschritte. Lediglich der Unterschied bezüglich des Math-Operation-Tests (Rechenoperationen) ist signifikant, die Veränderung in den anderen beiden Testverfahren ist in den beiden Gruppen vergleichbar. Die Effektstärke liegt bei $d_{\text{MOT-R}} = .40$. Da der Leistungszuwachs in der Kontrollgruppe von der Prä- zur Posttestung in absoluten Zahlen größer ist als in der Experimentalgruppe ($d_{\text{MCAT}} = -0.01$, $d_{\text{TCAP}} = -.29$) ermitteln Calhoon und Fuchs (2003) insgesamt eine kleine Gesamteffektstärke ($d = 0.01$). Als positiv heben die Autorinnen hervor, dass die Schülerinnen und Schüler angeben, dass die Methode des Peer-Tutoring Freude bereitet und der Einsatz curriculumbasierte Messungen motivierend auf sie wirkt.

Tsuei (2014) führt in seiner Studie über ein Schuljahr mit vier Schülerinnen und Schülern mit Lernbeeinträchtigungen in der vierten Klasse einmal in der Woche peergestütztes Lernen durch. Die Lehrkraft hält dabei zu Beginn der Stunde einen etwa zehnminütigen Input zum Thema. Anschließend werden Mathematikaufgaben durch ein Computerprogramm (G-Math) gestellt und zuerst von dem Tutor (als Modell) gelöst. Der Tutand beobachtet den Tutor beim Lösen und gibt bei richtig gelöster Aufgabe einen Punkt. Dann wechseln sie die Rollen und der Tutand wird zum Tutor. Tsuei (2014) begleitet die Schülerinnen und Schüler mit curriculumbasierten Messungen und kann bei drei Probanden signifikante Verbesserungen feststellen (Rosenthal's $r = .37$; $r = .53$; $r = .33$). Bezogen auf die Subtestebene zeigt sich in dieser Studie, dass kein Proband signifikante Verbesserungen im Bereich Rechenoperation zeigt. Bessere Ergebnisse werden von drei Probanden im Posttest in den Bereichen der Anwendungs- und Konzeptaufgaben (sog. „conceptual questions") erzielt. Insgesamt benötigen die Schülerinnen und Schüler mehr Zeit beim Lösen der Rechenaufgaben als bei den Anwendungs- und Konzeptaufgaben, was auch Hinweise dafür sein kann, dass das Format eines Computers im Kontext Rechenoperationen eher kein anwendungsfreundliches Medium darstellt (Tsuei, 2014, S. 124).

Spörer (2009) fördert in ihrer Studie die mathematischen Basiskompetenzen bei Zweitklässlerinnen und -klässlern über neun Schulstunden mit Hilfe des Peer-gestützten Lernens (PALS)

oder durch die Bearbeitung von Arbeitsblättern (Kontrollgruppe). Je eine Schulklasse wird einer Bedingung randomisiert zugewiesen ($n_{Kontrollgruppe} = 20$; $n_{PALS} = 19$). Bei dem Lernstoff handelt es sich um die Inhalte der ersten Klasse (Zahlen erkennen und darstellen, Vorgänger und Nachfolger, Zahlen vergleichen, addieren und subtrahieren). Die Ergebnisse des Prä-Post-Vergleichs zeigen, dass sich die Experimentalgruppe und Kontrollgruppe mit einem kleinen bis mittleren Effekt vergleichbar in ihren Leistungen, gemessen durch den DEMAT 1+, verbessern (d_{PALS} = .40, $d_{Kontrollgruppe}$ = .49). Aufschlussreich sind hier jedoch die Ergebnisse im Detail, die zeigen, dass gerade die leistungsschwachen Schülerinnen und Schüler (L–) hohe Leistungszuwächse erzielen ($d_{L–PALS}$ = 1.87, $d_{L–Kontrollgruppe}$ = 1.70) und dabei die Methode des peergestützen Lernens einen kleinen Vorsprung aufweist. Spörer (2009) diskutiert, ob die geringen Leistungszuwächse in der leistungsstarken Gruppe (L+) (d_{L+PALS} = .11, $d_{L+Kontrollgruppe}$ = .43) dadurch begründet werden können, dass die Wiederholung und Festigung der Inhalte aus der ersten Klasse eventuell zu Deckeneffekten führt.

Fuchs, Fuchs, Hamlett und Appleton (2002) zeigen in ihrer Untersuchung, dass eine Förderung mittels Peer-Tutoring gegenüber einer zusätzlichen Förderung mit einem Computer die größten Effekte erzielt. Dafür werden 40 Viertklässlerinnen und -klässler mit Rechenschwierigkeiten zu gleichen Teilen (n = 10) zu einer von vier Gruppen (Kontrollgruppe, Computerförderung, Peer-Tutoring, Peer-Tutoring + Computerförderung) zugewiesen. Alle Schülerinnen und Schüler nehmen an einer regulären Unterrichtseinheit über sechs Unterrichtsstunden zum Thema Problemlösen teil. Die beiden Experimentalgruppen „Peer-Tutoring“ und „Computerförderung“ erhalten zu jeder Unterrichtsstunde zusätzlich vier Übungsstunden, die kombinierte Gruppe erhält acht zusätzliche Stunden, alle Gruppen in einem Zeitraum von zwölf Wochen. Die Aufgaben zu den „Story Problems“ beinhalten die gleiche Struktur, wie die Aufgaben in den Übungsstunden. Die Aufgaben zum „Transfer story problem“ sind ebenso den Übungsaufgaben angelehnt, beinhalten jedoch eine Veränderung (zum Beispiel bzgl. des Vokabulars oder der Operationalisierung), die den Schülerinnen und Schülern noch nicht bekannt ist. „Real-world problem solvings“ sind Aufgaben, die in eine umfassende Alltagssituation eingebettet sind und die zum Beispiel zusätzliche, irrelevante Informationen enthalten. Alle Untersuchungsgruppen erhalten im regulären Mathematikunterricht (sechs Unterrichtsstunden in der Einheit) eine Strategieinstruktion mittels Direkter Instruktion. Beim „Peer-Tutoring“ gehen die Übungsstunden ebenso mit direkter Instruktion einher, bei der Computerförderung mit Feedback und Belohnung.

TABELLE 24 EFFEKTSTÄRKEN DER TREATMENTS (D) (FUCHS, FUCHS, HAMLETT, ET AL., 2002, S. 100)

	Tutoring + Computer vs.			**Computer vs.**		**Tutoring vs.**
Variable	Computer	Tutoring	Control	Tutoring	Control	Control
Story problems	1.28	– 0.03	1.90	– 1.60	0.51	2.10
Transfer story problems	0.83	– 0.27	1.31	– 1.31	0.61	1.67
Real-word problems solving	0.07	0.14	0.75	0.05	0.64	0.64

Die Ergebnisse der Untersuchung (siehe Tabelle 24) zeigen, dass die Methode des Peer-Tutorings inklusive Direkter Instruktion die höchste Wirksamkeit bezüglich der Aufgaben „Story

Problems“ und „Transfer story Problems“ auf zweiter Abstraktionsstufe zeigt. Bei der höchsten Abstraktionsstufe, die lebensweltbezogenen Aufgaben, scheint die Kombination aus Peer-Tutoring und Computerförderung den größten Effekt zu erzielen.

Eine weitere Studie zur Wirksamkeit peergestützten Lernens stammt von Fuchs, Fuchs, Yazdian und Powell (2002). Über einen Zeitraum von 16 Wochen führen 323 Erstklässlerinnen und -klässler dreimal in der Woche über 30 Minuten peergestütztes Lernens (PALS) im Mathematikunterricht durch. Eine vergleichbar große Kontrollgruppe erhält dabei den regulären Mathematikunterricht. Die heterogenen Tandems werden mittels Medianhalbierung der Prätest-Werte zusammengestellt. Die Zusammensetzung rotiert dabei alle drei Wochen nach einem bestimmten System. Die Übungen in den PALS-Einheiten gehen inhaltlich bis in den Zahlenraum bis zehn und thematisieren zum Beispiel Mehr/Weniger- oder Additions- und Subtraktions- Aufgaben. Der Stanford Achievement-Test, der in der Prä- und Posttestung eingesetzt wird, enthält dabei Aufgaben, die sich zum einen speziell auf die Übungen aus den peergestützten Einheiten beziehen (aligned Items) und zum anderen unspezifische Aufgaben, die dem Curriculum entsprechen, nach dem die Kontroll- und Experimentalgruppe unterrichtet werden (unaligned Items). Die Ergebnisse in Abbildung 34 zeigen, dass die Schülerinnen und Schüler, die die PALS-Methode anwenden (Experimentalgruppe EG) in der Posttestung sowohl bei den PALS-spezifischen, als auch bei den PALS-unspezifischen Aufgaben höhere Werte erzielen. Jedoch ist nur der Unterschied bei den spezifischen Aufgaben signifikant mir einer Effektstärke von $d = .31$. Für die Schülerinnen und Schüler mit einer Beeinträchtigung liegt die Effektstärke bei den spezifischen Aufgaben gar bei $d = .55$.

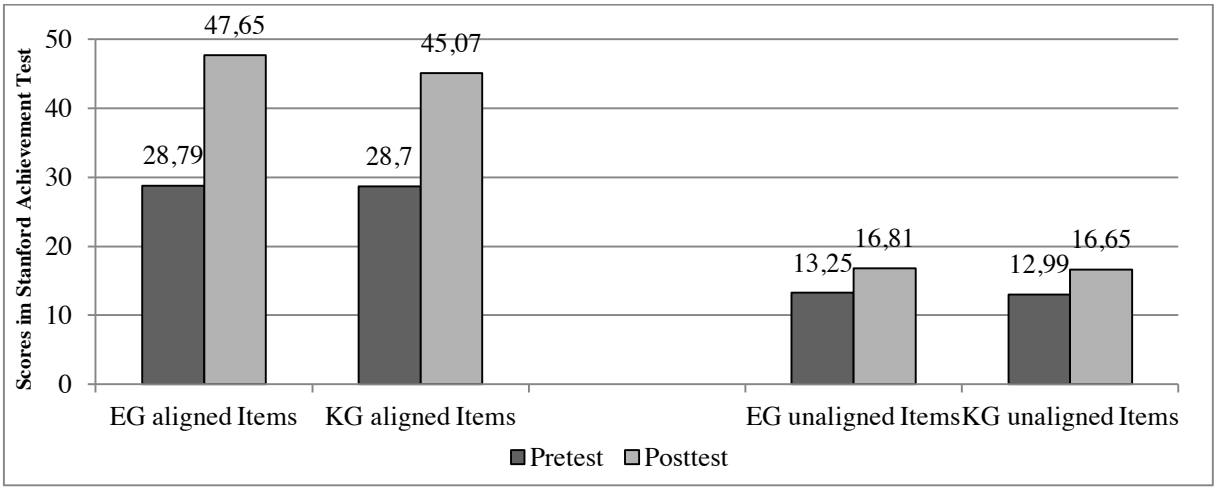


ABBILDUNG 34 ENTWICKLUNG NACH DER INTERVENTION (FUCHS, FUCHS, YAZDIAN, ET AL., 2002, S. 577)

FAZIT

Nach der Klassifikation Evidenzbasierter Praxis des Council for Exceptional Children (2014) liegen für die Methode des peergestützen Lernens im Bereich der Förderung mathematischer Basiskompetenzen ab dem Jahr 2000 zwei Gruppenvergleichsstudien vor, die zwar weniger als 60 Probanden haben, jedoch von positiven Befunden berichten (Fuchs, Fuchs, Hamlett, et al., 2002; Spörer, 2009). Einschränkend bezieht sich die Untersuchung von Spörer (2009) nur nachgeordnet auf die Schülerinnen und Schüler mit Lernbeeinträchtigungen, Fuchs und Kollegen (2002) konzentrieren sich auf Schülerinnen und Schüler mit Rechenschwierigkeiten. Eine Gruppenvergleichsstudie mir einer Stichprobengröße von mehr als 60 Probanden stammt von

Fuchs und Kollegen (2002) und berichtet – zumindest bezüglich PALS-spezifischer Aufgaben – von positiven Effekten. Bei unspezifischen Aufgaben hingegen sind die Befunde weniger eindeutig. Eine Einzelfallstudie mit ebenso positiven Effekten liegt von Tsuei (2014) vor. Die an vielen Stellen rezipierte Studie von Calhoon und Fuchs (2003) verfügt über die nötige Stichprobengröße, berichtet jedoch in den Domänen der Anwendungsaufgaben und den Aufgaben zu inhaltsbezogenen Kompetenzen wie zum Beispiel Raum und Form (außer Rechenoperationen) sogar von negativen Effekten.

Es zeigt sich zusammenfassend, dass die Forschungsergebnisse bezüglich der Methode des Peer-Tutoring sehr uneinheitlich sind (Büttner et al., 2012). In jüngster Zeit wird vermehrt diskutiert, die Wirksamkeit dieser Methode eher über Interviews oder Beobachtungen als über Leistungstests zu operationalisieren, um die vielschichtigen Ziele der Methode zuverlässiger erfassen zu können (Murphy, Grey & Honan, 2005). Eine eindeutig evidenzbasierte Methode liegt entsprechend dieser Analyse nicht vor (vgl. Tabelle 16).

Das Fazit der Wirksamkeitsanalyse nach dem Council for Exceptional Children (2014) für die Methode des peergestützten Lernens lautet: **Gemischte Evidenz**.

KOOPERATIVES LERNEN

Kooperatives Lernen unterscheidet sich vom peergestützten Lernen. Ein Unterschied zwischen den beiden Methoden liegt in der Größe der Lerngruppen. Peergestütztes Lernen erfolgt im Regelfall in Dyaden, kooperatives Lernen hingegen in Kleingruppen. Zum anderen unterscheidet sich die Zielsetzung der beiden Methoden. Peergestütztes Lernen verfolgt durch strukturierte Materialien und einer klaren Rollenverteilung das Ziel, Lerninhalte zu wiederholen und zu vertiefen. Die Aufgabenstellungen und Lösungsstrategien sind den Schülerinnen und Schülern bereits bekannt. Beim kooperativen Lernen hingegen wird Unterrichtsstoff von der Kleingruppe gemeinsam *erarbeitet*. Diese Erarbeitung ist dem peergestützten Lernen eher vorangestellt (Büttner et al., 2012). Beide Methoden verfolgen mehrere Ziele zugleich. So soll nicht nur die kognitive Aktivität gefördert, sondern auch die Motivation und Selbsterwartungshaltung gestärkt und die sozialen Kompetenzen ausgebaut werden (Büttner et al., 2012).

Kunter und Trautwein (2013) weisen darauf hin, dass kooperatives Lernen Fallstricke mit sich bringen kann (zum Beispiel Trittbrettfahren, Motivationsverluste, Koordinationsschwierigkeiten), wenn die Methode im Vorfeld nicht gründlich durchdacht ist. Folgende Aspekte gehören aus Sicht der Autoren zu den Gelingensbedingungen:

1. Verwendung echter Gruppenaufgaben: Echte Gruppenaufgaben sind additive Aufgaben, bei denen das Gruppenergebnis aus den verschiedenen Beiträgen der Gruppe besteht, oder konjunktive Aufgaben, bei denen es erst ein Ergebnis geben kann, wenn jedes Mitglied der Gruppe einen Teil dazu beigetragen hat.
2. Positive Interdependenz: Die Herstellung positiver Interdependenz bedeutet, dass die Gruppe gemeinsam einen Erfolg erreicht und nicht bloß eine Person, was auf Kosten (Misserfolg) anderer Gruppenmitglieder geht. Die wechselseitige Abhängigkeit ist folglich positiv besetzt und nicht negativ, wie zum Beispiel zwischen zwei kompetitiven Teams.

3. Individuelle Verantwortlichkeit: Für gelingendes kooperatives Lernen ist es notwendig, dass einzelne Gruppenmitglieder nicht die Möglichkeit haben, sich *nicht* aktiv am Gruppenprozess zu beteiligen. Jedes Mitglied der Gruppe benötigt eine individuelle Teilaufgabe, die bedeutsam und sichtbar ist, damit es zu einem Gruppenergebnis kommen kann.

„Kooperative Lernformen gehören zu den am besten erforschten Unterrichtsmethoden. Über viele hunderte Studien hinweg, die kooperatives Lernen in unterschiedlichen Altersgruppen [...] und Fächern untersucht haben, zeigt sich übereinstimmend die hohe Effektivität der Methode" (Kunter & Trautwein, 2013, S. 128).

Wirksamkeit

Ähnlich dem Vorgehen bei der Methode der Direkten Instruktion und des peergestützten Lernens wird im Folgenden die Wirksamkeit kooperativen Lernens dargestellt und analysiert. Nach der Darstellung allgemeiner Befunde wird die Wirksamkeit spezifisch im Bereich Mathematik fokussiert. Aufgrund der dünnen Forschungslage zum Einsatz im Bereich Mathematik muss hier auf Literatur vor dem Jahr 2000 zurückgegriffen werden.

Hattie und Kollegen (2013) ermitteln für die Methode des kooperativen Lernens im Vergleich zum Lernen in heterogenen Klassen eine Effektstärke von $d = .41$, im Vergleich zum individuellen Lernen eine Effektstärke von $d = .59$ und im Vergleich zum kompetitiven Lernen eine Effektstärke von $d = .54$. In der „Hitliste" verortet sich das kooperative Lernen zwei Plätze vor der Methode der Direkten Instruktion auf Platz 24. Einen weiteren allgemeinen Überblick über die Wirksamkeit von zehn verschiedenen Programmen zum kooperativen Lernen geben Johnson, Johnson und Stanne (2000) in ihrer Metaanalyse.

Slavin, Lake und Groff (2009) analysieren in ihrem Review die Wirksamkeit von Förderprogrammen im Bereich Mathematik für die Klassen fünf bis 12 (Middle and High School) anhand der Kriterien des „What Works Clearinghouse". Zwei Programme werden in den Bereich „Strong Evidence for Effectiveness" eingeordnet, die beide dem kooperativen Lernen zugeordnet werden: Das Programm IMPROVE ($d_{\text{gewichtet}} = .52$) besteht aus zehn Übungseinheiten, in denen Schülerinnen und Schüler in heterogenen Lerngruppen metakognitive Strategien üben. Die Lernentwicklung wird formativ begleitet und die Lerninhalte angepasst (siehe auch Schneider u. a., 2013, S. 232). Beim Student Teams-Achievement Divisions (STAD) ($d_{\text{gewichtet}} = .42$) bestehen die heterogenen Lerngruppen aus vier Schülerinnen und Schülern. Die Gruppen-Arbeitsphasen sind gepaart mit lehrerzentrierten Einheiten und individuellem Lernen. Je nach Erfolgsquote in einem wöchentlichen Quiz werden die Gruppen belohnt. Den stärksten Zusammenhang zwischen der Häufigkeit kooperativen Lernens und mathematischen Kompetenzerwerbs scheint es zu geben, wenn die Methode nur in wenigen Stunden und dabei maximal in jeder zweiten Stunde eingesetzt wird (Smith, McKenna & Hines, 2014).

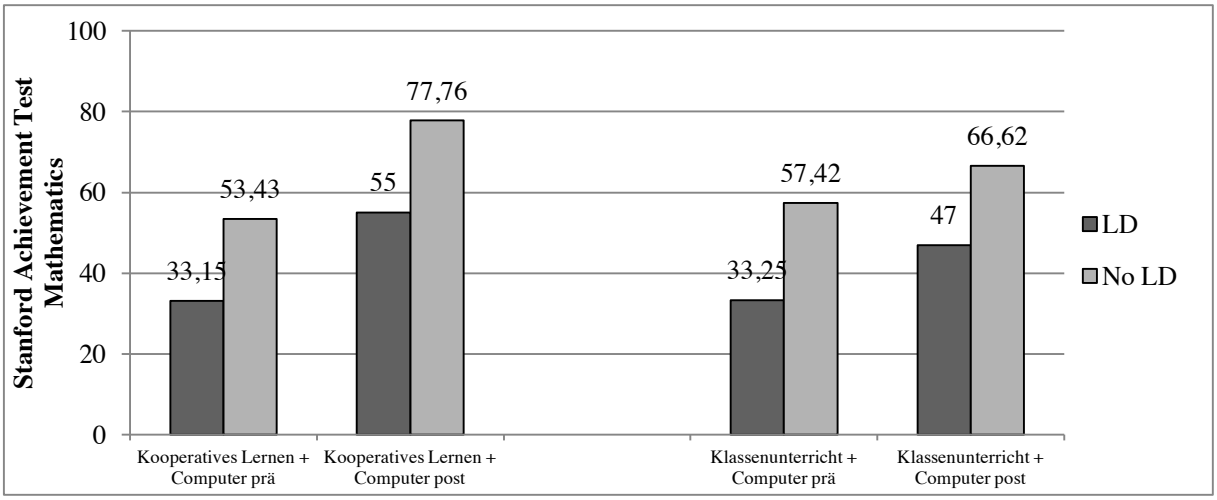
O'Connor und Jenkins (2013, S. 522) bemängeln, dass es nur sehr wenige Studien gibt, die das kooperative Lernen bei Schülerinnen und Schülern mit Lernschwierigkeiten bzw. Lernbeeinträchtigungen untersuchen. Domänenspezifisch zeigt sich auch hier eine bessere Forschungslage bezüglich Lesen und Schreiben als Mathematik.

McMaster und Fuchs (2002) analysieren die Wirksamkeit des kooperativen Lernens bei Schülerinnen und Schülern mit Lernbeeinträchtigungen bezüglich der Leistungen in den Kulturtechniken. Von den 15 identifizierten Studien beziehen sich fünf auf den Bereich Mathematik. Alle fünf Studien verzeichnen dabei eine Veröffentlichung vor dem Jahr 2000.

Eine Studie stammt dabei von Xin (1996a). Er vergleicht die Wirksamkeit kooperativen Lernens mit dem regulären Klassenunterricht. Sowohl im Setting des kooperativen Lernens sowie im Klassenunterricht erfolgen einzelne Sequenzen in Form einer Computerförderung. Die Stichprobe besteht aus 118 Drittklässlerinnen und -klässlern, von denen 25 eine Lernbeeinträchtigung (LD) aufweisen. Die randomisierte Verteilung auf die Gruppen ist wie folgt: $n_{LD+Koop.+Computer} = 13$, $n_{LD+Klasse+Computer} = 12$, $n_{No\text{-}LD+Koop.+Computer} = 46$, $n_{No\text{-}LD+Klasse+Computer} = 47$.

In der Gruppe des kooperativen Lernens werden zu Beginn organisatorische Rahmenbedingungen (Gruppengröße, Rotationsprinzip, Methode kooperatives Lernen) besprochen und erarbeitet. Danach durchlaufen die Schülerinnen und Schüler in den Unterrichtsstunden folgende Schritte: 1. Peergestütztes Arbeiten am Computer, 2. Ergebnisabgleich in der 4er-Kleingruppe, 3. wöchentliche Kompetition mit den anderen Teams. In der Gruppe des Klassenunterrichts erfolgt die erste Phase in den Unterrichtsstunden zum Thema Problemlösen lehrerzentriert (30 Minuten). Anschließend erfolgt Einzelarbeit am Computer unter Einbezug von Arbeitsblättern (20 Minuten). Die Intervention läuft im Mathematikunterricht über 20 Wochen mit vier Stunden pro Woche.

ABBILDUNG 35 WIRKSAMKEIT KOOPERATIVES LERNEN VS. KLASSENUNTERRICHT IN KLASSE 3

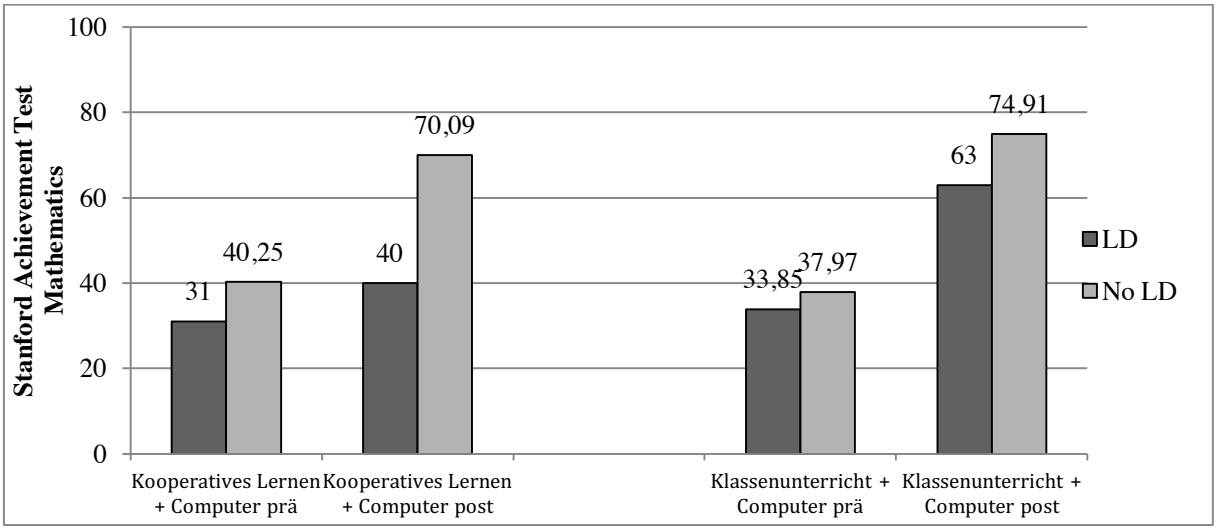


Die Ergebnisse (siehe Abbildung 35) zeigen, dass die Schülerinnen und Schüler mit Lernbeeinträchtigung stärker von der Methode des kooperativen Lernens profitieren ($d_{korr.} = .35$; eigene Berechnung). Bei den Schülerinnen und Schülern ohne Lernbeeinträchtigung liegt die Effektstärke des kooperativen Lernens bei $d_{korr.} = .60$ (eigene Berechnung). Die Entwicklung (Prä-Post-Vergleich) ist beim kooperativen Lernen sowohl bei denjenigen mit und ohne Lernbeeinträchtigungen signifikant. Im Zwischengruppenvergleich erzielt die Gruppe des kooperativen Lernens im Posttest signifikant bessere Ergebnisse als die Schülerinnen und Schüler im Klassenunterricht ($p = .042$).

Xin (1996b) führt eine zweite Studie durch, in der er die Wirksamkeit von kooperativem Lernen mit der des Einzelunterrichts bei Viertklässlerinnen und -klässlern vergleicht. Seine Stichprobe besteht aus 92 Schülerinnen und Schülern, von denen 16 eine Lernbeeinträchtigung aufweisen.

Das Treatment ist das gleiche wie in seiner ersten Studie (siehe oben), sodass es auch hier zwei Gruppen (Kooperatives Lernen + Computer vs. Klassenunterricht + Computer) gibt. Die randomisierte Verteilung auf die Gruppen ist wie folgt: $n_{LD+Koop.+Computer} = 6$, $n_{LD+Klasse+Computer} = 6$, $n_{No-LD+Koop.+Computer} = 43$, $n_{No-LD+Klasse+Computer} = 38$.

ABBILDUNG 36 WIRKSAMKEIT KOOPERATIVES LERNEN VS. KLASSENUNTERRICHT IN KLASSE 4



Die Ergebnisse in der zweiten Studie von Xin (1996b) zeigen, dass die Schülerinnen und Schüler mit Lernbeeinträchtigung in Klasse vier offenbar *nicht* stärker von der Methode des kooperativen Lernens profitieren ($d_{korr.} = -.99$; eigene Berechnung). Auch die Schülerinnen und Schüler ohne Lernbeeinträchtigung profitieren nicht stärker von dieser Methode ($d_{korr.} = -0.42$; eigene Berechnung). Xin (1996b) macht in dieser zweiten Studie keine konkreten Angaben zu den Signifikanzen im Prä-Post-Vergleich. In der Frage, welche der beiden Methoden insgesamt wirksamer ist, kann der Autor im Vergleich keine signifikanten Effekte feststellen. Berücksichtigt werden muss bei dieser zweiten Studie, dass die Stichprobe insgesamt geringer ist als in der ersten. Xin (1996b) vermutet, dass die Wirksamkeit der Methode im Laufe der Schuljahre sinkt, wie auch Kunsch, Jitendra und Sood (2007) sowie Rohrbeck und Kollegen (2003) für das peergestütze Lernen feststellen.

Eine weitere relevante Studie, die auch vor dem Jahr 2000 veröffentlicht ist, stammt von Stevens und Slavin (1995). Die Autoren begleiten über zwei Jahre zwei Grundschulen, die sich fächerübergreifend verstärkt auf die Methode des kooperativen Lernens fokussieren. Im Fach Mathematik wird dabei das TAI(Team-Accelerated Instruction)-Programm eingesetzt, das sowohl aus Einzel- als auch Kleingruppenarbeitsphasen besteht. Die Stichprobe besteht aus 1.012 Schülerinnen und Schüler von der zweiten bis zur sechsten Klasse. 21 Klassen stellen die Experimentalgruppe dar und 24 Klassen aus drei weiteren Schulen die Kontrollgruppe. Darunter befinden sich 40 (in der Experimentalgruppe) und 36 (in der Kontrollgruppe) Schülerinnen und Schüler mit Lernbeeinträchtigungen. Die Experimentalgruppe wird im Prä-Post-Design sowohl nach dem ersten als auch nach dem zweiten Versuchsjahr diagnostisch begleitet. Zur Erhebung der mathematischen Kompetenzen wird der California Achievement Test eingesetzt. Die Ergebnisse zeigen, dass die Effektstärken bei den Schülerinnen und Schülern mit Lernbeeinträchtigungen nach einem Jahr (Vortestunterschiede liegen nicht vor) $d = .33$ (computation) bzw. $d = .16$ (application) betragen und nach zwei Jahren $d = .59$ (computation) bzw. $d = .35$ (appli-

cation). Dabei sind lediglich die Effekte nach dem zweiten Jahr im Vergleich zur Kontrollgruppe signifikant; in diese Berechnung werden die Schülerinnen und Schüler mit Lernbeeinträchtigungen aller Klassenstufen einbezogen.

O'Melia und Rosenberg (1994) untersuchen eine weitere Modifikation des kooperativen Lernens durch Hausaufgabenteams. An ihrer Untersuchung nehmen 171 Schülerinnen und Schüler mit Lern- und/oder Verhaltensproblemen aus den Klassenstufen sechs, sieben und acht teil. 86 Schülerinnen und Schüler bilden die Experimentalgruppe und 85 die Kontrollgruppe. In der Experimentalgruppe werden die Schülerinnen und Schüler in leistungsheterogene Kleingruppen eingeteilt. Sie erhalten über acht Wochen am Ende der Mathematikstunde eine Hausaufgabe, die sie zu Hause erst einmal selbst lösen sollen. In der nächsten Mathematikstunde trifft sich die Kleingruppe für ca. zehn Minuten und der sog. „Checker" korrigiert die Hausaufgaben. Fehler werden dann innerhalb der Kleingruppe korrigiert, in der sich die Schülerinnen und Schüler gegenseitig unterstützen. Am Ende der Woche wird die beste Kleingruppe gekürt. Die Kontrollgruppe genießt den regulären Mathematikunterricht. Im Prä-Post-Vergleich zeigen sich keine signifikanten Unterschiede in den mathematischen Kompetenzen (California Achievement Test) (d = .14; eigene Berechnung), doch machen die Schülerinnen und Schüler häufiger (d = .41; eigene Berechnung) und gründlicher (mit weniger Fehlern) (d = .30; eigene Berechnung) ihre Hausaufgaben.

Fazit

Nach der Klassifikation Evidenzbasierter Praxis des Council for Exceptional Children (2014) liegen für die Methode des kooperativen Lernens im Bereich der Förderung mathematischer Basiskompetenzen bei Schülerinnen und Schülern mit Lernbeeinträchtigungen vier Gruppenvergleichsstudien vor, die alle mehr als 60 Probanden aufweisen (O'Melia & Rosenberg, 1994; Stevens & Slavin, 1995; Xin, 1996a, 1996b). Xin (1996a) berichtet von einer kleinen Effektstärke von d = .35 für die Förderung durch kooperatives Lernen, zusätzlich unterstützt durch einen Computer. Auch Stevens und Slavin (1995) berichten von kleinen bis mittleren Effekten nach einer zweijährigen Implementation des kooperativen Lernens für lern- bzw. verhaltensschwierige Schülerinnen und Schüler. Die Wirksamkeit der Methode wird unterstützt durch das Review von Slavin, Lake und Groff (2009), in dem die Autoren die IMPROVE- und STAD-Konzepte als „Strong Evidence for Effectiveness" verorten. Jedoch berichtet Xin (1996b) in seiner zweiten Studie eine negative Wirksamkeit für die Förderung durch kooperatives Lernen und dem Computer (d = –.99) in der vierten Jahrgangsstufe. Keinen Effekt bezüglich der mathematischen Kompetenzen finden O'Melia und Rosenberg (1994) in ihrer Studie zum kooperativen Lernen durch Hausaufgabenteams.

Die Analyse der Wirksamkeit wird massiv dadurch erschwert, dass speziell für das Thema des kooperativen Lernens im Fach Mathematik bei Schülerinnen und Schülern mit Lernschwierigkeiten keine jüngeren Studien aufzufinden sind (siehe z. B. Büttner u. a., 2012). Diejenigen identifizierten Studien evaluieren zum Teil Modifikationen der Methode (zum Beispiel durch das zusätzliche Hinzuziehen des Computers oder der Ausweitung auf Hausaufgabenteams) oder berufen sich auf ein Programm (TAI). An dieser Stelle eröffnet sich ein großes Forschungsdesiderat. Insbesondere aufgrund der negativen Effekte bei Xin (1996b) und des uneindeutigen Effekts bei O'Melia und Rosenberg (1994) liegt entsprechend dieser Analyse keine eindeutig evidenzbasierte Methode vor.

Das Fazit der Wirksamkeitsanalyse nach dem Council for Exceptional Children (2014) für die Methode des kooperativen Lernens lautet: **Gemischte Evidenz**.

FEEDBACK

Formen und Möglichkeiten des Feedbacks sind sehr vielfältig. Unterschieden wird, *wer* das Feedback (Schüler, Eltern, Lehrer) *wem* bezüglich *was* (Lernen, Verhalten) gibt oder erhält. Hattie und Kollegen (2013, S. 206) definieren anschaulich, was allgemein unter Feedback zu verstehen ist:

> „Zusammengefasst lässt sich sagen, dass Feedback eine Information ist, die von einem Akteur (zum Beispiel Lehrperson, Peer, Buch, Eltern oder die eigene Erfahrung) über Aspekte der eigenen Leistung oder des eigenen Verstehens gegeben wird."

Mittlerweile gibt es viele Forschungen, die sich mit den Kriterien guten Feedbacks beschäftigen. In ihrer bekannten Metaanalyse „The Power of Feedback" stellen Hattie und Timperley (2007) fest, dass Schülerfeedback dann erfolgreich ist, wenn dadurch drei Fragen beantwortet werden: 1. Wo stehe ich jetzt? (Feed Back) 2. Wo gehe ich hin? (Feed Up) 3. Wie komme ich dahin? (Feed Forward) (siehe auch Kunter & Trautwein, 2013, S. 98). Zudem schreibt Strahm (2008, S. 55), dass sich erfolgreiches (im Sinne von konstruktives) Feedback dadurch auszeichnet, dass es konkret, präzise und unmittelbar gegeben wird.

Im Gegensatz zum einfachen Feedback, das beinhaltet, ob etwas richtig oder falsch gelöst wird, enthält das „elaborierte Feedback" Hinweise darauf, wie etwas richtig oder besser gemacht werden kann (Hilfestellungen, Strategien). Das elaborierte Feedback hat zudem auch den Lernprozess im Blick und nicht nur ein einziges Ergebnis (Kunter & Trautwein, 2013, S. 98).

Die Methode des Feedbacks wird insbesondere in den Metaanalysen von Baker, Gersten und Lee (2002) und Gersten und Kollegen (2009) berücksichtigt. In einem Großteil der darin berücksichtigten Studien handelt es sich um die Methode des curriculumbasierten Messens, die neben der Funktion des Gewinnens diagnostischer Informationen vor allem die Funktion des Feedback-Gebens hat. Curriculumbasiertes Messen ist in Kapitel 4.2.2.4 im Rahmen des Responsiven Handlungsmodells unter der Überschrift Lernverlaufsdiagnostik bereits umfassend beschrieben, sodass darauf an dieser Stelle nicht weiter eingegangen wird.

Der Fokus liegt im Folgenden also auf der Form des Feedbacks von Lehrkräften an Schülerinnen und Schüler (Schülerfeedback) über ihre Lernleistung bzw. -entwicklung, was ein Feedback an die Lehrkraft über die Wirksamkeit ihres Unterrichts impliziert (Voß & Hartke, 2014, S. 85).

Unter anderem fassen Hattie und Kollegen (2013, S. 433) curriculumbasiertes Messen unter dem Oberbegriff der formativen Evaluation des Unterrichts. Die Begriffe „formative assessment" oder „formative evaluation" sind im angloamerikanischen Sprachgebrauch durchaus verbreitet (vgl. z. B. Keeley & Tobey, 2011) und gehen auf die Arbeiten Scrivens (1967) zurück, der erstmals die summative und formative Leistungsevaluation voneinander trennt. Deno (1985, 2003) greift mit seiner Forschergruppe zwei Jahrzehnte später Scrivens Konzept der formativen Leistungsevaluation auf und entwickelt es bis heute weiter zum curriculumbasierten Messen.

WIRKSAMKEIT

Analysiert wird im Folgenden die Wirksamkeit des curriculumbasierten Messens als eine Form von Feedback. Hattie und Kollegen (2013) verorten die formative Evaluation des Unterrichts in ihrer domänenübergreifenden Megaanalyse mit einer hohen Effektstärke von $d = .90$ auf Platz 3 von 138. In der spezifischen Kategorie „Unterrichten", in der die Methode der Direkten Instruktion zum Beispiel Platz 12 einnimmt, platziert sich die formative Evaluation des Unterrichts an erster Stelle. Grundlage für die Analyse bilden zwei Metaanalysen mit insgesamt 30 Studien. Einen weiteren allgemeinen Überblick über die Wirksamkeit curriculumbasierten Messens bei Schülerinnen und Schülern mit und ohne Lernbeeinträchtigungen ermöglichen Lembke und Stecker (2007) sowie Stecker, Fuchs und Fuchs (2005).

Spezifisch für den Einsatz im Bereich Mathematik ermitteln Baker, Gersten und Lee (2002, S. 60) in ihrer Metaanalyse eine mittlere Effektstärke von $d = .57$ für die Rückmeldung der individuellen Lernentwicklung an die Schülerinnen und Schüler selbst. Erhalten die Lehrkräfte zusätzlich datenbasiert Hinweise auf die Modifikation oder die Implementation einer Förderung (zum Beispiel durch ein Computerprogramm) liegt die Wirksamkeit bei $d = .51$. Gersten und Kollegen (2009) ermitteln für das Feedback an Schülerinnen und Schüler über ihre Lernentwicklung mittels grafischer Visualisierung eine Effektstärke von $d = .23$, für solches an Lehrkräfte eine Effektstärke von $d = .21$. Erhalten die Lehrkräfte zusätzlich Hinweise zur Modifikation des Fördersettings, liegt die Wirksamkeit bei $d = 34$. In der Metaanalyse von Methe und Kollegen (2012) über Einzelfallstudien zur Förderung mathematischer Kompetenzen liegt die Effektstärke des Einsatzes von CBM bei $d = .79$.

Zur differenzierten Analyse der Wirksamkeit im Bereich Mathematik bei Schülerinnen und Schülern mit Lernschwierigkeiten werden im Folgenden Studien dargestellt, die nach dem Jahr 2000 veröffentlicht sind. Ausgeschlossen werden in der folgenden Analyse diejenigen Studien, die sich ausschließlich mit der Güte curriculumbasierter Messverfahren beschäftigen.

Shapiro, Edwards und Zigmond (2005) implementieren bei 120 Schülerinnen und Schülern mit Lernbeeinträchtigungen (special education students) der Klassenstufen zwei bis sechs wöchentliche curriculumbasierte Messungen im Bereich Rechenoperationen (computation) und bei 102 Schülerinnen und Schülern curriculumbasierte Messungen im Bereich Anwendungsaufgaben (application). Der Einsatz erstreckt sich über neun Monate. Die Kurztests im Bereich Rechenoperationen bestehen aus 25 Aufgaben zu den Grundrechenarten und die Kurztests zu den Anwendungsaufgaben aus 18 bis 24 Aufgaben u. a. zu den Themen Größen und Längen, Diagramme oder Geld. Insgesamt gibt es fünf Anforderungsniveaus, in die die Schülerinnen und Schüler zu Beginn, unabhängig von ihrer Klassenstufe, je nach individuellen Kompetenzen eingestuft werden. Die Ergebnisse zeigen, dass die Schülerinnen und Schüler mit den curriculumbasierten Messungen sowohl im Bereich Rechenoperationen als auch im Bereich Anwendungsaufgaben einen wöchentlichen Zuwachs von 0,38 Aufgaben erzielen. Tabelle 25 gibt Aufschluss über die Zuwachsraten je nach Anforderungsniveau. Die Schülerinnen und Schüler, die auf dem höchsten Anforderungsniveau eingeordnet sind, scheinen lediglich noch Zuwächse im Bereich der Anwendungsaufgaben zu verzeichnen. In der Betrachtung der Ergebnisse im Bereich Rechenoperationen unter Berücksichtigung der Diskrepanz zwischen Klassenstufe und Anforderungsniveau haben die Autoren zudem herausgefunden, dass das curriculumbasierte Messen bei einer Differenz von 0, 1 und 2 (z. B. Klasse 3 und Niveau 1, 2 oder 3) sehr wirksam

ist, sich die Wirksamkeit jedoch bei zu großen Differenzen (< 2) verringert. Im Bereich der Anwendungsaufgaben scheint die Methode lediglich bei einer Diskrepanz von 5 (Klasse 5, Niveau 1) weniger wirksam zu sein.

TABELLE 25 ZUWACHS JE NACH ANFORDERUNGSNIVEAU (SHAPIRO ET AL., 2005, S. 23)

Niveau	n	Wöchentlicher Zuwachs Rechenoperation	n	Wöchentlicher Zuwachs Anwendungsaufgaben	SD
1	15	+ 0,32	-	-	-
2	42	+ 0,28	47	+ 0,36	0.20
3	50	+ 0,43	49	+ 0,37	0.20
4	10	+ 0,72	10	+ 0,44	0.25
5	3	+ 0,00	3	+ 0,52	0.08

66 % der Schülerinnen und Schüler im Bereich Rechenoperationen und 37 % derjenigen mit Anwendungsaufgaben entwickeln sich während der Studie auf einem Niveau, das in vorherigen Studien für Schülerinnen und Schüler *ohne* Lernbeeinträchtigung ermittelt wurde (Zuwachs von ca. + 0,50 Aufgaben pro Woche). Auch im Vergleich mit Zuwachsraten bezogen auf Klassenstufen, ordnen sich die Ergebnisse in bisherige Erkenntnisse ein: Bei Hosp, Hosp und Howell (2007, S. 109) liegen die wöchentlichen Zuwachsraten für Schülerinnen und Schüler der Allgemeinen Schule bei + 0,30 (Klasse 1), + 0,30 (Klasse 2), + 0,30 (Klasse 3), + 0,70 (Klasse 4), + 0,75 (Klasse 5), + 0,45 (Klasse 6).

Eine weitere Studie zur Wirksamkeit curriculumbasierten Messens stammt von Lambert, Algozzine und Mcgee (2014). Die Autoren setzen das Programm „Accelerated Math" (AM) ein, das auch positiv durch das „What Works Clearinghouse" analysiert ist und vom National Center on Response to Intervention zum Einsatz empfohlen wird. AM gibt Schülerinnen und Schülern auf Grundlage der eingegebenen Schülerdaten einen Überblick über ihre Lernentwicklung und Lehrkräften zudem Hinweise zum benötigten Fördersetting. An der Untersuchung nehmen insgesamt 762 Schülerinnen und Schüler von der zweiten bis zur fünften Jahrgangsstufe teil, die aufgrund unterschiedlicher Faktoren von Schulversagen bedroht sind (at risk-students). Zu den Gründen gehört, ob die Schülerinnen und Schüler kostenfrei Mittagessen bekommen, einen geringen sozioökonomischen Hintergrund aufweisen oder einen schulischen Unterstützungsbedarf haben. Die Schülerinnen und Schüler werden pro Jahrgang randomisiert in eine Experimental- und Kontrollgruppe eingeordnet. Alle Schülerinnen und Schüler führen im Herbst, Winter und im Frühjahr eines Schuljahrs curriculumbasierte Messungen durch (STAR Math & *Terra Nova*). Lediglich in der Experimentalgruppe werden die Ergebnisse durch das Accelerated-Math-Programm evaluiert. Die Ergebnisse zeigen sowohl im STAR Math- sowie im *Terra Nova-Programm*, dass sich die Schülerinnen und Schüler der Experimentalgruppen stärker entwickeln als diejenigen in der Kontrollgruppe. Die Effektstärken liegen bei $d_{EG} = .79$ und $d_{KG} = .32$. Die Schülerinnen und Schüler in der Jahrgangsstufe zwei und fünf entwickeln sich dabei weniger stark, was durch Motivationsgründe oder auch Aspekte erster Erfahrungen mit Testdurchführungen begründet wird. Die Wirksamkeit des Accelerated-Math-Programms, als Methode des curriculumbasierten Messens, gerade für rechenschwache Schülerinnen und Schüler wird zudem durch eine repräsentative Stichprobe durch Spicuzza und Kollegen (2001) belegt.

Auch Allinder und Kollegen (2000) vergleichen die Wirksamkeit curriculumbasierten Messens mit und ohne strukturierter Anleitung zur Auswertung der Schülerdaten durch die Lehrkräfte.

Dafür teilen sie Lehrkräfte in drei Gruppen ein. 12 Lehrkräfte bilden eine Kontrollgruppe. Zehn Lehrkräfte führten zweimal in der Woche curriculumbasierte Messungen durch und überprüfen zweimal im Monat, ob das gesetzte Jahresziel erreicht werden kann und ob eine Modifikation des Unterrichts und/oder einer Förderung stattfinden muss. Dafür können die Lehrkräfte die Hinweise des Computerprogramms „Monitoring Basic Skills" (MBSP) berücksichtigen. Weitere acht Lehrkräfte hingegen führen ebenso die curriculumbasierten Messungen durch, analysieren die Daten jedoch detaillierter nach einer vorgegebenen Schrittfolge und folgenden Kernfragen:

1. On what skill(s) has the student done well in the preceding 2 weeks?
2. On what skill(s) has the student improved compared to the previous 2-week-period?
3. How will the teacher attempt to improve student performance on the target skill(s)?

Insgesamt befinden sich 54 Schülerinnen und Schüler mit Lernbeeinträchtigungen aus der vierten Jahrgangsstufe in der Stichprobe. Jede Lehrkraft verfügt in der Studie folglich über ein bis zwei Schüler. Die Ergebnisse zeigen, dass sich die Schülerinnen und Schüler der Kontrollgruppe als auch diejenigen mit den curriculumbasierten Messungen (ohne strukturierte Anleitung) gleich stark im Math Computation Test-Revised (MCTR) entwickeln. Signifikante Unterschiede hingegen gibt es zu den Schülerinnen und Schülern, deren Lehrkräfte die CBM-Daten strukturiert auswerten.

Stecker und Fuchs (2000) begleiten die Leistungsentwicklung von 84 Schülerinnen und Schülern mit Lernbeeinträchtigungen in den Jahrgängen zwei bis acht unter Berücksichtigung curriculumbasierter Messverfahren. 22 Förderschullehrkräfte entwickeln über 20 Wochen für 42 Schülerinnen und Schülern Förderangebote, die sich von dem Leistungsstand der curriculumbasierten Messverfahren ableiten. Die weiteren 42 Schülerinnen und Schüler bearbeiten keine curriculumbasierten Messverfahren und erhalten dieselben Angebote. Die mathematischen Kompetenzen unterscheiden sich im Vortest nicht. Die Schülerinnen und Schüler, bei denen eine Passung zwischen CBM und Förderung stattfindet, zeigen nach der Untersuchung im Mathematics Operations Test-Revised (MOT-R) signifikant bessere Leistungen.

Stecker, Fuchs und Fuchs (2005) zeigen in ihrem Review, dass die Implementation curriculumbasierten Messens erst dann zu Leistungsverbesserungen der Schüler führt, wenn die Daten grafisch aufgearbeitet, nach einem bestimmten System interpretiert und dann als Feedback an die Schülerinnen und Schüler gegeben werden, damit negative Entwicklungstrends vermieden werden. Dass die grafische Rückmeldung zwar im Entscheidungsprozess über Fördermaßnahmen unterstützend, aber nicht ausreichend ist, um eine Leistungsverbesserung zu erzielen, zeigen auch Graney und Shinn (2005).

Die Forschungslage zum curriculumbasierten Messen ist insgesamt sehr umfangreich und erfährt beständig eine inhaltliche Ausweitung. Lowenthal und Basset (2012) stellen die Relevanz für den Einsatz curriculumbasierter Messungen für gelingende Transitionsprozesse heraus, Foegen (2006) zeigt zudem, dass der Einsatz curriculumbasierter Messungen auch in höheren mathematischen Kontexten wie Algebra erfolgreich sein kann. Die Forschungen konzentrieren sich jedoch insgesamt eher auf den Bereich der allgemeinen Schule, weniger auf den Bereich bei Lernbeeinträchtigungen, wo das Fach Mathematik dann den geringsten Anteil einnimmt, was durch einschlägige Reviews und Metaanalysen belegt wird (Foegen, Jiban & Deno, 2007).

FAZIT

Nach der Klassifikation Evidenzbasierter Praxis des Council for Exceptional Children (2014) liegen für die Methode des Feedbacks, hier speziell in Form curriculumbasierter Messungen, im Bereich der Förderung mathematischer Basiskompetenzen bei Schülerinnen und Schülern mit Lernbeeinträchtigungen drei gut fundierte randomisierte Gruppenvergleichsstudien mit überaus positiven Effekten und jeweils 60 Probanden vor (Lambert et al., 2014; Shapiro et al., 2005; Stecker & Fuchs, 2000). Zudem werden weitere Studien mit entsprechenden Kriterien, jedoch mit geringeren Stichproben identifiziert, die ebenso von einer positiven Wirksamkeit berichten (Allinder et al., 2000). Curriculumbasiertes Messen scheint zudem in weiteren Kontexten wie Transitionen (Lowenthal & Bassett, 2012) oder auch weiterführenden mathematischen Inhalten (Foegen, 2006) unterstützend zu wirken. Es wird keine Studie identifiziert, die von einer negativen Wirksamkeit berichtet. Die Methode des Feedbacks durch CBM kann als evidenzbasiert deklariert werden (vgl. Tabelle 16).

Das Fazit der Wirksamkeitsanalyse nach dem Council for Exceptional Children (2014) für die Methode des Feedbacks lautet: **Evidenzbasierte Praxis**.

4.3.3.3 Zusammenfassung und Forschungsdesiderata

SCHRITT 7: ERGEBNIS: LISTE EVIDENZBASIERTER VERFAHREN

Im vorherigen Kapitel sind die als zuvor potentiell evidenzbasierten identifizierten Methoden der Direkten Instruktion, des peergestützten Lernens, des kooperativen Lernens und des Feedbacks eingehend bezüglich ihrer Wirksamkeit entsprechend der Kriterien des Council for Exceptional Children (2014) analysiert worden. Die dafür herangezogenen Studien sind im Regelfall nach dem Jahr 2000 veröffentlicht, betrachten (auch) Schülerinnen und Schüler mit Lern- oder Rechenschwierigkeiten und fokussieren die Wirksamkeit in der Domäne Mathematik.

Folgende Einstufung kann vorgenommen werden:

1. Direkte Instruktion:	Evidenzbasierte Praxis
2. Peergestütztes Lernen:	Gemischte Evidenz
3. Kooperatives Lernen:	Gemischte Evidenz
4. Feedback:	Evidenzbasierte Praxis

Für keine der vier Methoden wird eine ungenügende oder negative Evidenz festgestellt. Für die Methode der Direkten Instruktion und des Feedbacks ist die Forschungslage aktuell, umfassend und wenig widersprüchlich. Keine identifizierte Studie berichtet von negativen Effekten. Es profitieren offenbar sowohl lernstarke als auch lernschwache Schülerinnen und Schüler von der Methode. Weniger eindeutig zeigt sich die Forschungslage bei dem peergestützten sowie kooperativen Lernen. Die meisten Studien sind zwischen 1980 und 2000 veröffentlicht, da es zu dieser Zeit offenbar einen Forschungsboom in diesen Themen gibt. Die vorliegenden aktuellen Studien sind sehr uneinheitlich und berichten unterschiedlich von positiven, neutralen oder gar

negativen Effekten. Dennoch überwiegt eindeutig die Anzahl an Studien mit positiven Effekten. Allen Methoden gemeinsam ist, dass die Wirksamkeit steigt, je akkurater und fundierter sie von der Lehrkraft durchgeführt wird.

Im Anschluss an die Analysen können Forschungsdesiderata festgestellt werden. Methoden und Konzepte verändern ihre Wirksamkeit bei sich wandelnden Klassensettings und Schulsystemen. Erkenntnisse über eine Wirksamkeit aus dem Jahr 1980 können heute anders ausfallen. Im Zuge eines verstärkt inklusiven Bildungssystems müssen die Wirksamkeiten der Methoden beständig evaluiert werden. Verwiesen werden sollte an dieser Stelle an die Aufgabe der Wissenschaft, Erkenntnisse nicht nur zu gewinnen, sondern auch beständig zu überprüfen. Die Evaluation der Methoden sollte dabei nicht immer in Kombination mit weiteren Maßnahmen (zum Beispiel peergestütztes Lernen + Einbezug von Eltern, zusätzliche Computerförderung o. ä.) erfolgen. Zur Einschätzung der Evidenzen sind Erkenntnisse über die genuinen Methoden mit der Gewinnung isolierter Effekte hilfreich und angezeigt. Hierfür werden insbesondere Replikationsstudien gefordert (Mulcahy, Krezmien & Travers, 2012). Die Kombinationen mit weiteren Medien, Materialien oder Methoden, wie auch z. T. in den hier einbezogenen Studien, werden in Fragen notwendiger Grundlagenforschung kritisch gesehen. Eine derzeit zu wenig verfolgte Möglichkeit ist dabei, Lehrkräfte, die lernwirksamen Unterricht machen, bezüglich ihres Unterrichtsstils, der eingesetzten Methoden o. ä. zu befragen (Poplin et al., 2011; Whitaker, 2009).

Die Domäne Mathematik stellt insgesamt ein vernachlässigtes Forschungsgebiet dar. In den Bereich Lesen, Rechtschreiben oder auch emotional-soziale Entwicklung ist die Forschungslage umfassender. Die Wirksamkeit der Methoden sollte folglich nicht nur in inklusiven Settings, sondern vor allem in der schulischen Domäne Mathematik untersucht werden.

In einem weiteren Schritt müssen Forschungserkenntnisse für die schulische Praxis aufgearbeitet und verständlich zugänglich gemacht werden. Die „Grüne Liste Prävention“ ist ein wichtiger Schritt bezüglich der Verzahnung von Wissenschaft und Praxis, doch sind die Forscherkolleginnen und -kollegen in den Vereinigten Staaten weiter, die mittels Webseiten und Tutorials evidenzbasierte Methoden im Schulalltag implementieren (siehe Tabelle 17). Eine Untersuchung von Hintz und Kollegen (2015) zeigt zudem, dass sowohl Förderschul-lehrkräfte als auch Lehrkräfte der allgemeinen Schule (N = 116) ihr Wissen über wirksame Unterrichtsmethoden eher gering einschätzen und dabei zugleich einen persönlichen Bedarf an Informationen und Fortbildung sehen. Hier sind Wissenschaftlerinnen und Wissenschaftler aufgerufen, ihre Erkenntnisse, seien sie positiv, negativ oder uneindeutig, Praktikern zugänglich zu machen. Denn: „Overall, when teachers are using research- or evidence-based practices, they are using their time wisely“ (Lembke & Stormont, 2005, S. 762).

4.4 Implementation von Fördermaßnahmen

„Aufgrund der aktuellen Output-Orientierung des Bildungswesens und einer Häufung an Bildungsprogrammen und -reformen gewinnt die Implementationsforschung in der Unterrichtswissenschaft zunehmend an Bedeutung.“
(Oswald, Valkanover & Conzelmann, 2013, S. 256)

In der Frage der Implementation, also die Passung und Umsetzungstauglichkeit von Programmen oder Methoden in die schulische Praxis, können verschiedene Aspekte betrachtet werden.

Mögliche Kriterien sind nicht nur die Angemessenheit oder die Akzeptanz der Förderung aus Sicht der Lehrkräfte oder der Schülerinnen und Schüler, sondern auch die Machbarkeit oder die Übernahme der zugrundeliegenden Konzeption bei der Umsetzung. Die Beurteilung der Implementationsqualität wird bisher im Gros der Evaluationsstudien von Förderprogrammen oder -konzepten vernachlässigt. Berücksichtigung hat sie jedoch zum Beispiel gefunden bei der Evaluation des Programms „Calcularis" (Käser et al., 2013) oder auch bei der Evaluation der Methode des peergestützten Lernens bei Calhoon und Fuchs (2003).

Im Folgenden werden vorerst die Begrifflichkeiten im Kontext der Implementationsforschung vorgestellt. Im Anschluss werden die Bedeutsamkeit der Analyse der Implementationsqualität und mögliche Kriterien zur Erfassung diskutiert.

4.4.1 Begriffliches und Konzeption

Der Begriff der Implementation wird in der Literatur unterschiedlich definiert; dabei werden zugleich auch verschiedene Implementationsdimensionen betrachtet. Eine oftmals zitierte Definition stammt von Euler und Sloane (1998, S. 312): „‚Implementation soll verstanden werden als die Umsetzung von wissenschaftlichen Erkenntnissen in die gesellschaftliche Praxis". Petermann (2015b, S. 122; Herv. d. Verf.) stellt ca. 20 Jahre später in seiner Definition den Herausforderungscharakter von Implementation heraus, in dem er schreibt, dass sich die Implementationsforschung allgemein mit der „*Diskrepanz* zwischen der Planung und Verwirklichung von Konzepten" auseinandersetzt. Im Regelfall bezieht sich Implementation auf die Verwurzelung von Innovationen in die sog. pädagogische Praxis, „in a given organization or community setting" (Mihalic, 2004, S. 83). Tippelt und Tenorth (2007, S. 333; Herv. d. Verf.) bezeichnen als Implementation „die *dauerhafte* Einführung bestimmter Maßnahmen in bestehende Organisationsformen und -prozesse" und betonen damit, ebenso wie Oswald und Kollegen (2013, S. 256), die Langfristigkeit des Ansatzes. Es wird jedoch nicht näher bezeichnet, was unter „dauerhaft" zu verstehen ist. Die zeitliche Verortung der Erhebung von Implementationsqualität ist aktuell ein Aspekt, über den in der Fachliteratur noch Uneinheitlichkeit besteht. So wird nur von wenigen Autoren erwähnt, dass die Implementations*qualität* auch in ersten Erprobungen im Feld erhoben werden sollte, bevor die allgemeine Implementation und Vermarktung in die pädagogische Praxis zum Beispiel durch Verlage oder Fortbildungsinstitute verfolgt wird, was eher der Erhebung der Implementations*quantität* nahekommt (Petermann, 2015b, S. 126). „Weitere Forschung ist nötig, um die Frage nach dem angemessenen Messzeitpunkt zu klären. Diese Klärung ist wichtig, da förderliche und hinderliche Faktoren der Implementierung nur durch klare Vorgaben präzise aufgedeckt werden können" (Petermann, 2015b, S. 125).

In dem methodischen Zugang zur Erhebung der Implementationsqualität wird unterschieden zwischen einer „top-down"-Strategie und einer „symbiotischen" Strategie (Gräsel & Parchmann, 2004; Petermann, 2015b). Bei der top-down-Strategie wird eine Innovation von einer externen Instanz geplant und hierarchisch „von oben herab" in die pädagogische Praxis eingeführt und verbreitet. Oftmals wird dafür als Beispiel ein neues Curriculum genannt, das durch das Kultusministerium erarbeitet und verpflichtend erlassen wird. „Erfolgreich" implementiert ist die Innovation (hier das Curriculum) dann, wenn die Umsetzung im Unterricht möglichst wie vorgesehen erfolgt. Erst nach erfolgter Umsetzung wird dann die Wirksamkeit

evaluiert. Bei der symbiotischen Strategie hingegen arbeiten verschiedene Akteure aus unterschiedlichen Bereichen zusammen an einer Innovation. So kooperieren Wissenschaftler mit Lehrkräften und Personen aus der Bildungspolitik auf Grundlage von Fragen und Problemen der pädagogischen Praxis. In Rahmen dieser Netzwerke werden zum Beispiel Unterrichtskonzepte oder -reihen erarbeitet. „Erfolgreich" implementiert ist die Innovation dann, wenn a) die Kooperation zwischen den beteiligten Entwicklern als gelungen bezeichnet werden kann, b) wenn sich die beschriebenen Ausgangsprobleme verbessert haben und c) wenn weitere Akteure (zum Beispiel Schulen) in Folge der Verbreitung der Innovation aktiv werden. Reinmann-Rothmeier und Mandl (1998, S. 309) haben auf Grundlage von Erfahrungen aus einem eigenen Forschungsprojekt ein Implementationsmodell aufgestellt, das die verschiedenen Akteure in Fragen der Implementation berücksichtigt.

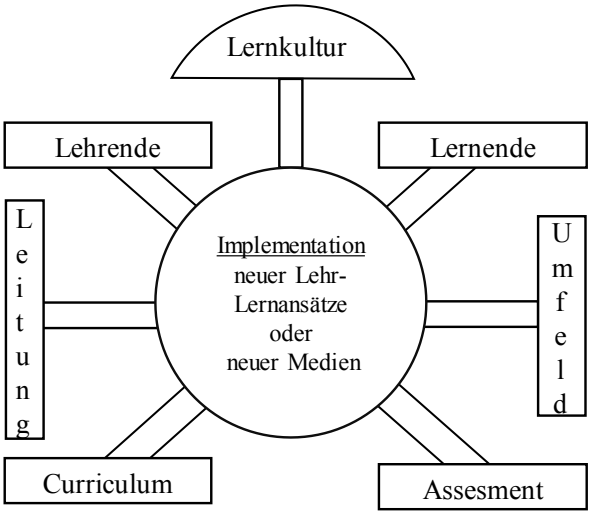


ABBILDUNG 37 AKTEURE VON IMPLEMENTATION (REINMANN-ROTHMEIER & MANDL, 1998, S. 308)

Reinmann-Rothmeier und Mandl (1998) verstehen die Implementation von neuen Lehr- und Lernansätzen oder auch von neuen Medien im Kontext eines systemischen Wandels. Der Wandel kann sich in mehreren Bereichen vollziehen: zum Beispiel auf curricularer Ebene oder auch in Fragen der Leistungsmessung, dem sog. Assessment. Im Kontext dieser Veränderungen besteht die Notwendigkeit von Innovationen und neuen Ideen bzw. Ansätzen, um dem systemischen Wandel begegnen zu können. Von diesen Innovationen, so die Autoren, sollten die beteiligten Akteure überzeugt und ausreichend informiert sein. Als Akteure verstehen die Autoren Leitungspersonen (Schulleitung, Hochschulleitung), die Lehrenden (z. B. Lehrkräfte), die Lernenden und auch Personen des Umfelds (z. B. Eltern). In diesem Modell ist nicht eindeutig benannt, wer konkret für die Entwicklung von Innovationen zuständig ist. Hier sind sowohl die Vorgänge im Rahmen der „top-down"-Strategie sowie der „symbiotischen Strategie" (siehe Gräsel & Parchmann, 2004) denkbar.

In Fragen von Implementation ist es notwendig, dass „Betroffene zu Beteiligten werden" (Reinmann-Rothmeier & Mandl, 1998, S. 308). Zudem sollten im Zuge von Implementation alle Akteure zu Lernenden in einem permanenten Evaluationsprozess werden, was bedeutet, dass bereits zu Beginn eines Vorhabens Unstimmigkeiten erkannt und korrigiert werden (Petermann, 2015b, S. 122). Die Dialoge und die Kooperation aller Beteiligten führen letztlich erst dazu, dass es zu den geplanten Zielen kommen kann. „Solange Schulen, Hochschulen und

die berufliche Weiterbildung nicht selbst [...] ‚mitlernen' und auf diese Weise die Rahmenbedingungen für neue Entwicklungen [...] schaffen, werden Energien und Ideenreichtum motivierte Lehrer, Dozenten und Trainer im wahrsten Sinne des Wortes verschlissen" (Reinmann-Rothmeier & Mandl, 1998, S. 310).

4.4.2 Bedeutsamkeit und Desiderata

Unbestritten ist, dass „die Relevanz für die Praxis ein notwendiges Qualitätskriterium für unterrichtswissenschaftliche Forschung darstellt" (Gräsel & Parchmann, 2004, S. 197). Die Erhebung und Formulierung von Implementationsqualität sollte ein zentraler Bestandteil von Forschungsprojekten sein, um zu prüfen, inwiefern sich Konzepte in die Praxis umsetzen lassen und wo evtl. Fallstricke liegen (Gräsel & Parchmann, 2004). „Erst durch das Vorhandensein von Informationen über das Ausmaß der Implementation wird es möglich, Aussagen über die Wirksamkeit und den Erfolg einer Maßnahme zu treffen" (Petermann, 2015b, S. 122). Dass die Relevanz von Hinweisen zur Implementation auch auf bildungspolitischer Ebene verstanden wird, zeigt das geplante Vorgehen des Landespräventionsrates Niedersachsen, die Programme der „Grünen Liste Prävention" um Hinweise zur Implementation zu ergänzen (Groeger-Roth, 2014).

Die Implementationsforschung erhält derzeit eine größere Aufmerksamkeit. Dennoch ist die Frage der Implementation ein Aspekt, der in den meisten (Interventions-)Studien vernachlässigt wird (Durlak, 1997a; Oswald et al., 2013). „[Es] besteht nach wie vor ein Mangel an Angaben zur Implementation von Programmen und an Implementationsforschung" (Oswald et al., 2013, S. 256). Wenn Hinweise auf eine unzureichende oder ungenaue Implementation vorliegen, lassen Autoren dieses Ergebnis in Publikationen oftmals sogar unerwähnt (Križan & Vossen, 2016, S. 83; Mihalic, 2004, S. 83). Mittels Metaanalysen wurde bereits festgestellt, wie eng umrissen die Angaben zur Implementationsqualität in publizierten Studien sind. Domitrovich und Greenberg (2000) untersuchten 34 Studien zu etablierten Präventionsprogrammen psychischer Störungen im Schulalter. Nur sieben der Studien gingen über den Aspekt der Implementationstreue hinaus, indem zum Beispiel auch die Verantwortlichkeit der Durchführenden oder der Grad der Anpassung erhoben wird. Insgesamt wird nur in 1/3 der Studien die Wirksamkeit der Präventionsmaßnahme im Kontext der Implementationsqualität diskutiert. Dass Implementationsprobleme (zum Beispiel in der Mitarbeit der Adressaten, dem Engagement der Programm-Administratoren oder die Rahmenbedingungen der Programm-Durchführung) durchaus zu einer geringeren Programmwirksamkeit führen können, zeigen Durlak und Kollegen (2011) in ihrer Metaanalyse (s. auch Lambert et al., 2014).

4.4.3 Kriterien von Implementation

Der am häufigsten genannte Bereich der Implementationsqualität ist die Wiedergabetreue, die „fidelity" (Lambert et al., 2014; Mihalic, 2004). Sie wird übersetzt als „the degree of fit between the developer-defined elements of a prevention program, and its actual implementation in a given organization or community setting" (Mihalic, 2004, S. 83). Doch zeigt die aktuelle Fachliteratur zum Thema, dass Implementationsqualität auf mehreren Säulen beruht, die weitere

Aspekte umfassen. Oswald und Kollegen (2013, S. 257 f.) nennen drei zentrale Bereiche, die zu einer möglichst optimalen Implementation führen:

1. Die Nähe der Interventionsinhalte zur untersuchten Praxis,
2. der Einbezug von Lehrerkollegien zur gegenseitigen Unterstützung und
3. die langfristige Unterstützung der Lehrperson.

Eine weitere umfassendere und detailliertere Aufschlüsselung von Implementationsvariablen stammt von Petermann (2015b, S. 123 ff.).

TABELLE 26 IMPLEMENTATIONSVARIABLEN (PETERMANN, 2015B, S. 123 FF.)

Implementationsvariable	Beschreibung
1. Akzeptanz (Acceptability)	Die Auffassung der an der Implementation Beteiligten, dass oder in welchem Maße eine bestimmte Intervention oder Behandlung zufriedenstellend ist
2. Übernahme (Adoption)	Die Absicht oder Entscheidung, eine Intervention in die Praxis umzusetzen
3. Angemessenheit (Appropriateness)	Die wahrgenommene Passung, Aktualität und Kompatibilität der Intervention für ein bestimmtes Setting, einen Anbieter oder einen Nutzer
4. Machbarkeit (Feasibility)	Das Ausmaß, in dem ein Angebot erfolgreich in einem definierten Setting eingesetzt werden kann
5. Wiedergabetreue (Fidelity)	Das Maß, in dem eine Intervention so umgesetzt wird, wie es im Original vorgesehen ist
6. Implementationskosten (Cost)	Die Kosten einer Implementierung
7. Durchdringung (Penetration)	Die Integration einer Maßnahme in die betreffende Institution
8. Nachhaltigkeit (Sustainability)	Das Ausmaß, in dem ein Angebot erhalten oder institutionalisiert wird

Die acht genannten Implementationsvariablen können unterschiedlich erhoben werden. Grundsätzlich bieten sich Umfragen/Fragebögen, Interviews, Beobachtungen oder auch die Erhebung administrativer Daten an. Je nach fokussierter Dimension (Implementationsqualität vs. -quantität) eignen sich die methodischen Zugänge besser oder schlechter. Petermann (2015b) hat die acht Parameter nicht explizit den Dimensionen Implementationsqualität bzw. -quantität zugeordnet, doch werden die Aspekte Akzeptanz, Übernahme, Angemessenheit, Machbarkeit und Wiedergabetreue eher zum *qualitativen* Paradigma gezählt, Implementationskosten, Durchdringung und Nachhaltigkeit hingegen eher zum *quantitativen* Paradigma. Nicht explizit erwähnt, jedoch an verschiedenen Stellen berücksichtigt, ist der Implementationsaspekt „Spaß" bzw. „Lernfreude" der Schülerinnen und Schüler mit bzw. an einem stimmten Programm, der schon verschiedentlich erhoben und diskutiert ist (siehe z. B. für Peer-Tutoring Calhoon & Fuchs, 2003; für das Programm Calcularis: Käser u. a., 2013). Dieser Aspekt könnte im weiteren Sinne unter die Kategorie „Akzeptanz" (siehe Tabelle 26) fallen.

Durlak (1997a, S. 194) empfiehlt die Berücksichtigung folgender fünf Aspekte, um eine gute Implementation zu erreichen:

1. „Insuring that program practices and goals are acceptable to the host community.

2. Carefully training those who will deliver the program so that they are able to conduct the program as planned.
3. ‚Manualizing' the program so that its key elements can be replicated by other adopting agencies.
4. Having mechanisms in place for ongoing monitoring of program implementation.
5. Remaining available for continual consultation and problem solving."

Ein limitierender Faktor bei allen Implementationsbemühungen ist dabei die subjektive Theorie einer Lehrkraft, die bei der Übernahme einer unterrichtlichen Neuerung eine zentral Rolle spielt (Gräsel & Parchmann, 2004). Die Lehrkraft übernimmt langfristig Veränderungsvorschläge, sobald sie diese als realisierbar, nützlich und sinnvoll empfindet. Auch die Selbstwirksamkeitserwartung spielt eine bedeutsame Rolle, da sie positiv mit dem Grad an pädagogischem Engagement und negativ mit dem Belastungsempfinden korreliert (Schmitz & Schwarzer, 2002).

4.4.4 Zusammenfassung

Aktuell besteht eine eindeutige Forderung danach, die Implementation von Innovationen in die pädagogische Praxis in Forschungsprozessen zu berücksichtigen. Dies soll vermeiden, dass die offensichtliche Relevanz unterrichtswissenschaftlicher Forschung für die Praxis nicht in leeren Worthülsen mündet (siehe Kapitel 4.4.2). Dabei geht es zum einen um die Implementations*qualität* zum Beispiel im Sinne der Wiedergabetreue oder der Machbarkeit von Konzepten oder Maßnahmen sowie um die sich anschließende Erhebung der Implementations*quantität*, sprich die Umsetzungshäufigkeit von innovativen Konzepten oder Maßnahmen im weiten pädagogischen Feld.

Das Konzept der Implementation basiert dabei auf verschiedenen Säulen, die Petermann (2015b) in acht Implementationsvariablen aufgeschlüsselt hat: Akzeptanz, Übernahme, Angemessenheit, Machbarkeit, Wiedergabetreue, Implementationskosten, Durchdringung und Nachhaltigkeit (siehe Kapitel 4.4.3). Uneinigkeit besteht derzeit darüber, zu welchem Zeitpunkt die Erhebung der Implementationsqualität angezeigt ist. Einige Autoren verorten diese erst zum Zeitpunkt der Verankerung einer Maßnahme in die pädagogische Praxis, unterstützt durch Verlage oder Fortbildungsinstitute. Hier wird jedoch das Plädoyer eröffnet, bereits früher, also in Pilotphasen oder ersten Evaluationsversuchen im Feld, die Implementationsqualität zu berücksichtigen. Denn Forschungsergebnisse zeigen, dass letztlich die Wirksamkeit einer Maßnahme stark mit der Implementationsqualität (Durlak et al., 2011; Lambert et al., 2014) und der Überzeugung der beteiligten Akteure (Oswald et al., 2013) zusammenhängt. Zusammenhänge zwischen dem Ausmaß der Wirksamkeit und Implementationsquantität konnten bisher nicht festgestellt werden (Petermann, 2015b). Sobald Hinweise zur Implementationsqualität vorliegen, kann in diesem Kontext die Wirksamkeit einer Innovation beurteilt werden.

5 Zusammenfassung der theoretischen Grundlagen

Die PISA-Ergebnisse aus dem Jahr 2015 zeigen, dass fast ein Fünftel der Schülerinnen und Schüler in der Sekundarstufe I massive Schwierigkeiten in ihrer mathematischen Kompetenzentwicklung zeigen. Ein vertiefter Blick in mathematische Entwicklung zeigt dabei, dass diese bereits weit vor Schuleintritt beginnt. Geary (2000) spricht dabei von der Entwicklung der Vorläuferfertigkeiten als „biologically primary quantitative abilities". Das ZGV-Modell nach Krajewski (2013) und das Entwicklungsmodell nach Fritz und Ricken (2008, 2009) bilden erste mathematische Entwicklungsschritte empirisch abgesichert ab (siehe Kapitel 2.2). Doch auch wenn der Einfluss früher mathematischer Kompetenzen auf die weitere Entwicklung eine große Rolle spielt, ist der Mathematikunterricht der zentrale Lernkontext, in dem Schülerinnen und Schüler die Welt der Mathematik begreifen und verstehen (Fritz & Ricken, 2008, S. 78; Schneider et al., 2013, S. 209). Mathematische Lernprozesse werden im schulischen Kontext durch Curricula und Bildungsstandards systematisiert und gesteuert, sodass Geary (2000) diese dann als „biologically secondary abilities" bezeichnet. Die Inhalte der Grundschulmathematik stellen dabei die wichtige Grundlage für den weiteren schulischen Kompetenzerwerb dar (siehe Kapitel 2.3.3).

Die pädagogisch-psychologische Perspektive versucht durch Kompetenzstufenmodelle und Bildungsstandards den mathematischen Kompetenzerwerb in der Schule zu konkretisieren, um vor allem Lernrückständen identifizieren zu können und eine entsprechende Unterrichtsgestaltung bzw. Förderung zu ermöglichen. Zentraler Inhalt in der Grundschule ist dabei die „Arithmetik". Insbesondere die Grundrechenarten sollen in den ersten vier Jahrgangsstufen verstanden und angewendet werden können. Ein gefestigtes Verständnis der Grundschulmathematik ist unumgänglich für den weiteren mathematischen Kompetenzerwerb. Das Verständnis des systematischen Rechnens und das Regelwissen in der Grundschulzeit bilden die Grundlage für komplexe Rechenoperationen im weiteren Mathematikunterricht (Gaidoschik, 2008; Grube, 2005; Moser Opitz, 2007).

Als besonders herausfordernd stellt sich im Anschluss an die Grundschulzeit eine gelingende Transition *in* die Sekundarstufe I dar, die in Deutschland im Regelfall nach der vierten Jahrgangsstufe erfolgt (siehe Kapitel 2.4.1). Weißbach schreibt schon 1986 vom sog. „Sekundarstufenschock". Hier zeigen anspruchsvolle Längsschnittstudie wie u. a. die PALMA-Untersuchung (Vom Hofe et al., 2009; Pekrun et al., 2006) oder die KESS-, LAU- oder LOGIK-Studie (KESS: Bos & Gröhlich, 2010; LAU: Lehmann, Husfeldt, et al., 2001; LOGIK: Stern, 2008), dass die Entwicklung grundlegender Basiskompetenzen, die schwerpunktmäßig in der Grundschule erworben werden, nach der Transition weitergeht und keinesfalls abgeschlossen ist (siehe Kapitel 2.4.2 & 2.4.3). Für die auf die Grundschule folgende Sekundarstufe I gibt es derzeit kein empirisch abgesichertes Erwerbs- oder Kompetenzstufenmodell. Einen Überblick über die zu erwerbenden Kompetenzen *am Ende* der Sekundarstufe I ermöglichen zwar die Bildungsstandards für den Mittleren Schulabschluss (Kultusministerkonferenz, 2003), jedoch sind bisher kein jahrgangsübergreifendes bzw. an die Primarstufe anknüpfendes Kompetenzstufenmodell entwickelt.

Zahlreiche Befunde gibt es dahingehend, dass ein nicht zu missachtender Teil der Schülerinnen und Schüler in der Sekundarstufe I Defizite in den mathematischen Basiskompetenzen aufweisen (siehe Kapitel 2.4.2). Diese manifestieren sich im Laufe der Entwicklung und es kann zu massiven Schwierigkeiten im weiteren Erwerb mathematischer Kompetenzen in der Sekundarstufe I kommen. Befunde der Transitionsforschung zeigen sogar, dass Schülerinnen und Schüler einen Leistungsrückgang nach dem Übergang in die Sekundarstufe I verzeichnen können (siehe Kapitel 2.4.1).

Die Auseinandersetzung mit Schwierigkeiten in der Entwicklung mathematischer Kompetenzen zeigt, dass – neben den Schülerinnen und Schülern mit massiven Rechenstörungen – weitere 15–20 % der Schülerinnen und Schüler allgemeine Rechenschwierigkeiten zeigen (Fischbach et al., 2013; Lorenz, 1993; Mazzocco & Myers, 2003) (siehe Kapitel 3.2). Die Ursachen für allgemeine Rechenschwierigkeiten können im biologischen, familiären oder schulischen Kontext oder auch in Bedingungen des Umfelds liegen. In Fragen der Förderung gerade im inklusiven Setting ist zu bedenken, dass sich die Spannweite an mathematischen Kompetenzen weiter vergrößern wird (Schels et al., 2015). Bezüglich der pädagogischen Folgen besteht Konsens darüber, dass *alle* Schülerinnen und Schüler mit Schwierigkeiten im Rechnen einer Förderung bedürfen und sich diese nicht auf diejenigen mit einer Rechenstörung begrenzen sollte (Gold, 2015). Standardisierte Rechentests und informelle Verfahren ermöglichen eine differenzierte Beschreibung und Identifizierung der Problemstellen beim Kompetenzerwerb. Diese mehrschrittige Diagnostik ermöglicht die Entwicklung einer Förderung, die am individuellen Kompetenzstand der rechenschwachen Schülerin oder des rechenschwachen Schülers anknüpft (siehe Kapitel 3.4). Dabei sollte die Unterstützung grundsätzlich curriculare Inhalte fördern sowie die Entwicklung kognitiver und emotional-sozialer Kompetenzen berücksichtigen (Lorenz, 2014) (siehe Kapitel 3.5).

In Fragen der Förderung mathematischer Basiskompetenzen gerade in inklusiven Lerngruppen wird derzeit die Forderung nach Evidenzbasierung lauter. *Evidenzbasiert* bedeutet hier, dass in Unterricht- und Fördersettings Maßnahmen und Methoden zum Einsatz kommen, deren Wirksamkeit durch wissenschaftliche Forschung bestätigt werden und deren Einsatz sich im jeweiligen Einzelfall eignet (Hillenbrand, 2015, S. 314). Das Konzept der Evidenzbasierung stammt ursprünglich aus den Vereinigten Staaten und genießt in Deutschland aktuell große Aufmerksamkeit. Die Aufarbeitung des Forschungsstandes bzgl. Möglichkeiten der Förderung mathematischer Kompetenzen – auf Grundlage des Handlungsleitfadens zur Identifikation und Analyse evidenzbasierter Verfahren und Methoden nach Freeman und Sugai (2013) – ergibt, dass es zahlreiche evaluierte mathematische Förderprogramme für den *Elementar-* und *Primarstufenbereich* gibt. Zu nennen wäre zum Beispiel „Mengen, zählen, Zahlen“ (Krajewski et al., 2013).

Im Bereich der Sekundarstufe I stellt sich die Befundlage als sehr defizitär dar. Die Förderung des Kompetenzerwerbs in der Sekundarstufe I ist bisher sehr vernachlässigt worden (Ehlert et al., 2013; Krajewski & Ennemoser, 2010; Mastropieri et al., 2008; Ufer, 2009). Als potenziell evidenzbasiert gelten hier die „Rechenspiele mit Elfe und Mathis II“ (Lenhard et al., 2010) und als eindeutig evidenzbasiert das Programm „Calcularis“ (2015). Doch sind beide Verfahren nur im Einzelsetting oder in Kleingruppen einsetzbar. Ein Förderprogramm, das im inklusiven Setting mit allen Schülerinnen und Schülern einer Klasse durchgeführt werden kann, liegt bisher nicht vor. Der Handlungsbedarf diesbezüglich ist immens. Identifiziert werden können für den

Rahmen einer Förderung jedoch geeignete Förder*methoden* (siehe Kapitel 4.3.3). Diese „[sind] nicht *explizit* in Form von geschlossenen Konzeptionen, als Förderprogramme für den mathematischen Bereich ausgewiesen [...], sondern [werden] zur Leistungsförderung *auch* im mathematischen Bereich empfohlen“ (Koch, 2008, S. 100, Herv. i. Orig.). Eine umfassende Analyse des Forschungsstandes ergibt die Direkte Instruktion, peergestütztes Lernen, kooperatives Lernen und Feedback als wissenschaftlich fundierte Methoden, die sich zur Förderung mathematischer Kompetenzen in der Sekundarstufe I anbieten.

Neben der Frage der Wirksamkeit spielt in der Evaluation von Förderkonzepten oder -maßnahmen verstärkt auch die Implementation eine Rolle (siehe Kapitel 4.4). Petermann (2015b, S. 123) schlägt acht Implementationsvariablen vor, die zu berücksichtigen sind: Akzeptanz, Übernahme, Angemessenheit, Machbarkeit, Wiedergabetreue, Implementationskosten, Durchdringung und Nachhaltigkeit. Es gilt folglich, die Implementationsqualität einer Fördermaßnahme zu erheben, um die Wirksamkeit im gemeinsamen Zusammenspiel verstehen und interpretieren zu können.

Im folgenden Kapitel wird eine Maßnahme zur Förderung mathematischer Basiskompetenzen für inklusive Settings der Sekundarstufe I vorgestellt, die durch die Implementation der evidenzbasierten Fördermethoden der Direkten Instruktion, des peergestützten und kooperatives Lernens sowie des Feedbacks gekennzeichnet ist. Mit dieser Fördermaßnahme wird dem großen Bedarf an evidenzbasierter Förderung mathematischer Basiskompetenzen in der Sekundarstufe I begegnet.

6 Grundlagen der Maßnahme zur Förderung mathematischer Basiskompetenzen zu Beginn der Sekundarstufe I

Das Kapitel 4.3 zum Thema der evidenzbasierten Förderung mathematischer Basiskompetenzen in der Sekundarstufe I gibt einen Überblick über den aktuellen Forschungsstand und ermöglicht zentrale Erkenntnisse darüber, durch welche Methoden mathematische Basiskompetenzen nach dem Übergang in die Sekundarstufe I gefördert werden können. Im Rahmen eines universitären Forschungsprojekts werden diese Erkenntnisse gebündelt und in den Rahmen einer Fördermaßnahme implementiert. Das Konzept dieser Maßnahme wird im Folgenden vorgestellt.

Zu Beginn werden die grundlegenden Ziele der Maßnahme formuliert. Im Anschluss werden die Entwicklungsschritte der Maßnahme beschrieben, beginnend mit der Erarbeitung des Konzepts in einer Projektgruppe über die Überarbeitung der Maßnahme im Anschluss an eine Pilotstudie bis hin zur Durchführung der Hauptstudie. Nachdem der Entwicklungsprozess dargestellt ist, wird das inhaltliche Konzept vorgestellt. Es werden neben den Inhalten und der Struktur die zentralen didaktischen und methodischen Prinzipien vorgestellt.

6.1 Ziele der Förderung

Grundlegend für eine Fördermaßnahme sind, noch vor Ausarbeitung jeglicher methodischer und didaktischer Inhalte, die Ziele, die mit der Maßnahme verbunden sind und durch sie erreicht werden sollen (Gugel, 2006, S. 13). Die hier entwickelte Maßnahme verfolgt Ziele auf drei verschiedenen Ebenen: der Mikro-, Meso- und Makroebene (Hascher & Schmitz, 2010). Dabei können die im Folgenden genannten Ziele nicht zwingend *einer* einzigen Ebene zugeordnet werden.

Die *Mikroebene* bezieht sich auf die Ziele für die Schülerinnen und Schüler selbst. Ziel der Maßnahme ist es, bei ihnen die mathematischen Basiskompetenzen zum Eintritt in die Sekundarstufe I zu fördern und zu festigen, damit sie mit einer gesicherten Ausgangsbasis Anschluss an den Mathematikunterricht der Sekundarstufe I finden. Dabei muss die Leistungsheterogenität der Schülerinnen und Schüler berücksichtigt werden, sodass rechenstarke Schülerinnen und Schüler ge*fordert* und rechenschwache Schülerinnen und Schüler ge*fördert* werden. Aber auch den durchschnittlich rechenbegabten Schülerinnen und Schülern muss die Konzeption gerecht werden. Ziel ist es, mit den Schülerinnen und Schülern geeignete Rechenstrategien zu wiederholen und ggf. auch zu vermitteln, die sie dann auf einem eigenen Niveau verfestigen und anwenden können.

Die *Mesoebene* betrifft zum einen die Klasse als Ganzes und zum anderen den Lernort Schule. Die Maßnahme ermöglicht es herauszufinden, welche Rechenstrategien bisher in der Grundschule erworben wurden und welche Kompetenzstände die Schülerinnen und Schüler mit in die Sekundarstufe I bringen. Das gemeinsame Wiederholen und Abgleichen zentraler Rechenstrategien befördert eine solide Ausgangsbasis bzgl. des mathematischen Strategiewissens für die Sekundarstufe I. Der Einsatz peer-mediierter Methoden ermöglicht die Schaffung eines positiven Lernklimas innerhalb der Klasse.

Letztlich verfolgt die Maßnahme auch Ziele auf der *Makroebene* – bezogen auf die Wissenschaft und die schulische Praxis. Ziel ist es, auf dieser Ebene theoretisch evidenzbasierte Methoden umzusetzen und die Praxistauglichkeit zu evaluieren. So kann eine Verzahnung von Wissenschaft und Praxis gelingen, in der das Methodenrepertoire der Lehrkräfte vergrößert werden kann. Zudem wird mit der Entwicklung einer univserellen Fördermaßnahme im Bereich der mathemaischen Basiskompetenzen zu Beginn der Sekundarstufe I eine Forschungslücke geschlossen (vgl. Tabelle 20).

TABELLE 27 ZIELE DER MAßNAHME

Ebene	Ziele
Mikroebene (Schülerinnen und Schüler)	• Festigung und Sicherung des Stellenwertsystems und der grundlegenden Rechenstrategien (halbschriftlich, schriftlich, Kopfrechnen) in den Grundrechenarten • Bewertung und zielführender Einsatz von Rechenstrategien • Förderung rechenschwacher Schüler/Schülerinnen • Forderung rechenstarker Schüler/Schülerinnen • Erleichterung der Transition • Ausbau sozialer Kompetenzen
Mesoebene (Klasse & Schule)	• Vergleich und Abgleich bisher gelernter Rechenstrategien der Schülerinnen und Schüler • Identifikation der individuellen Kompetenzstände nach Transition • Ermöglichung einer gemeinsamen Basis für den Mathematik-unterricht in der Sekundarstufe I • Schaffung eines positiven Lernklimas in der Klasse • Aufzeigen einer lernförderlichen Unterrichtseinheit zum Einstieg in den Mathematikunterricht nach der Transition
Makroebene (Wissenschaft & schulische Praxis)	• Erarbeitung und Evaluierung mathematischer Förderung in inklusiven Lernsettings • Überprüfung der Einsatztauglichkeit evidenzbasierter Methoden • Aufzeigen von Chancen und Fallstricken der Methoden • Erweiterung des Methodenrepertoires der Lehrkräfte • Verzahnung von Wissenschaft und Praxis • Schließen einer Forschungslücke

Damit die Maßnahme langfristig im schulischen Kontext Anwendung findet *(Mesobene: Aufzeigen einer lernförderlichen Unterrichtseinheit zum Einstieg in den Mathematikunterricht nach der Transition)*, sollte im Anschluss an die Evaluation der Maßnahme eine strukturierte Implementation in die Schullandschaft erwogen werden und erfolgen.

6.2 Entwicklungsprozess

Im Folgenden werden der zeitliche und organisatorische Entwicklungsprozess der Maßnahme vorgestellt. Zur detaillierten methodisch-didaktischen Ausgestaltung der Maßnahme wird auf das darauffolgende Kapitel 6.3 verwiesen.

Die Maßnahme wird nach der Idee und unter der Leitung von Prof. Dr. Clemens Hillenbrand, Lehrstuhlinhaber der Pädagogik und Didaktik bei Beeinträchtigungen des Lernens, im Jahr 2014 an der Carl von Ossietzky Universität Oldenburg entwickelt. Am Ende des Jahres 2013

bildet sich nach Vorstellung des Projekts innerhalb des Instituts für Sonder- und Rehabilitationspädagogik eine Projektgruppe von neun Studierenden des Master of Education und einer wissenschaftlichen Mitarbeiterin des Lehrstuhls der Pädagogik und Didaktik bei Beeinträchtigungen des Lernens, um auf Grundlage wissenschaftlicher Forschungsbefunde (siehe Kapitel 4.3) an der grundlegenden Konzeption der Maßnahme zu arbeiten. Es folgen zahlreiche Projekttreffen mit der Gesamtgruppe (siehe Tabelle 28). Nachdem im Januar 2014 die grundsätzliche Struktur der vierwöchigen Maßnahme (16 Stunden à 45 Minuten) erarbeitet ist, werden kleinere Arbeitsgruppen gebildet, die sich entweder mit spezifischen Themeneinheiten innerhalb der Maßnahme oder der Umsetzung der Methoden beschäftigen. Die Konzeption und die Entwicklung der Maßnahme belaufen sich über ca. sechs Monate im Zeitraum von Januar bis Juni 2014.

TABELLE 28 ZEITLICHER ABLAUF DER KONZEPTION UND ENTWICKLUNG DER MAßNAHME

Datum/Zeitraum	Schritte im Projekt
Dezember 2013	Bildung der Projektgruppe
15.01.2014	Projektrahmen & Zeitplanung
30.01.2014	Konzeption der Maßnahme & Erstellen von Arbeitsgruppen
bis 10.02.2014	Weiterarbeit in Arbeitsgruppen
10.02.2014	Unterrichtsplanung I
bis 24.02.2014	Weiterarbeit in Arbeitsgruppen
24.02.2014	Unterrichtsplanung II
bis 04.03.2014	Weiterarbeit in Arbeitsgruppen
04.03.2014	Fertigstellung der Unterrichtsplanung
11.04.2014	Abschließende Feinabstimmung
06.06.2014	Formatierung & Erstellung des Manuals
Juni 2014	Pilotstudie
04.07.2014	Vorstellung des Konzepts auf einer deutschen Tagung
September/Oktober 2014	Hauptstudie

Im Juni 2014 liegt die Maßnahme in manualisierter Form vor, sodass sie im Anschluss im Rahmen einer Pilotstudie umgesetzt werden kann (für Details siehe Kapitel 8.4). Im Juli 2014 wird das Konzept der Maßnahme auf einer deutschen Tagung vor wissenschaftlichem Publikum vorgestellt, um Feedback zur Konzeption zu erhalten. Im September und Oktober findet dann die Hauptstudie statt (für Details siehe Kapitel 8.4 Untersuchungsdurchführung).

6.3 Darstellung der Fördermaßnahme

Da die Intervention eine möglichst andauernde Wirksamkeit erzielen soll, handelt es sich um eine sog. „Langzeitintervention“ (Hagenauer, 2010, S. 244). Im Folgenden werden die didaktischen Kernfragen erläutert (Kapitel 6.3.1), denn „im Unterricht stehen natürlich nicht die Methoden, sondern die Inhalte im Vordergrund“ (Gugel, 2006, S. 13). Die Frage nach den eingesetzten Methoden in der entwickelten Maßnahme wird aufgrund der inhaltlichen Gewichtung

und der umfassenden theoretischen Herleitung getrennt in einem zweiten Unterkapitel beantwortet (Kapitel 6.3.2).

6.3.1 Didaktische Prinzipien

Allgemein bezeichnet „Didaktik“ die Theorie der Praxis des Lehrens und Lernens. Jank und Meyer (1994, S. 16) schreiben konkret: „Die Didaktik kümmert sich um die Frage, wer was, wann, mit wem, wo, wie, womit, warum und wozu lernen soll“. Diese zentralen didaktischen Fragen werden im Folgenden beantwortet.

WER?

Die Maßnahme zur Förderung mathematischer Basiskompetenzen richtet sich an Schülerinnen und Schüler in der fünften Klasse, die den Prozess der Transition in die Sekundarstufe I vollzogen haben. Sie wird im Klassenkontext mit allen Schülerinnen und Schülern durchgeführt (siehe Kapitel 4.2). Der Einsatz evidenzbasierter Methoden und kooperativer und peergestützter Lernformen soll dabei gerade Schülerinnen und Schüler unter ungünstigen Lernausgangslagen für mathematischen Kompetenzerwerb unterstützen. In Deutschland liegen mit den „Rechenspielen mit Elfe und Mathis II“, dem Programm „Mathe sicher können Band 1 und 2“, dem Computerprogramm „Calcularis“ und „Kalkulie I, II, III“Förderprogramme für die Sekundarstufe I für Schülerinnen und Schüler mit Rechenschwierigkeiten auf *selektiver und indizierter Ebene* vor. Somit füllt die Fördermaßnahme die Lücke für ein Konzept, von dem die *gesamte Klasse auf universeller Ebene* im Anschluss an die Transition in die Sekundarstufe I profitieren kann und die auf wissenschaftlichen Erkenntnissen zur wirksamen Förderung mathematischer Basiskompetenzen basiert.

Durlak (1997a) ermöglicht mit seinem übersichtlichen Klassifikationsschema von Interventions- und Präventionsmaßnahmen die Verortung der Zielgruppe der Fördermaßnahme. Die Abbildung 38 platziert die hier geplante Fördermaßnahme innerhalb des Klassifikationsschemas (grau schraffierte Felder).

			Selection of Target Groups		
			Universal	High-Risk Groups	Those undergoing Transitions
Level of Intervention	Environmental	Community			
		School			
		Peer			
		Family			
		Individual			

ABBILDUNG 38 EINORDNUNG DER MAßNAHME IN DIE KLASSIFIKATION VON PRÄVENTIONS- UND INTERVENTIONSMAßNAHMEN (DURLAK, 1997A, S. 3)

Nach Durlak (1997b, S. 287) kann eine Maßnahme oftmals nicht immer explizit nur einer Kategorie zugeordnet werden. Denkbar sind zum Beispiel Maßnahmen, die Förderaspekte sowohl in den schulischen als auch in den familiären Kontext (z. B. über Elterntrainings) einflechten. Für die hier vorzustellende Maßnahme treffen die Kategorien „School“ und „Those undergoing

Transitions“ zu, da die Durchführung ausschließlich im schulischen Setting erfolgt und speziell im Anschluss an die Transition in die Sekundarstufe I durchgeführt wird, um einen guten Übergang zu ermöglichen. Die Kategorie „Individual“ trifft in diesem Kontext nicht zu, da es sich um eine universelle Maßnahme für die gesamt Klasse handelt.

WAS?

Bezüglich der Förderinhalte der Maßnahme zeigt die Aufarbeitung der mathematischen Kompetenzen von Schülerinnen und Schülern in der Sekundarstufe I durch nationale und internationale Studien (siehe 2.4 Erwerb mathematischer Kompetenzen in der Sekundarstufe I), dass einige Schülerinnen und Schüler in der Sekundarstufe I massive Schwierigkeiten in der Beherrschung der mathematischen Basiskompetenzen zeigen. Darunter wird das Verständnis (Konventions- und Regelwissen) und die Anwendung (Bearbeiten von Sachaufgaben) arithmetischer Operationen und des Stellenwertsystems verstanden (Ehlert et al., 2013; Ennemoser & Krajewski, 2013; Freesemann, 2014; Leiß & Blum, 2006; Reiss & Winkelmann, 2008).
Innerhalb der Maßnahme werden vor allem die Bereiche der technischen Grundlagen in den Grundrechenarten und des Stellenwertsystems fokussiert. Diese werden auch als Rechenfertigkeiten bezeichnet (Krauthausen & Scherer, 2008, S. 304) und sind notwendige – wenn auch keine hinreichende – Voraussetzung für den weiteren mathematischen Kompetenzerwerb. Technische Grundlagen werden sowohl in den Bildungsstandards für die Primarstufe (Kultusministerkonferenz, 2004) sowie in dem Kompetenzstufenmodell nach Reiss und Winkelmann (2009) als Fundament bzw. eine Art Grundvoraussetzung für die anwendungsbezogenen mathematischen Kompetenzen gesehen, zu denen auch die Modellierung von Sachaufgaben zählt.
So fokussiert die Maßnahme auf eine intensive, kontextbezogene Förderung dieser Basiskompetenzen, an die sich der Aufbau von „innermathematischer“ Problemlösekompetenz und „außermathematischer“ Modellierungskompetenz auf gefestigter Wissensgrundlage anschließen soll. Als Rechenverfahren werden dabei das halbschriftliche und schriftliche Rechnen und auch das Kopfrechnen gefördert. Diese drei Verfahren zählen, neben dem Einsatz des Taschenrechners, zu den „grundsätzliche[n] Methoden zur Bewältigung von Rechenanforderungen“ (Krauthausen & Scherer, 2008, S. 43).
Die Bedeutsamkeit des Kopfrechnens als „Grundbaustein des Rechnens“ (Krauthausen & Scherer, 2008, S. 51) ist in der Fachdidaktik und sonderpädagogischen Praxis unbestritten. Die Relevanz dieses Verfahrens lässt sich u. a. durch die Cognitive-Load-Theorie begründen, die zeigt, dass Arbeitsgedächtnisressourcen bei anspruchsvollen Aufgaben wie Problemlöseaufgaben unnötig beansprucht werden, wenn kein schneller Abruf von Rechenstrategien und -fakten erfolgt (Ayres & Paas, 2012; Fritz & Ricken, 2008). Im Rahmen der Maßnahme wird das Kopfrechnen zum Beispiel mittels Rechenmauern oder sog. „schönen Päckchen“ (Gaidoschik, 2009, S. 87 & S. 141) via Aufgabenreihen gefördert. „Schön sind Aufgabenreihen, an denen es etwas zu entdecken und begreifen gibt“ (Gaidoschik, 2009, S. 87).
Halbschriftliches Rechnen wird oftmals auch als gestütztes Kopfrechnen bezeichnet. Es zeichnet sich vor allem durch seine Flexibilität aus. Beim halbschriftlichen Lösen einer Rechenaufgabe gibt es diverse Wege, Zwischenschritte und -ergebnisse weiter zu verarbeiten, um am Ende auf das richtige Ergebnis zu kommen.

> „Entscheidend ist dabei, dass die Art und Weise der *Notation nicht festgelegt* ist, auch wenn die in manchen Schulbüchern vorzufindende Praxis der ‚Musteraufgaben' das Gegenteil suggerieren kann. Beim halbschriftlichen Rechnen sind auch die Wege zur Lösung nicht vorgeschrieben, was dem Aufgabenlöser größere Freiräume beim Verfolgen eigener Wege erlaubt" (Krauthausen & Scherer, 2008, S. 46; Herv. i. Orig.).

Halbschriftliche Rechenverfahren ermöglichen es Schülerinnen und Schülern, eigene Rechenwege zu verfolgen, anstatt vorgegebene, die eventuell nicht verstanden sind. Durch geschicktes Rechnen werden dabei Zahlvorstellungen, Zahlbeziehungen und Rechengesetze miteinander verknüpft.

Beim schriftlichen Rechnen hingegen handelt es sich um Algorithmen, die in ihrer Abfolge festgelegt sind und zu einer „eindeutig bestimmten Lösung" (Gerster, 2009, S. 269) führen. Moser Opitz (2007b, S. 269) empfiehlt, „halbschriftliche Strategien im Unterricht mit Schülerinnen und Schülern mit besonderen Förderbedarf besonders zu gewichten". Schriftliches Rechnen wird lange Zeit als „Krönung oder Höhepunkt des Rechenunterrichts" (Krauthausen & Scherer, 2008, S. 51) aufgefasst, was durch seine Effizienz begründet wird. Schriftliche Rechenverfahren sollten verständnisgestützt unter Verknüpfung mit Kopfrechenstrategien und halbschriftlichen Strategien vermittelt werden, um zu vermeiden, dass die Schülerinnen und Schüler routiniert und ohne mathematische Einsicht rechnen.

WARUM?

„Die zunehmende Beherrschung grundlegender Rechenfertigkeiten (erworben durch vielfältiges aktiv-entdeckendes Lernen und produktives Üben) soll zu *Routinen* führen, durch die Rechenanforderungen kognitiv handhabbar werden" (Krauthausen & Scherer, 2008, S. 44; Herv. i. Orig.). Schriftliches Rechnen hat eine „exemplarische [...] Bedeutung für den gesamten algorithmisch-kalkülhaften Teil der Schulmathematik" (Gerster, 2009, S. 269). In der fachdidaktischen Diskussion werden dem Kopfrechnen und dem halbschriftlichen Rechnen jedoch eine größere Relevanz zugeschrieben, da schriftliches Rechnen im Alltag durch elektronisches Rechnen abgelöst wird (Krauthausen, 2009, S. 101; Krauthausen & Scherer, 2008, S. 52; Schipper, 2009). Schipper (2009, S. 133) verweist darauf, dass schriftliche Rechenverfahren für lern- und/oder rechenschwache Schülerinnen und Schüler, die massive Probleme damit haben, Aufgaben im Kopf zu rechnen, ein „vorletzter Rettungsanker vor der Bewältigung [...] mithilfe eines Taschenrechners sein [kann]".

Auch wissenschaftliche Befunde sprechen für die Fokussierung der Förderung auf technische Fertigkeiten und Strategien. Vom Hofe und Kollegen (2009) analysieren die Modellierungskompetenz von Sekundarstufenschülerinnen und -schülern und stellen dabei fest, dass die meisten Defizite auf unvollständig ausgebildete oder gar fehlende Grundvorstellungen zurückzuführen sind. Bestätigt wird dieser Befund durch eine vergleichbare Untersuchung von Ehlert und Kollegen (2013). Dass zudem eine Förderung der grundlegenden Rechenoperationen bei Schülerinnen und Schülern einen sehr großen Lerneffekt zeigt, zeigt die Metaanalyse von Kroesbergen und Van Luit (2003). Die Autoren ermitteln für die Förderung der sog. „basic facts" eine Effektstärke von $d_{\text{gewichtet}} = 1.14$, für die Förderung von Problemlöseaufgaben hingegen eine Effektstärke von $d_{\text{gewichtet}} = .63$. Letztlich unterstützt auch der Modellierungskreislauf von Blum und Leiß (2005) die hier geplante Fokussierung auf Rechenoperationen. Der Kreislauf besteht

aus sieben Schritten: 1. Verstehen, 2. Vereinfachen/Strukturieren, 3. Mathematisieren, 4. Mathematisch arbeiten, 5. Interpretieren, 6. Validieren, 7. Darlegen/Erklären. Insbesondere Schritt 3 und Schritt 4 machen es dabei erforderlich, dass Schülerinnen und Schüler die zur Fragestellung passenden Rechenoperationen wählen und diese dann korrekt anwenden (Verständnis und Anwendung, siehe Kapitel 2.1). „Zur Bewältigung dieses Teilschrittes [Schritt 4; d. Verf.] ist hier insbesondere die Kompetenz *Mit symbolischen, formalen und technischen Elementen der Mathematik umgehen* erforderlich“ (Institut zur Qualitätsentwicklung im Bildungswesen, 2009, S. 80; Herv. i. Orig.). Ohne Rechenstrategien und Faktenwissen ist es letztlich nicht möglich, den Modellierungskreislauf erfolgreich zu durchlaufen (siehe auch Vollrath & Roth, 2012, S. 262).

Wann?

Der Zeitpunkt der Implementation in den Mathematikunterricht unmittelbar nach der Transition ermöglicht es, die Rechenstrategien, die die Schülerinnen und Schüler in der Grundschule gelernt haben, untereinander zu vergleichen und den zielgerichteten Einsatz durch die Schülerinnen und Schüler abwägen zu lassen. Gerade beim halbschriftlichen Rechnen ist entscheidend, dass die Flexibilität der Rechenwege berücksichtigt wird (Krauthausen & Scherer, 2008, S. 46). Der Einsatz der Maßnahme am Anfang der Sekundarstufe I ermöglicht, dass die Schülerinnen und Schüler mit einer gesicherten Ausgangsbasis in den Mathematikunterricht der Sekundarstufe I übergehen können.

Wie?

In der Entwicklungsphase der Maßnahme wird mit regulär durchschnittlich vier Stunden Mathematikunterricht pro Woche gerechnet (Niedersächsische Landesschulbehörde, 2016). Insgesamt werden 16 Förderstunden geplant und gestaltet. Der Förderzeitraum umfasst vier Wochen. Die Einheit zum Stellenwertsystem umfasst eine Förderstunde, die zu den Grundrechenarten jeweils drei Stunden. Die Fördermaßnahme ist gerahmt durch eine Einführungs- und eine Abschlussstunde. Die 15. Sitzung wird als didaktische Reserve eingebaut und dient der Vertiefung oder Wiederholung ausgewählter Inhalte, je nach Bedarf der Interventionsleiterinnen und -leiter[1] und der Schülerinnen und Schüler. Die Förderdauer beläuft sich mit geplanten 720 Minuten (16 x 45 Minuten) und damit auf mehr als 600 Minuten, was nach Ise und Kollegen (2012, S. 189) den größten Interventionseffekt erwarten lässt. Einen Überblick über die Verteilung der Themen auf die 16 Stunden bietet Tabelle 29.

[1]Die Studierenden, die die Fördermaßnahme in der Pilot- oder Hauptuntersuchung umsetzen, werden im Folgenden angelehnt an Ise und Kollegen (2012) als „Interventionsleiterinnen und -leiter“ bezeichnet.

TABELLE 29 ÜBERBLICK ÜBER DIE INHALTE DER FÖRDERSTUNDEN

Stunde	Inhalte
1	Einführungsstunde
2	Übungen zum Stellenwertsystem
3	Strategien der Addition (Schrittweises Rechnen, Nachbaraufgaben, Auffüllen) Festigung geschickten Kopfrechnens
4	Einführung und Vertiefung der halbschriftlichen Addition
5	Einführung und Vertiefung der schriftlichen Addition
6	Einführung und Vertiefung der halbschriftlichen Subtraktion
7	Einführung und Vertiefung der schriftlichen Subtraktion
8	Gemischte Übungen zur Subtraktion
9	Strategien und Gesetze der Multiplikation Einführung in die halbschriftliche Multiplikation
10	Fortführung der halbschriftlichen Multiplikation Einführung der schriftlichen Multiplikation
11	Fortführung der schriftlichen Multiplikation Gemischte Übungen zur Multiplikation
12	Einführung und Vertiefung der schriftlichen Division
13	Strategien der Division (Rechnen mit Rest, mit Null im Dividenden, mit Null im Quotienten)
14	Gemischte Übungen zur Division
15	Didaktische Reserve
16	Abschlussstunde

Jede Förderstunde wird strukturell nach den Unterrichtsschritten nach Meyer (2010, S. 40) gestaltet. Nach einer Stundeneröffnung bzw. Begrüßung erfolgt eine Vorwissensaktivierung, dann gemeinsam mit den Schülerinnen und Schülern eine Hinführungsphase, an die sich eine Übungsphase anschließt. Nach einer Zusammenführung der Stundenergebnisse wird ein Ausblick auf die nächste Stunde gegeben und die Schülerinnen und Schüler verabschiedet. Wie diese Phasen konkret ausgestaltet werden, wird im folgenden Kapitel erläutert.

6.3.2 Methodische Prinzipien

„Unterrichtsmethoden sind die Formen und Verfahren, mit denen sich die Lehrerinnen, Lehrer, Schülerinnen und Schüler die sie umgebende natürliche und gesellschaftliche Wirklichkeit unter Beachtung der institutionellen Rahmenbedingungen der Schule aneignen“ (Meyer, 2002, S. 109). Es bestehen hohe Anforderungen an passende – also lernförderliche – Methoden im Mathematikunterricht. Diese Methoden müssen auf drei Ebenen sog. Mindeststandards erfüllen: auf der fachlichen, pädagogischen und methodenkritischen Ebene (Gugel, 2006, S. 12). So sollen Methoden dem Stand der Wissenschaft entsprechen, Gegenpositionen berücksichtigen und Begründungszusammenhänge ermöglichen (fachliche Ebene). Zudem sollten sie das Vorwissen berücksichtigen, Lernbereitschaft wecken, Selbstreflexion und Eigeninitiative fördern sowie die spezifischen Bedingungen der Teilnehmerinnen und Teilnehmer berücksichtigen (pädagogische Ebene). Auf der dritten, methodenkritischen Ebene wird eine Reduzierung der Komplexität von Zusammenhängen der Wirklichkeit, eine Kontrastierung bestimmter Gesichtspunkte, Handlungsorientierung, Verfremdung und Anschaulichkeit gefordert (Gugel, 2006, S. 12).

Die Wahl der Methoden im Rahmen der Maßnahme geschieht auf der Grundlage der Analyse des aktuellen Forschungsstandes zu evidenzbasierten Methoden zur Förderung mathematischer Basiskompetenzen (fachliche Ebene) (siehe Kapitel 4.3 Evidenzbasierte Förderung mathematischer Basiskompetenzen in der Sek. I). Im Rahmen der Analyse ist zum einen festgestellt, ob und welche standardisierten Förderprogramme derzeit für die Sekundarstufe I im deutschsprachigen Raum vorliegen und zum anderen, welche Methoden sich für den Einsatz anbieten. Die Auswahl der Methoden im Rahmen der Maßnahme basiert zum einen auf diesen Rechercheergebnissen. Zum anderen werden pädagogische und methodenkritische Aspekte berücksichtigt, sodass auch methodische Ansätze in die Maßnahme implementiert werden, obwohl sie nicht explizit im Rahmen der Recherche identifiziert sind (Advance Organizer und Spiele).

Die Verankerung der Methoden wird im Folgenden chronologisch, bezogen auf die Verortung innerhalb einer Förderstunde, vorgestellt. Tabelle 30 gibt einen Überblick über einen typischen Stundenverlauf und die darin eingesetzten Methoden. Für die mit einem * versehenen Unterrichtsphasen ist keine „explizite" Methode vorgesehen: die Stundeneröffnung und die Zusammenführung (Ergebnissicherung) erfolgen flexibel durch Lehrervorträge bzw. Gruppengespräche.

TABELLE 30 ÜBERBLICK ÜBER DIE EINGESETZTEN (EVIDENZBASIERTEN) METHODEN IN DEN UNTERRICHTSPHASEN

Unterrichtsphase	Methode	Dauer
Stundeneröffnung/Begrüßung	Lehrervortrag*	3 Minuten
Vorwissensaktivierung	Advance Organizer	5 Minuten
Hinführung	Direkte Instruktion	5–10 Minuten
Übung	Einzelarbeit oder	20 Minuten
	Peer-Tutoring oder	
	Kooperatives Lernen	
Zusammenführung	Gruppengespräch*	max. 5 Minuten
Abschluss	Verhaltensfeedback Spiel	max. 5 Minuten
Nach jeder Einheit: Feeback in der Abschlussphase zum Lernverlauf (insgesamt 6x)		3 Minuten

Jede Stunde wird durch eine Begrüßung durch die Interventionsleiterinnen und -leiter der Lerngruppe eröffnet. Es folgt auf die Begrüßung die Phase der Vorwissensaktivierung.

VORWISSENSAKTIVIERUNG: DER ADVANCE ORGANIZER

Für die Phase der Vorwissensaktivierung wird die Methode des Advance Organizers gewählt, die auf Ausubel (1960) zurückgeht. Es gibt zahlreiche Befunde, die eine Wirksamkeit der Methode belegen. Einen Überblick über Forschungsbefunde bietet Reinck (2016). Folgende Definition verdeutlicht die Methode des Advance Organizers:

> „Der Advance Organizer ist eine Organisationshilfe, die einem neuen Lerninhalt vorausgeht und dessen Verinnerlichung unterstützt. Er ist eine lehrerzentrierte Unterrichtsmethode, die von ihnen [den Lehrkräften, d. Verf.] als Unterrichtseinstieg dargeboten wird." (Martenstein, Melzer & Reinck, 2013, S. 3)

Der Advance Organizer wird folglich einem neuen Lerninhalt vorangestellt mit dem „Ziel [...], die Lücke, die zwischen dem Vorwissen der Lernenden und dem neu zu lernenden Inhalt besteht, gezielt zu überbrücken. Mit ihm [...] [werden] die Lerninhalte in einem übergeordneten Kontext dar[gestellt], auf Basis dessen der nachfolgende Lernstoff besser mit bereits Bekanntem verknüpft und verankert werden kann [...]“ (Martenstein et al., 2013, S. 2). Die Inhalte eines Advance Organizers sind dabei abstrakter, umfassender und allgemeiner als die zu lernenden Inhalte, damit diese nicht vorweggenommen werden.

Im Rahmen der Maßnahme besteht der Advance Organizer aus einem DIN-A-0-Plakat, auf dem fünf Ringe abgedruckt sind, die den Olympischen Ringen ähnlich sind (siehe Abbildung 39). Jeder Ring steht dabei für eine thematische Einheit (schwarz: Stellenwertsystem, blau: Addition, gelb: Subtraktion, grün: Multiplikation, rot: Division). Die Inhalte der Ringe, die die Struktur der Maßnahme repräsentieren, werden im Laufe der 16 Unterrichtsstunden gemeinsam mit den Schülerinnen und Schülern erarbeitet und das Plakat somit Schritt für Schritt gefüllt.

Jeweils in der ersten Stunde einer neuen Einheit wird in einem Interventionsleiter/innen-Schüler/innen-Dialog das Vorwissen zu dem Thema erarbeitet. Folgender Beispieldialog aus der ersten Stunde der Additionseinheit verdeutlicht das Vorgehen. Die Daumenabfrage bedeutet, dass die Schülerinnen und Schüler im Anschluss an die Frage der Interventionsleiterinnen und -leiter (zum Beispiel *Wird etwas hinzugerechnet?)* je nach Zustimmung ihren Daumen nach oben, zur Seite oder nach unten zeigen können.

> „Heute starten wir mit einer neuen Disziplin. Wir befinden uns ab heute im blauen Ring, in dem es um das Plus-Rechnen geht *(+ anheften)*. Daumenabfrage: Was passiert, wenn wir plus rechnen? Wird dann etwas hinzugerechnet? *(Daumenabfrage abwarten und Begriffskarte anheften)* Wird etwas weniger? Wird etwas zusammengezählt? *(Daumenabfrage abwarten und Begriffskarte anheften)* Wird etwas weggenommen?
> Das Plus-Rechnen wird auch „Und-Rechnen“ genannt *(Begriffskarte anheften)*. Der Fachbegriff für das Plus-Rechnen lautet „Addition“ und das passende Verb dazu lautet „addieren“ *(Begriffskarten anheften)*.
> Wir schauen uns ein Beispiel an. Es gibt eine Disziplin, die heißt „Triathlon“: Erst schwimmt man 2 km *(Schwimmer und Zahl anheften)* und fährt dann 80 km Fahrrad *(Fahrrad und Zahl anheften)*. Zum Schluss läuft man 20 km *(Läufer und Zahl anheften)*.
> Hier ist eine Plus-Aufgabe vorhanden. Um herauszufinden, wie viele Kilometer ein Triathlet zurückgelegt hat, werden die Zahlen addiert.
> Man rechnet 2 km+80 km+20 km. Insgesamt legt der Sportler also 102 km zurück *(Aufgabe vervollständigen)*. Ihr seht also, dass hier die Kilometer hinzugerechnet und zusammengezählt wurden *(auf die Begriffe zeigen)*.
> Die einzelnen Teile einer Plus-Rechnung kann man übrigens auch so bezeichnen: Summand + Summand + Summand = Summe *(Begriffe anheften)*.
> Am Ende der nächsten drei Stunden wisst ihr ganz genau, welche Strategien es gibt, Plus-Aufgaben zu rechnen. Darin seid ihr dann ohne Taschenrechner fit!“

Bei der Darstellung eines Advance Organizers ist es von großer Bedeutung, dass die Grundgedanken mit einer hohen Verständlichkeit präsentiert werden, damit die Verknüpfung mit dem Vorwissen erleichtert und nicht erschwert wird (Wahl, 2005, S. 140).

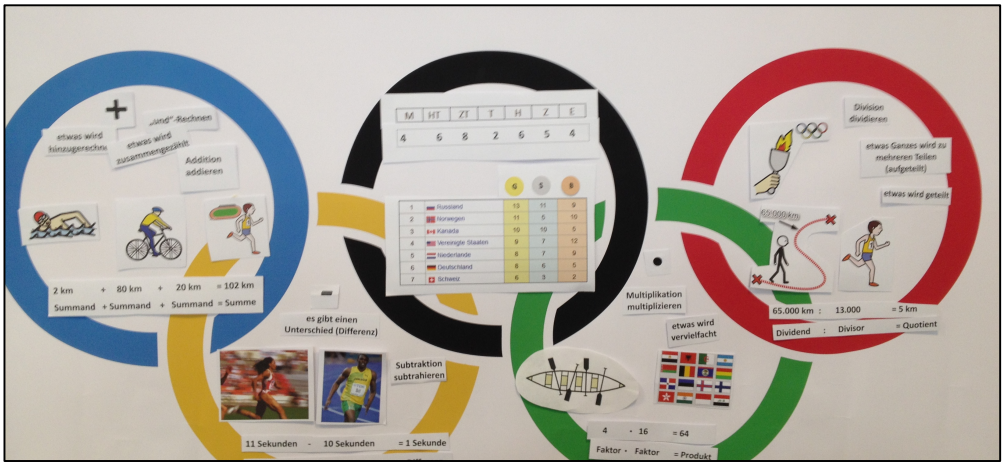


ABBILDUNG 39 VORWISSENSAKTIVIERUNG DURCH DEN ADVANCE-ORGANIZER[2]

Der umfassende Dialog (wie oben) wird nur in der ersten Stunde der jeweiligen Einheit durchgeführt. In den beiden folgenden Stunden wird lediglich vor dem Plakat ein Rückbezug auf den Advance Organizer genommen:

> „Gestern sind wir mit einer neuen Grundrechenart angefangen. Wir befinden uns momentan im blauen Ring *(auf Ring zeigen)*. Das Thema ist also noch Plus-Rechnen. Heute beschäftigen wir uns weiterhin mit Aufgaben, in denen *(Begriffskarten nennen und darauf zeigen)* etwas hinzugerechnet oder etwas zusammengerechnet wird. Das nennen wir auch „Und"-Rechnen oder addieren. Ganz speziell geht es heute um: Halbschriftliches Addieren."

Die Inhalte des Advance Organizer auf dem Plakat füllen sich im zeitlichen Verlauf der Maßnahme. Der Advance Organizer gibt den Schülerinnen und Schülern die Möglichkeit, die neuen Inhalte an bereits vorhandenes Wissen anzuknüpfen und gibt ihnen zudem eine Orientierung über die inhaltliche Struktur der Förderung. Nach Wahl (2005, S. 141 f.) erbringt die Methode des Advance Organizers folgende Wirkungen:

- „Erhöhtes Interesse für die neuen Inhalte,
- gerichtete Aufmerksamkeit,
- gesteigerte Selbstwirksamkeit,
- bessere Orientierung,
- besseres Behalten,
- weniger Missverständnisse,
- besserer Transfer."

Im Anschluss an die Vorwissensaktivierung mit dem Advance Organizer, die ca. fünf Minuten dauert, erfolgt die Hinführungsphase.

HINFÜHRUNG: DIE DIREKTE INSTRUKTION

Die Methode der Direkten Instruktion ist in Kapitel 4.3.3.2 vor allem auf Grundlage der Gruppenvergleichsstudien von Engelmann (2007) und Tournaki (2003) als eindeutig evidenzbasierte

[2] Die in den Materialien verwendeten Piktogramme stammen aus der Freeware „Picto Selector".

Methode identifiziert, die sich für den Einsatz im Bereich Mathematik eignet. Auf die detaillierte theoretische und konzeptionelle Beschreibung der Methode der Direkten Instruktion wird auf das Kapitel 4.3.2.2 verwiesen.

Die Direkte Instruktion wird in der Maßnahme zur Wiederholung der zentralen Rechenstrategien im Rahmen der Hinführungsphase verwendet. Nachdem die Vorwissensaktivierung durch den Advance Organizer durchgeführt ist, erfolgt eine ca. fünfminütige lehrerzentrierte Phase der Direkten Instruktion an der Tafel. Dabei wird die Schrittfolge nach Lebens und Lauth (2014, S. 424) verfolgt (siehe Kapitel 4.3.3.2). Folgender Auszug aus der zweiten Stunde der Additionseinheit verdeutlicht das Vorgehen. Das Tafelbild sollte dabei bereits vor Beginn der Stunde vorbereitet werden.

> „Heute ist unser Ziel, dass alle Schülerinnen und Schüler am Ende der Unterrichtsstunde halbschriftlich, also in kleinen Schritten, Plusrechnen können. Dazu habe ich euch zwei mögliche Strategien mitgebracht: stellenweises Rechnen und schrittweises Rechnen. Diese beiden Strategien möchte ich euch jetzt an der Tafel zeigen:
>
> *(Tafelbild)*

Stellenweises Rechnen										
5	4	7	+	1	4	6	=			
5	0	0	+	1	0	0	=	6	0	0
	4	0	+		4	0	=		8	0
		7	+			6	=		1	3
							=	**6**	**9**	**3**

> „Die Aufgabe lautet: 547+146. Eine Vorgehensweise zum Lösen der Aufgabe ist, dass ihr zunächst die Hunderter aufschreibt und zusammenrechnet (*oder: addiert*) – also 500 plus 100. Das ergibt 600. Danach schaut ihr euch die Zehner an und rechnet diese zusammen (40+40). Das ergibt 80. Als letztes rechnet ihr die Einer zusammen 7 plus 6. Das ergibt 13. Hier müsst ihr aufpassen, da ihr die 1 unter die Zehner schreiben müsst und die 3 unter die Einer. Die Hunderterstelle bleibt also frei. Wenn wir unten alle Teilergebnisse zusammenrechnen (600+80+13), erhalten wir am Ende 693 als Endergebnis."
>
> *(Lehrkraft spricht vor, die Schülerinnen und Schüler sprechen im Chor nach:)*
>
> „Beim *stellenweisen* Plusrechnen zerlege ich zunächst die Summanden in Hunderter, Zehner und Einer. Ich rechne dann die Hunderter, Zehner und Einer zusammen. Am Ende addiere ich alle Teilergebnisse und erhalte das Ergebnis der Plusaufgabe."
>
> „Ich zeige euch jetzt noch eine zweite Möglichkeit, wie ihr diese Aufgaben lösen könnt."
>
> *(Tafelbild)*

Schrittweises Rechnen										
5	4	7	+	1	4	6	=			
5	4	7	+	1	0	0	=	6	4	7
6	4	7	+		4	0	=	6	8	7
6	8	7	+			6	=	**6**	**9**	**3**

> „Die Aufgabe lautet auch hier wieder: 547+146. Hier rechnet ihr zu 547 zunächst die Hunderter – also 100 dazu. Ihr erhaltet 647. Zu diesen 647 rechnet ihr in der nächsten Zeile die Zehner

> dazu, das sind hier 40. Ihr erhaltet also 687. Zu diesen 687 rechnet ihr im letzten Schritt die Einer. Bei unserer Aufgabe sind das 6. Achtet auf den Übertrag! Wir erhalten als Ergebnis auch hier 693."
>
> *(Lehrkraft spricht vor, die Schülerinnen und Schüler sprechen im Chor nach:)*
>
> „Beim *schrittweisen* Plusrechnen, rechne ich zum Summanden erst die Hunderter, zu dem neuen Ergebnis dann die Zehner und zum Schluss die Einer. Somit erhalte ich das Endergebnis"

Da Gersten und Kollegen (2009) für Visualisierungshilfen im Mathematikunterricht eine Effekstärke von $d = .47$ ermitteln (siehe Tabelle 22), wird ergänzend zur Direkten Instruktion die erarbeiteten Rechenstrategien und Hinweise auf Fallstricke (zum Beispiel *„Genau untereinanderschreiben!"*, *„Den Übertrag nicht vergessen!"* u. ä.) auf DIN-A-3-Plakaten an der Tafel fixiert. Neben dem Advance Organizer bleiben auch diese Hinweisschilder während der gesamten Maßnahme für die Schülerinnen und Schüler sichtbar, um eine lernförderliche Unterstützung anzubieten.

Im Anschlus an die Hinführungsphase folgt die Übungsphase.

Übung: Einzelarbeit, Peer-Tutoring oder kooperatives Lernen

Die Übungsphase wird verschiedentlich entweder als Einzelarbeitsphase, in Form des Peer-Tutorings oder in kooperativen Lerngruppen durchgeführt. In der Übungsphase stehen die Anwendung und die Vertiefung der entsprechenden Rechenstrategien der Stunde im Fokus.

In der Einzelarbeitsphase kommen Arbeitsblätter zum Einsatz, die ein bestimmtes Gestaltungsprinzip verfolgen (siehe Abbildung 40): Oben rechts in der Ecke auf dem Arbeitsblatt ist ein Symbol angegeben, das den Schülerinnen und Schülern verdeutlicht, in welcher Einheit sie sich gerade befinden (siehe a) in Abbildung 40). Gersten und Kollegen (2009) ermitteln zudem für das Lernen anhand von Beispielen eine Effektstärke von $d = .82$, sodass Beispielaufgaben auf den Arbeitsblättern ein lernförderliches Element darstellen. Somit befinden sich auf den Arbeitsblättern auch Beispiele, wie die Aufgaben zu lösen sind (siehe b) in Abbildung 40. Die dargebotenen Beispiele ähneln den Aufgaben, die zuvor im Rahmen der Direkten Instruktion präsentiert wurden. Zudem ermöglichen verschiedene Formen der Selbstkontrolle auf den Arbeitsblättern das Überprüfen der eigenen Lösungen (siehe c) in Abbildung 40).

a)

b)

1. Beispielaufgabe:

(Stellenweises Rechnen)

5	**2**	**3**	**+**	**1**	**6**	**7**	**=**	**6**	**9**	**0**
5	0	0	+	1	0	0	=	6	0	0
	2	0	+		6	0	=		8	0
		3	+			7	=		1	0
							=	**6**	**9**	**0**

2. Beispielaufgabe:

(Schrittweises Rechnen)

5	**2**	**3**	**+**	**1**	**6**	**7**	**=**	**6**	**9**	0
5	2	3	+	1	0	0	=	6	2	3
6	2	3	+		6	0	=	6	8	3
6	8	3	+			7	=	**6**	**9**	**0**

c)

1. Aufgabe: Rechne mit den Strategien! Lies die Aufgabe. Wähle die passende Strategie und notiere diese in das Kästchen. Kontrolliere deine Ergebnisse unten.
Tipp: Es gibt mehr Zahlen in der Kontrolle als du brauchst.

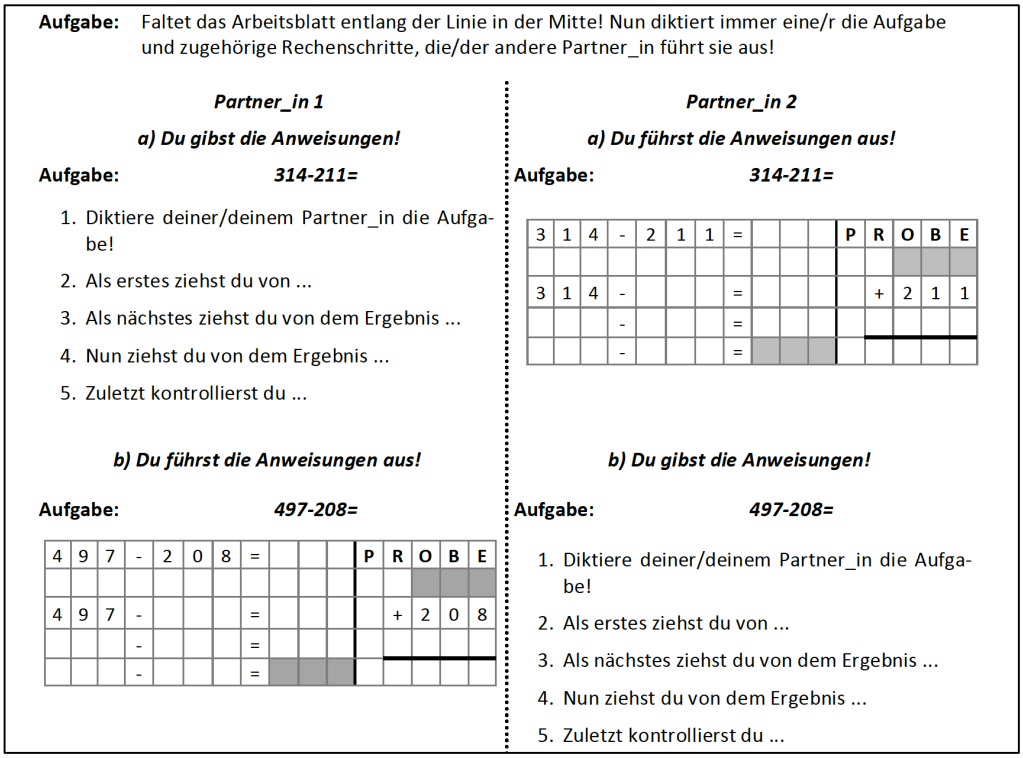
2. Aufgabe: Partnercheck! Wähle eine Aufgabe aus und besprich deine Strategie

ABBILDUNG 40 ELEMENTE DER ARBEITSBLÄTTER A) ORIENTIERUNGSHILFE, B) BEISPIELAUFGABEN, C) INSTRUKTION UND SELBSTKONTROLLE

Alle Arbeitsblätter liegen für drei Differenzierungsstufen (rechenschwach, durchschnittlich, rechenstark) vor. Die Differenzierungsstufe ist ggf. unten auf dem Arbeitsblatt notiert (l. für leicht oder s. für schwer). Dabei gilt es nicht immer, Aufgaben eigens zu lösen, sondern zum Beispiel auch bereits gelöste Aufgaben zu korrigieren, in die Fehler eingebaut sind (sog. Fehlerteufel-Aufgaben) (Gaidoschik, 2010).

Neben Einzelarbeitsphasen wird auch die Methode des peergestützten Lernens umgesetzt. Grundsätzlich können Gruppen- und Partnerarbeiten soziale Kompetenzen und die Übernahme von Verantwortung fördern (Thal & Ebert, 2001, S. 61). Durch die gegenseitige Verantwortlichkeit für den Lernerfolg des Anderen wird eine aktive Lernhaltung eingenommen und Wissen vertieft (Haag, 2014, S. 463). Die peergestützte Methode kann zudem ein Klima unterstützenden und hilfsbereiten Lernens innerhalb der Klasse fördern (Souvignier, 2007). Für eine weiterführende Beschreibung der Methode wird auf das Kapitel 4.3.3.2 verwiesen.

Die Schülerinnen und Schüler werden für die Methode des peergestützten Lernens direkt zu Beginn in feste heterogene Tandems entsprechend ihrer Mathematikleistungen aus der Prä-Testung eingeteilt. Dafür erfolgt eine Medianaufteilung der Klasse entsprechend der im Vorfeld erhobenen Mathematikleistungen (vgl. Vorgehen bei Fuchs, Fuchs, Yazdian, et al., 2002; Spörer, 2009). Die Schülerinnen und Schüler werden dann in eine Rangfolge gebracht und gematcht (zum Beispiel Reihenfolge: 1, 2, 3, 4, 5 $|_{Median}$ 6, 7, 8, 9, 10 → Tandems: 1 & 6; 2 & 7; 3 & 8 etc.). Die Zusammensetzung der Tandems ändert sich im Laufe der Maßnahme nicht. Die Tandems lösen unterschiedliche Aufgabenformate. Eine ausgewählte Aufgabe für die Tandems ist in Abbildung 41 dargestellt. Auch für diese Übung liegt den Schülerinnen und Schülern eine Beispiellösung vor.



Aufgabe: Faltet das Arbeitsblatt entlang der Linie in der Mitte! Nun diktiert immer eine/r die Aufgabe und zugehörige Rechenschritte, die/der andere Partner_in führt sie aus!

Partner_in 1

a) Du gibst die Anweisungen!

Aufgabe: ***314-211=***

1. Diktiere deiner/deinem Partner_in die Aufgabe!
2. Als erstes ziehst du von ...
3. Als nächstes ziehst du von dem Ergebnis ...
4. Nun ziehst du von dem Ergebnis ...
5. Zuletzt kontrollierst du ...

b) Du führst die Anweisungen aus!

Aufgabe: ***497-208=***

										P	R	O	B	E
4	9	7	-				=				+	2	0	8
			-				=							
			-				=							

Partner_in 2

a) Du führst die Anweisungen aus!

Aufgabe: ***314-211=***

3	1	4	-	2	1	1	=				P	R	O	B	E
3	1	4	-				=					+	2	1	1
			-				=								
			-				=								

b) Du gibst die Anweisungen!

Aufgabe: ***497-208=***

1. Diktiere deiner/deinem Partner_in die Aufgabe!
2. Als erstes ziehst du von ...
3. Als nächstes ziehst du von dem Ergebnis ...
4. Nun ziehst du von dem Ergebnis ...
5. Zuletzt kontrollierst du ...

ABBILDUNG 41 AUSZUG AUS EINER PEERGESTÜTZTEN ÜBUNG

Neben Einzelarbeit und peergestütztem Lernen werden auch kooperative Lernformen in Form von Kleingruppenarbeiten umgesetzt. Die Interventionsleiterinnen und -leiter teilen dabei die Schülerinnen und Schüler entsprechend ihrer Leistungen der Prä-Testung in Quartile ein und setzen so die Gruppen zusammen (zum Beispiel Reihenfolge: 1, 2, 3, 4 $|_{1.\text{Quartil}}$ 5, 6, 7, 8 $|_{2.\text{Quartil}}$ 9, 10, 11, 12 $|_{3.\text{Quartil}}$ 13, 14, 15, 16 $|_{4.\text{Quartil}}$ → Kleingruppen: 1, 5, 9, 13; 2, 6, 10, 14 etc.). Auch dieses Vorgehen ermöglicht das Arbeiten in heterogenen Kleingruppen. Für die Arbeitsphasen in den kooperativen Lerngruppen erhalten die Schülerinnen und Schüler Rollenkarten, die ihnen zeigen, welche Funktion und Aufgabe sie beim Lösen der vorliegenden Aufgabe haben. Die Verteilung der Rollenkarten vor den Arbeitsphasen rotiert dabei lehrergesteuert von Kleingruppenphase zu Kleingruppenphase, damit die Schülerinnen und Schüler nicht dauerhaft in einer Rolle verbleiben und die Möglichkeit haben, eine andere Rolle einzunehmen (s.Vorgehen bei Fuchs, Fuchs, Yazdian, et al., 2002; Tsuei, 2014).

ABBILDUNG 42 DIE ROLLENKARTEN DER KOOPERATIVEN LERNGRUPPEN

Eine Aufgabe der kooperativen Kleingruppen ist zum Beispiel eine Rechenpyramide. Diese wird erst in Einzelarbeit gelöst, dann klebt sie jede/r auf ein gemeinsames Placemat und dann

werden die verschiedenen Lösungen in der Kleingruppe besprochen. Sowohl für die peergestützten als auch für die kooperativen Arbeitsphasen erhalten die Schülerinnen und Schüler zu Beginn einen kurzen Handlungsleitfaden, auf dem das Vorgehen noch einmal beschrieben ist.

Abschluss einer Stunde: Verhaltensfeedback & Spiel

In der Abschlussphase der Unterrichtsstunde erfolgen das Verhaltensfeedback auf dem Stadionplan und ein an den Stundeninhalt angelehntes Abschlussspiel. Im Rahmen der Maßnahme werden in der ersten Sitzung vier Regeln aufgestellt: „Ich lasse jedes Mitglied aussprechen“, „Ich höre zu, wenn jemand anderes spricht“, „Ich lache nicht über Antworten oder Lösungsvorschläge“ und „Ich beteilige mich am Unterricht“. Die Interventionsleiterinnen und -leiter verfügen während der Durchführung über eine gelbe und rote Karte. Hält sich ein Schüler oder eine Schülerin nicht an die besprochenen Regeln, wird er oder sie mit einer gelben Karte verwarnt. Erfolgt erneut ein Regelverstoß, wird die rote Karte ausgehändigt. Auf dem individuellen Stadionplan (siehe Abbildung 43) erhalten die Schülerinnen und Schüler am Ende der Stunde einen Stempel, der jedoch bei Erhalt einer roten Karte nicht mehr gegeben wird. Die Schülerinnen und Schüler haben die Möglichkeit, in den 16 Stunden Stempel zu sammeln und am Ende der Maßnahme einen Rang zu erlangen. Nur wer alle Stempel erhält, wird zum Olympioniken bzw. zur Olympionikin gekürt (zum Beispiel durche eine Medaille).

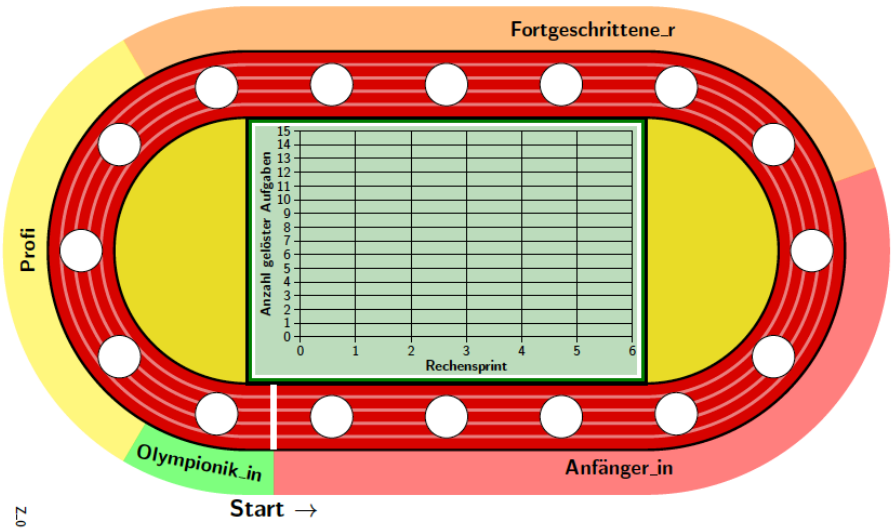


Abbildung 43 Der Stadionplan mit Verhaltensfeedback und dem Lernverlauf

Das Abschlussspiel, das stets mit Rechenaufgaben verknüpft ist (zum Beispiel Tafelfußball o. ä.) dient der Entspannung der Schülerinnen und Schüler und einem positiven Stundenausklang. Die Durchführung der Maßnahme in Form von Interventionsleiter/innen-Tandems ermöglicht es, diese beiden Phasen parallel stattfinden zu lassen.

Neben dem Verhaltensfeedback durch Stempel erfolgt in der Mitte des Stadionplans ebenso die Lernverlaufsdiagnostik.

Abschluss einer Einheit: Lernverlaufsdiagnostik

In der Analyse der potenziell evidenzbasierten Methoden wird das curriculumbasierte Messen als eine lernförderliche Feedbackmethode identifiziert, sodass es auch im Rahmen der Maßnahme Anwendungen findet. Die curriculumbasierten Messverfahren werden parallel zur Maßnahme entwickelt und im Vorfeld der Hauptstudie mit einer unabhängigen Stichprobe positiv pilotiert (siehe Tabelle 42). Die eingesetzten Testbögen entsprechen den Gütekriterien Reliabilität, Validität und Änderungssensitivität (Reliabilität: Käter, 2014; Änderungssensitivität: Rensing, 2014; Validität: Wunram, 2014).

Die Schülerinnen und Schüler lösen in der Fördermaßnahme nach der Stunde zum Stellenwertsystem (zweite Stunde), nach jeweils der dritten Stunde der Einheit zu den Grundrechenarten (Stunde 5, 8, 11 und 14) und im Anschluss an die Abschlussstunde (Stunde 16) einen „Rechensprint". Dieser dauert drei Minuten und umfasst Aufgaben, für die im Rahmen der Maßnahme Rechenstrategien geübt und gefördert werden. Ein Rechensprint enthält insgesamt 15 Aufgaben zum Stellenwertsystem und zu den Grundrechenarten (je 3 Aufgaben), die innerhalb des CBM-Testbogens in der Schwierigkeit ansteigen. Die Anzahl richtig gerechneter Aufgaben wird den Schülerinnen und Schülern in der folgenden Stunde genannt, sodass diese ihr Ergebnis in ihr Diagramm innerhalb des Stadionplans eintragen und einen individuellen Lernverlauf erkennen können. Da nicht davon ausgegangen werden kann, dass alle Schülerinnen und Schüler am Anfang der fünften Klasse über die benötigte Diagrammkompetenz verfügen, wird das Eintragen der Werte und die Interpretation beispielhaft an der Tafel vorgemacht. Folgende Tabelle 31 gibt einen Überblick über den zeitlichen Einsatz der curriculumbasierten Messverfahren innerhalb der Maßnahme.

Tabelle 31 Überblick über den Einsatz der CBM

Stunde 1	Eröffnungsfeier		
Stunde 2	Stellenwertsystem **CBM 1**		
Stunden 3–5	Addition I	Addition II	Addition III **CBM 2**
Stunden 6–8	Subtraktion I	Subtraktion II	Subtraktion III **CBM 3**
Stunden 9–11	Multiplikation I	Multiplikation II	Multiplikation III **CBM 4**
Stunden 12–14	Division I	Division II	Division III **CBM 5**
Stunden 15	Didaktische Reserve		
Stunde 16	Abschlussfeier **CBM 6**		

Es kommen also insgesamt sechs Rechensprints zum Einsatz. Es liegt für jeden Rechensprint eine a- und eine b-Version vor, um ein gegenseitiges Abschauen zu vermeiden.

7 Forschungsziel, -fragen und Hypothesen

Nachdem die didaktischen und methodischen Prinzipien der Fördermaßnahme dargestellt sind, wird im Folgenden die empirische Untersuchung zur Evaluation und Implementation der Fördermaßnahme dargelegt. Dafür erfolgt eingangs die Formulierung des Forschungsziels und der Forschungsfragen (Kapitel 7.1) sowie der Hypothesen (Kapitel 7.2).

7.1 Forschungsziel und -fragen

Im vorherigen Kapitel 6 wird die Entwicklung einer Maßnahme zur Förderung mathematischer Basiskompetenzen für Schülerinnen und Schüler zu Beginn der Sekundarstufe I vorgestellt, die sich durch die nach spezifischen Kriterien ausgewählten Implementation evidenzbasierter Fördermethoden auszeichnet. Neben der Direkten Instruktion, peergestützten und kooperativen Lernformen sowie dem formativen Leistungsfeedback kommen ein Advance Organizer zur Vorwissensaktivierung, ein Verhaltensfeedback und Spiele zum Stundenabschluss im Rahmen der Fördermaßnahme zum Einsatz.

Forschungsziel dieser Arbeit ist es, die kurz- und langfristige Wirksamkeit der Maßnahme sowie deren Implementationsqualität zu evaluieren und damit einen praxiserprobten Beitrag zu Möglichkeiten evidenzbasierter Förderung mathematischer Basiskompetenzen in der Sekundarstufe I zu leisten.

Es leiten sich drei Forschungsfragen ab:

Forschungsfrage I: Führt die vierwöchige Maßnahme zur Förderung mathematischer Basiskompetenzen zu Beginn der Sekundarstufe I zu einer kurz und/oder langfristigen Verbesserung der mathematischen Basiskompetenzen bei den Schülerinnen und Schülern und dabei insbesondere bei denjenigen, die ungünstige Lernvoraussetzungen für ihren mathematischen Kompetenzerwerb aufweisen?

Forschungsfrage II*:* Können Unterschiede in der Lernfreude oder in dem mathematischen Leistungsspektrum der Klasse die kurzfristige Wirksamkeit der Fördermaßnahme vorhersagen?

Forschungsfrage III: Verfügt die Fördermaßnahme hinsichtlich verschiedener Implementationsvariablen aus Sicht der Schülerinnen und Schüler und aus Sicht der Interventionsleiterinnen und -leiter über eine gute Implementationsqualität?

Entsprechend der theoretischen Fundierung werden im Folgenden 15 Hypothesen bzgl. der Wirksamkeit, der potenziellen Prädiktoren und der Implementationsqualität literaturbasiert begründet und aufgestellt.

7.2 Hypothesen

Der Hypothesenblock I widmet sich der Wirksamkeit der Maßnahme bezogen auf alle Schülerinnen und Schüler, die an der Fördermaßnahme teilnehmen, und der Hypothesenblock II auf

diejenigen Schülerinnen und Schüler, die ungünstige Lernvoraussetzungen für ihren mathematischen Kompetenzerwerb aufweisen (Evaluation der differenziellen Wirksamkeit). Der Hypothesenblock III umfasst potenzielle Prädiktoren und der Hypothesenblock IV die Implementationsqualität der Fördermaßnahme.

7.2.1 Hypothesenblock I: Wirksamkeit auf universeller Ebene

Die Fördermaßnahme verfolgt das Ziel, auf universeller Ebene (National Center on Response to Intervention, 2010) die mathematischen Basiskompetenzen der Schülerinnen und Schüler der Experimentalgruppe zu Beginn der Sekundarstufe I sowohl kurz als auch langfristig zu verbessern. Im Vergleich zur Kontrollgruppe wird eine positivere mathematische Entwicklung erwartet.

Hypothese 1: Die Schülerinnen und Schüler, die an der vierwöchigen Maßnahme zur Förderung der mathematischen Basiskompetenzen teilnehmen, zeigen im Prä-Post-Vergleich (kurzfristig) und im Zeitraum zwischen der Post-, Follow-Up-I- und Follow-Up-II-Erhebung (langfristig) eine bessere Entwicklung in den mathematischen Basiskompetenzen als die Schülerinnen und Schüler der Kontrollgruppe.

Prä-Post-Vergleich:

Mit $g \in \{EG, KG\}$ und $t \in \{t_1, t_2\}$ gilt

$\boldsymbol{H_0^{1a}}$: $\forall g, t: \mu_{g,t} = \mu_g + \mu_t - \mu$
$\vee\ \mu_{EG,t_2} - \mu_{EG,t_1} \not> \mu_{KG,t_2} - \mu_{KG,t_1}$

$\boldsymbol{H_1^{1a}}$: $\exists g, t: \mu_{g,t} \neq \mu_g + \mu_t - \mu$
$\wedge\ \mu_{EG,t_2} - \mu_{EG,t_1} > \mu_{KG,t_2} - \mu_{KG,t_1}$

Post-Follow-Up-I-Follow-Up-II-Vergleich:

Mit $g \in \{EG, KG\}$ und $t \in \{t_2, t_3, t_4\}$ gilt

$\boldsymbol{H_0^{1b}}$: $\forall g, t: \mu_{g,t} = \mu_g + \mu_t - \mu$
$\vee\ \mu_{EG,t_3} - \mu_{EG,t_2} \not> \mu_{KG,t_3} - \mu_{KG,t_2}$
$\vee\ \mu_{EG,t_4} - \mu_{EG,t_3} \not> \mu_{KG,t_4} - \mu_{KG,t_3}$
$\wedge$
$(\exists t: \mu_{EG,t} \not> \mu_{KG,t}$
$\vee\ H_0^{1a})$

$\boldsymbol{H_1^{1b}}$: $(\exists g, t: \mu_{g,t} = \mu_g + \mu_t - \mu$
$\wedge\ \mu_{EG,t_3} - \mu_{EG,t_2} > \mu_{KG,t_3} - \mu_{KG,t_2}$
$\wedge\ \mu_{EG,t_4} - \mu_{EG,t_3} > \mu_{KG,t_4} - \mu_{KG,t_3})$
$\vee$
$(\forall t: \mu_{EG,t} > \mu_{KG,t}$
$\wedge\ H_1^{1a})$

Anmerkung: EG = Experimentalgruppe, KG = Kontrollgruppe, t = Messzeitpunkt, g = Gruppe

7.2.2 Hypothesenblock II: Wirksamkeit bei Schülerinnen und Schüler mit ungünstigen Lernvoraussetzungen

Bestimmte Faktoren können den (mathematischen) Kompetenzerwerb bei Schülerinnen und Schülern erschweren. Dazu gehören zum Beispiel ein geringes Vorwissen, geringe kognitive Fähigkeiten, eine geringe Lesekompetenz oder ein risikobehaftetes Sozial- und/oder Lernverhalten. Auch ein vorliegender sonderpädagogischer Unterstützungsbedarf weist auf ein erschwertes mathematisches Lernen hin.

Im Hypothesenblock II wird die differenzielle kurz- und langfristige Wirksamkeit der Maßnahme bei Schülerinnen und Schülern mit eben diesen ungünstigen Lernvoraussetzungen betrachtet. Folgende Tabelle 32 gibt einen Überblick über die Cut-Off-Werte, die – im Regelfall

auf Grundlage der Testmanuale – begründet und gewählt werden, um die Subgruppen zu bestimmen. Die Tabelle dient dabei einem Überblick; auf die konkrete Beschreibung der Messinstrumente wird auf Kapitel 8.2 verwiesen.

TABELLE 32 THEORIEBASIERTE BESTIMMUNG DER SCHÜLERINNEN UND SCHÜLER MIT UNGÜNSTIGEN LERNVORAUSSETZUNGEN

Risikobedingung	Cut-Off-Werte zur Bestimmung der Kriterien
Geringes Vorwissen	mT1-Wert im Heidelberger Rechentest 1–4 (Subskala Rechenoperationen) in der Präerhebung < 40
Geringe kognitive Fähigkeiten	Intelligenzquotient (IQ) im Culture Fair Intelligence Test 20-R (Teil 1) in der Präerhebung < 85
Geringe Lesekompetenz	Lesequotient (LQ) im Salzburger Lesescreening 1–4 in der Präerhebung < 90
Unzureichendes Sozialverhalten	T-Wert in der Skala Sozialverhalten in der Lehrereinschätzliste für das Sozial- und Lernverhalten in der Post-Erhebung < 40
Unzureichendes Lernverhalten	T-Wert in der Skala Lernverhalten in der Lehrereinschätzliste für das Sozial- und Lernverhalten in der Post-Erhebung < 40
Sonderpädagogischer Unterstützungsbedarf	Es liegt ein diagnostizierter sonderpädagogischer Unterstützungsbedarf im Bereich Lernen vor.

Unter die Subgruppe „geringes Vorwissen“ fallen die Schülerinnen und Schüler, die in der Erhebung der mathematischen Kompetenzen (HRT 1–4 in der Prä-Erhebung) einen T-Wert von kleiner als 40 erreichen. Die Testautoren Haffner und Kollegen (2005, S. 19) schreiben, dass „anhand der T-Werte [...] Testleistungen direkt miteinander verglichen werden [können]“. Der durchschnittliche T-Wert im HRT 1–4 liegt bei 50, eine Standardabweichung beträgt 10 T-Wert-Punkte. Angenommen wird, dass eine Abweichung um mehr als eine Standardabweichung von dem Mittelwert als Risikobereich gilt, sodass der Cut-Off-Wert für ein geringes Vorwissen bei 40 oder kleiner liegt. Gleiches gilt für die Skalen Sozial- und Lernverhalten in der Lehrereinschätzliste für Sozial- und Lernverhalten (Petermann & Petermann, 2013). Bezüglich der Lesekompetenz wird der Risikobereich bei einem Lesequotienten von kleiner als 90 angelegt, da die Testautoren des Salzburger Lesescreenings 1–4 Mayringer und Wimmer (2003, S. 22) die Leseleistungen ab diesem Wert als unterdurchschnittlich bezeichnen (siehe Tabelle 36). Beim IQ beträgt der Mittelwert 100 und die Standardabweichung 15 IQ-Punkte (Schmidt-Atzert & Amelang, 2012, S. 161). Somit wird hier ein IQ von kleiner als 85 als Cut-Off-Wert für geringe kognitive Fähigkeiten angenommen. Der sonderpädagogische Unterstützungsbedarf wird hier auf den Bereich Lernen bezogen und liegt zum Zeitpunkt der Befragung mittels eines Gutachtens diagnostiziert vor.

7.2.2.1 Geringes Vorwissen

„Vorwissen und begriffliche Schemata sind zwingend erforderlich, um sich neue Kenntnisse zuverlässig und dauerhaft aneignen zu können“ (Lauth et al., 2014, S. 21). Schülerinnen und Schüler mit einem geringen Vorwissen haben oftmals Probleme, sich neue Wissensinhalte anzueignen. Dies kann gerade in Transitionen, die für einige Schülerinnen und Schüler ein kritisches Lebensereignis darstellen, deutlich werden (Ufer, 2009). Oftmals ist das geringe Vorwissen verknüpft mit unzulänglichen Lernstrategien, die sich schon über eine lange Zeit verfestigt haben (Lauth & Mackowiak, 2006). Die SCHOLASTIK-Studie zeigt für den Erwerb mathematischer Kompetenzen innerhalb der Primarstufe, dass der Einfluss der Intelligenz auf die Mathematikleistung zurückgeht und die Bedeutsamkeit des Vorwissens steigt. „Eine hohe Intelligenzleistung kann offensichtlich nur in geringem Maße Defizite in der Lerngeschichte kompensieren“ (Stern, 1997, S. 160).

Die hier entwickelte Maßnahme versucht einem unzureichenden Vorwissen auf verschiedenen Ebenen zu begegnen. Zum einen sind die Inhalte der Fördermaßnahme eine Wiederholung des Grundschulstoffs, damit gerade die Schülerinnen und Schüler mit einem geringen Vorwissen Anschluss an die Inhalte der Sekundarstufe I finden. Zum anderen eignen sich der Advance Organizer und auch die Direkte Instruktion mit den sich anschließenden Übungen besonders dazu, Vorwissen zu aktivieren und zu strukturieren (Reinck, 2016). Aufgrund der Ausrichtung der Maßnahme u. a. auf Schülerinnen und Schüler mit einem geringen Vorwissen wird angenommen, dass sich diese Schülerinnen und Schüler in ihren mathematischen Basiskompetenzen besser entwickeln als diejenigen mit einem geringen Vorwissen in der Kontrollgruppe.

Hypothese 2: Die Schülerinnen und Schüler mit einem geringen Vorwissen, die an der vierwöchigen Maßnahme zur Förderung der mathematischen Basiskompetenzen teilnehmen, zeigen im Prä-Post-Vergleich (kurzfristig) und im Zeitraum zwischen der Post-, Follow-Up-I- und Follow-Up-II-Erhebung (langfristig) eine bessere Entwicklung in den mathematischen Basiskompetenzen als die Schülerinnen und Schüler mit einem geringen Vorwissen in der Kontrollgruppe.

Prä-Post-Vergleich:
Mit $g \in \{EG_{gV}, KG_{gV}\}$ und $t \in \{t_1, t_2\}$ gilt

$\boldsymbol{H_0^{2a}}$: $\forall g, t: \mu_{g,t} = \mu_g + \mu_t - \mu$
$\vee\ \mu_{EG_{gV},t_2} - \mu_{EG_{gV},t_1} \not> \mu_{KG_{gV},t_2} - \mu_{KG_{gV},t_1}$

$\boldsymbol{H_1^{2a}}$: $\exists g, t: \mu_{g,t} \neq \mu_g + \mu_t - \mu$
$\wedge\ \mu_{EG_{gV},t_2} - \mu_{EG_{gV},t_1} > \mu_{KG_{gV},t_2} - \mu_{KG_{gV},t_1}$

Post-Follow-Up-I-Follow-Up-II-Vergleich:
Mit $g \in \{EG_{gV}, KG_{gV}\}$ und $t \in \{t_2, t_3, t_4\}$ gilt

$\boldsymbol{H_0^{2b}}$: $\forall g, t: \mu_{g,t} = \mu_g + \mu_t - \mu$
$\vee\ \mu_{EG_{gV},t_3} - \mu_{EG_{gV},t_2} \not> \mu_{KG_{gV},t_3} - \mu_{KG_{gV},t_2}$
$\vee\ \mu_{EG_{gV},t_4} - \mu_{EG_{gV},t_3} \not> \mu_{KG_{gV},t_4} - \mu_{KG_{gV},t_3}$
$\wedge$
$(\exists t: \mu_{EG_{gV},t} \not> \mu_{KG_{gV},t}$
$\vee\ H_0^{2a})$

$\boldsymbol{H_1^{2b}}$: $(\exists g, t: \mu_{g,t} = \mu_g + \mu_t - \mu$
$\wedge\ \mu_{EG_{gV},t_3} - \mu_{EG_{gV},t_2} > \mu_{KG_{gV},t_3} - \mu_{KG_{gV},t_2}$
$\wedge\ \mu_{EG_{gV},t_4} - \mu_{EG_{gV},t_3} > \mu_{KG_{gV},t_4} - \mu_{KG_{gV},t_3})$
$\vee$
$(\forall t: \mu_{EG_{gV},t} > \mu_{KG_{gV},t}$
$\wedge\ H_1^{2a})$

Anmerkung: EG = Experimentalgruppe, KG = Kontrollgruppe, t = Messzeitpunkt, g = Gruppe, gV = geringes Vorwissen

7.2.2.2 Geringe kognitive Fähigkeiten

Mathematische Fähigkeiten hängen eng mit kognitiven Fähigkeiten zusammen, wie zum Beispiel die QuaSUM-Studie (Lehmann, Peek, et al., 2001) zeigt. Sie belegt, dass die Leistungen im Fach Mathematik in der fünften Klasse am stärksten durch kognitive Voraussetzungen aufgeklärt werden können. Moser Opitz und Kollegen (2010, S. 32) zeigen, dass rechenschwache Schülerinnen und Schüler geringer ausgeprägte kognitive Fähigkeiten aufweisen können (Moser Opitz et al., 2010, S. 32). Krajewski (2008) stellt die Intelligenz schon im Vorschulalter als den wichtigsten Prädiktor mathematischen Kompetenzerwerbs heraus und Stern (2008) belegt eine Korrelation von $r = .43$ zwischen dem numerischen Verständnis im Alter von fünf Jahren und der mathematischen Intelligenz im Alter von 23 Jahren (vgl. Abbildung 4). Schwierigkeiten im Rechnen können dabei mit einer Überlastung des Arbeitsgedächtnisses begründet werden (siehe Kapitel 3.3).

Um die mathematischen Basiskompetenzen bei Schülerinnen und Schülern mit geringen kognitiven Leistungen fördern zu können, werden in der Fördermaßnahme Methoden ausgewählt, die eine Entlastung des Arbeitsgedächtnisses ermöglichen: Der Advance Organizer strukturiert zum Beispiel das Vorwissen und ermöglicht Schülerinnen und Schülern mit geringen kognitiven Fähigkeiten eine leichte Orientierung innerhalb der Fördermaßnahme. Zudem entlasten auch die Direkte Instruktion und die wiederkehrende Struktur der Arbeitsblätter das Arbeitsgedächtnis. Durch eine Orientierung an evidenzbasierten Methoden gerade für lern oder rechenschwache Schülerinnen und Schüler wird angenommen, dass die Schülerinnen und Schüler mit geringen kognitiven Fähigkeiten von der Fördermaßnahme profitieren.

Hypothese 3: Die Schülerinnen und Schüler mit geringen kognitiven Fähigkeiten, die an der vierwöchigen Maßnahme zur Förderung der mathematischen Basiskompetenzen teilnehmen, zeigen im Prä-Post-Vergleich (kurzfristig) und im Zeitraum zwischen der Post-, Follow-Up-I- und Follow-Up-II Erhebung (langfristig) eine bessere Entwicklung in den mathematischen Basiskompetenzen als die Schülerinnen und Schüler mit geringen kognitiven Fähigkeiten in der Kontrollgruppe.

Prä-Post-Vergleich:

Mit $g \in \{EG_{gkF}, KG_{gkF}\}$ und $t \in \{t_1, t_2\}$ gilt

$\boldsymbol{H_0^{3a}}$: $\forall g, t: \mu_{g,t} = \mu_g + \mu_t - \mu$
$\vee\ \mu_{EG_{gkF},t_2} - \mu_{EG_{gkF},t_1} \not> \mu_{KG_{gkF},t_2} - \mu_{KG_{gkF},t_1}$

$\boldsymbol{H_1^{3a}}$: $\exists g, t: \mu_{g,t} \neq \mu_g + \mu_t - \mu$
$\wedge\ \mu_{EG_{gkF},t_2} - \mu_{EG_{gkF},t_1} > \mu_{KG_{gkF},t_2} - \mu_{KG_{gkF},t_1}$

Post-Follow-Up-I-Follow-Up-II-Vergleich:

Mit $g \in \{EG_{gkF}, KG_{gkF}\}$ und $t \in \{t_2, t_3, t_4\}$ gilt

$\boldsymbol{H_0^{3b}}$: $\forall g, t: \mu_{g,t} = \mu_g + \mu_t - \mu$
$\vee\ \mu_{EG_{gkF},t_3} - \mu_{EG_{gkF},t_2} \not> \mu_{KG_{gkF},t_3} - \mu_{KG_{gkF},t_2}$
$\vee\ \mu_{EG_{gkF},t_4} - \mu_{EG_{gkF},t_3} \not> \mu_{KG_{gkF},t_4} - \mu_{KG_{gkF},t_3}$
$\wedge$
$(\exists t: \mu_{EG_{gkF},t} \not> \mu_{KG_{gkF},t}$
$\vee\ H_0^{3a}$

$\boldsymbol{H_1^{3b}}$: $(\exists g, t: \mu_{g,t} = \mu_g + \mu_t - \mu$
$\wedge\ \mu_{EG_{gkF},t_3} - \mu_{EG_{gkF},t_2} > \mu_{KG_{gkF},t_3} - \mu_{KG_{gkF},t_2}$
$\wedge\ \mu_{EG_{gkF},t_4} - \mu_{EG_{gkF},t_3} > \mu_{KG_{gkF},t_4} - \mu_{KG_{gkF},t_3})$
$\vee$
$(\forall t: \mu_{EG_{gkF},t} > \mu_{KG_{gkF},t}$
$\wedge\ H_1^{3a})$

Anmerkung: EG = Experimentalgruppe, KG = Kontrollgruppe, t = Messzeitpunkt, gkF = geringe kognitive Fähigkeiten

7.2.2.3 Geringe Lesekompetenz

Schülerinnen und Schüler mit einer geringen Lesekompetenz zeichnen sich oftmals dadurch aus, dass ihre Lesegeschwindigkeit verlangsamt ist und der Inhalt des Gelesenen eher schwer verstanden wird. Zudem ist die Zuordnung der Buchstaben zu den Lauten fehlerhaft; bei schwierigen Wörter werden vermeintlich richtige, oftmals verwandte Wörter, gelesen (Schulte-Körne, 2010). Schülerinnen und Schüler mit einer geringen Lesekompetenz zeigen oftmals, ebenso wie Schülerinnen und Schüler mit Rechenschwierigkeiten, Probleme im Bereich der grundlegenden Rechenoperationen (Jordan et al., 2003). Schuchardt und Mähler (2010) begründen ähnliche – und auch nicht ähnliche – Problemstellen durch begrenzte Kapazitäten und strukturelle Unterschiede im Arbeitsgedächtnis, insbesondere in der phonologischen Schleife und der zentralen Exekutive.

Die strategieorientierte, arbeitsgedächtnisentlastende Förderung mathematischer Basiskompetenzen bei Schülerinnen und Schüler mit geringen Lesekompetenzen kann zu einer Verbesserung ihrer mathematischen Basiskompetenzen führen. Auch der weitgehende Verzicht auf Schriftsprache innerhalb der Fördermaßnahme ist für Schülerinnen und Schüler mit einer geringen Lesekompetenz unterstützend und entspricht der aktuellen Forderung, sich in mathematischer Förderung „barrierefrei" auf mathematische – und nicht schriftsprachliche – Inhalte zu konzentrieren (Landerl & Fussenegger, 2006).

Hypothese 4: Die Schülerinnen und Schüler mit einer geringen Lesekompetenz, die an der vierwöchigen Maßnahme zur Förderung der mathematischen Basiskompetenzen teilnehmen, zeigen im Prä-Post-Vergleich (kurzfristig) und im Zeitraum zwischen der Post-, Follow-Up-I- und Follow-Up-II-Erhebung (langfristig) eine bessere Entwicklung in den mathematischen Basiskompetenzen als die Schülerinnen und Schüler mit einer geringen Lesekompetenz in der Kontrollgruppe.

Prä-Post-Vergleich:

Mit $g \in \{EG_{gLk}, KG_{gLk}\}$ und $t \in \{t_1, t_2\}$ gilt

$\boldsymbol{H_0^{4a}}$: $\forall g,t\colon \mu_{g,t} = \mu_g + \mu_t - \mu$
$\vee\ \mu_{EG_{gLk},t_2} - \mu_{EG_{gLk},t_1} \not> \mu_{KG_{gLk},t_2} - \mu_{KG_{gLk},t_1}$

$\boldsymbol{H_1^{4a}}$: $\exists g,t\colon \mu_{g,t} \neq \mu_g + \mu_t - \mu$
$\wedge\ \mu_{EG_{gLk},t_2} - \mu_{EG_{gLk},t_1} > \mu_{KG_{gLk},t_2} - \mu_{KG_{gLk},t_1}$

Post-Follow-Up-I-Follow-Up-II-Vergleich:

Mit $g \in \{EG_{gLk}, KG_{gLk}\}$ und $t \in \{t_2, t_3, t_4\}$ gilt

$\boldsymbol{H_0^{4b}}$: $\forall g,t\colon \mu_{g,t} = \mu_g + \mu_t - \mu$
$\vee\ \mu_{EG_{gLk},t_3} - \mu_{EG_{gLk},t_2} \not> \mu_{KG_{gLk},t_3} - \mu_{KG_{gLk},t_2}$
$\vee\ \mu_{EG_{gLk},t_4} - \mu_{EG_{gLk},t_3} \not> \mu_{KG_{gLk},t_4} - \mu_{KG_{gLk},t_3}$
$\wedge$
$(\exists t\colon \mu_{EG_{gLk},t} \not> \mu_{KG_{gLk},t}$
$\vee\ H_0^{4a})$

$\boldsymbol{H_1^{4b}}$: $(\exists g,t\colon \mu_{g,t} = \mu_g + \mu_t - \mu$
$\wedge\ \mu_{EG_{gLk},t_3} - \mu_{EG_{gLk},t_2} > \mu_{KG_{gLk},t_3} - \mu_{KG_{gLk},t_2}$
$\wedge\ \mu_{EG_{gLk},t_4} - \mu_{EG_{gLk},t_3} > \mu_{KG_{gLk},t_4} - \mu_{KG_{gLk},t_3})$
$\vee$
$(\forall t\colon \mu_{EG_{gLk},t} > \mu_{KG_{gLk},t}$
$\wedge\ H_1^{4a})$

Anmerkung: EG = Experimentalgruppe, KG = Kontrollgruppe, t = Messzeitpunkt, gLk = geringe Lesekompetenz

7.2.2.4 Risikobehaftetes Sozialverhalten

Erfolgreiches (mathematisches) Lernen von Schülerinnen und Schülern steht in einem engen Zusammenhang mit ihrem Sozialverhalten (Neukäter & David, 1996; Rolus-Borgward, 1997; Schröder & Wittrock, 2002). Die im Rahmen der Fördermaßnahme eingesetzten Methoden des peergestützten und kooperativen Lernens sollen gerade Schülerinnen und Schülern mit einem schwierigen Sozialverhalten in ihrem mathematischen Kompetenzerwerb unterstützen (Haag & Streber, 2011; Souvignier, 2007). Das Hineinschlüpfen in eine für die Problemstellung relevante „Rolle", das authentische Geben – und Erhalten – von Feedback, das sich selbst oder dem Gegenüber beim Lernen unterstützen soll, und auch das gemeinsame Lösen von Aufgaben im Team oder in einer Gruppe, kann gerade bei Schülerinnen und Schülern mit einem eher unzureichenden Sozialverhalten lernförderlich sein. Somit kann angenommen werden, dass eben diese Schülerinnen und Schüler mit einem risikobehafteten Sozialverhalten von der Fördermaßnahme profitieren.

Hypothese 5: Die Schülerinnen und Schüler mit einem risikobehafteten Sozialverhalten, die an der vierwöchigen Maßnahme zur Förderung der mathematischen Basiskompetenzen teilnehmen, zeigen im Prä-Post-Vergleich (kurzfristig) und im Zeitraum zwischen der Post-, Follow-Up-I- und Follow-Up-II-Erhebung (langfristig) eine bessere Entwicklung in den mathematischen Basiskompetenzen als die Schülerinnen und Schüler mit einem risikobehafteten Sozialverhalten in der Kontrollgruppe.

Prä-Post-Vergleich:

Mit $g \in \{EG_{rS}, KG_{rS}\}$ und $t \in \{t_1, t_2\}$ gilt

$\boldsymbol{H_0^{5a}}$**:** $\forall g,t\colon \mu_{g,t} = \mu_g + \mu_t - \mu$
$\vee\ \mu_{EG_{rS},t_2} - \mu_{EG_{rS},t_1} \not> \mu_{KG_{rS},t_2} - \mu_{KG_{rS},t_1}$

$\boldsymbol{H_1^{5a}}$: $\exists g,t\colon \mu_{g,t} \neq \mu_g + \mu_t - \mu$
$\wedge\ \mu_{EG_{rS},t_2} - \mu_{EG_{rS},t_1} > \mu_{KG_{rS},t_2} - \mu_{KG_{rS},t_1}$

Post-Follow-Up-I-Follow-Up-II-Vergleich:

Mit $g \in \{EG_{rS}, KG_{rS}\}$ und $t \in \{t_2, t_3, t_4\}$ gilt

$\boldsymbol{H_0^{5b}}$: $\forall g,t\colon \mu_{g,t} = \mu_g + \mu_t - \mu$
$\vee\ \mu_{EG_{rS},t_3} - \mu_{EG_{rS},t_2} \not> \mu_{KG_{rS},t_3} - \mu_{KG_{rS},t_2}$
$\vee\ \mu_{EG_{rS},t_4} - \mu_{EG_{rS},t_3} \not> \mu_{KG_{rS},t_4} - \mu_{KG_{rS},t_3}$
$\wedge$
$(\exists t\colon \mu_{EG_{rS},t} \not> \mu_{KG_{rS},t}$
$\vee\ H_0^{5a})$

$\boldsymbol{H_1^{5b}}$: $(\exists g,t\colon \mu_{g,t} = \mu_g + \mu_t - \mu$
$\wedge\ \mu_{EG_{rS},t_3} - \mu_{EG_{rS},t_2} > \mu_{KG_{rS},t_3} - \mu_{KG_{rS},t_2}$
$\wedge\ \mu_{EG_{rS},t_4} - \mu_{EG_{rS},t_3} > \mu_{KG_{rS},t_4} - \mu_{KG_{rS},t_3})$
$\vee$
$(\forall t\colon \mu_{EG_{rS},t} > \mu_{KG_{rS},t}$
$\wedge\ H_1^{5a})$

Anmerkung: EG = Experimentalgruppe, KG = Kontrollgruppe, t = Messzeitpunkt, rS = risikobehaftetes Sozialverhalten

7.2.2.5 Risikobehaftetes Lernverhalten

Neben dem Sozialverhalten spielt auch das Lernverhalten für die akademische Leistungsentwicklung eine bedeutsame Rolle (Haag & Streber, 2011). Gerade diejenigen Schülerinnen und Schüler, die keine gute Handlungsplanung haben, um Aufgaben zu bewältigen, benötigen Strategien und klare Strukturen, die sie beim Lernen unterstützen (Matthes, 2006). Im Kontext der Maßnahme wird eben dieser Aspekt insbesondere durch den Advance Organizer (Wahl, 2005) (siehe Abbildung 39) und durch die Strategieinstruktion durch die Direkte Instruktion (Lebens & Lauth, 2014) berücksichtigt. Um die Handlungsplanung auch in Einzelarbeitsphasen zu unterstützen, sind die Arbeitsblätter stets ähnlich gestaltet und sind mit Beispielen versehen, mit denen die Schülerinnen und Schüler die angeforderten Rechenstrategien noch einmal nachvollziehen können. Eine hohe Lernaktivität, in der nur wenig Passivität zulässig ist, erfordern zudem die peergestützten und kooperativen Arbeitsphasen, in denen auch die Schülerinnen und Schüler mit einem unzureichenden Lernverhalten einen beständigen Part einnehmen (Büttner et al., 2012). Somit kann angenommen werden, dass sich eben diese Schülerinnen und Schüler mit einem risikobehafteten Lernverhalten der Experimentalgruppe in ihren mathematischen Kompetenzen stärker entwickeln als diejenigen mit einem risikobehafteten Lernverhalten in der Kontrollgruppe.

Hypothese 6: Die Schülerinnen und Schüler mit einem risikobehafteten Lernverhalten, die an der vierwöchigen Maßnahme zur Förderung der mathematischen Basiskompetenzen teilnehmen, zeigen im Prä-Post-Vergleich (kurzfristig) und im Zeitraum zwischen der Post-, Follow-Up-I- und Follow-Up-II-Erhebung (langfristig) eine bessere Entwicklung in den mathematischen Basiskompetenzen als die Schülerinnen und Schüler der Kontrollgruppe mit einem risikobehafteten Lernverhalten.

Prä-Post-Vergleich:

Mit $g \in \{EG_{rL}, KG_{rL}\}$ und $t \in \{t_1, t_2\}$ gilt

H_0^{6a}: $\forall g, t: \mu_{g,t} = \mu_g + \mu_t - \mu$
$\vee\ \mu_{EG_{rL},t_2} - \mu_{EG_{rL},t_1} \not> \mu_{KG_{rL},t_2} - \mu_{KG_{rL},t_1}$

H_1^{6a}: $\exists g, t: \mu_{g,t} \neq \mu_g + \mu_t - \mu$
$\wedge\ \mu_{EG_{rL},t_2} - \mu_{EG_{rL},t_1} > \mu_{KG_{rL},t_2} - \mu_{KG_{rL},t_1}$

Post-Follow-Up-I-Follow-Up-II-Vergleich:

Mit $g \in \{EG_{rL}, KG_{rL}\}$ und $t \in \{t_2, t_3, t_4\}$ gilt

H_0^{6b}: $\forall g, t: \mu_{g,t} = \mu_g + \mu_t - \mu$
$\vee\ \mu_{EG_{rL},t_3} - \mu_{EG_{rL},t_2} \not> \mu_{KG_{rL},t_3} - \mu_{KG_{rL},t_2}$
$\vee\ \mu_{EG_{rL},t_4} - \mu_{EG_{rL},t_3} \not> \mu_{KG_{rL},t_4} - \mu_{KG_{rL},t_3}$
$\wedge$
$(\exists t: \mu_{EG_{rL},t} \not> \mu_{KG_{rL},t}$
$\vee\ H_0^{6a})$

H_1^{6b}: $(\exists g, t: \mu_{g,t} = \mu_g + \mu_t - \mu$
$\wedge\ \mu_{EG_{rL},t_3} - \mu_{EG_{rL},t_2} > \mu_{KG_{rL},t_3} - \mu_{KG_{rL},t_2}$
$\wedge\ \mu_{EG_{rL},t_4} - \mu_{EG_{rL},t_3} > \mu_{KG_{rL},t_4} - \mu_{KG_{rL},t_3})$
$\vee$
$(\forall t: \mu_{EG_{rL},t} > \mu_{KG_{rL},t}$
$\wedge\ H_1^{6a})$

Anmerkung: EG = Experimentalgruppe, KG = Kontrollgruppe, t = Messzeitpunkt, rL = risikobehaftetes Lernverhalten

7.2.2.6 Sonderpädagogischer Unterstützungsbedarf im Bereich Lernen

Schülerinnen und Schüler mit einem Unterstützungsbedarf im Bereich Lernen zeigen im Regelfall ein erhebliches Schulleistungsversagen, das sich aufgrund begrenzter Arbeitsgedächtniskapazitäten in Kombination mit unzureichenden Lernstrategien oftmals in der Nichtbeherrschung der Kulturtechniken niederschlägt. Der Begriff „Unterstützungsbedarf im Bereich Lernen" impliziert eine „schulorganisatorische Setzung" (Linderkamp & Grünke, 2007, S. 18), in der ein Leistungsrückstand von mehr als zwei Schuljahren angenommen werden kann (Grünke & Grosche, 2014). Auch für diese Schülerinnen und Schüler ist eine Wiederholung des Grundschulstoffs verbunden mit einer arbeitsgedächtnisentlastenden strukturierten Vorwissensaktivierung durch den Advance Organizer und einer Strategieinstruktion durch die Direkte Instruktion im Erwerb mathematischer Kompetenzen zum Zeitpunkt des Übergangs eine lernförderliche Unterstützung.

Hypothese 7: Die Schülerinnen und Schüler mit einem sonderpädagogischen Unterstützungsbedarf im Bereich Lernen, die an der vierwöchigen Maßnahme zur Förderung der mathematischen Basiskompetenzen teilnehmen, zeigen im Prä-Post-Vergleich (kurzfristig) und im Zeitraum zwischen der Post-, Follow-Up-I- und Follow-Up-II-Erhebung (langfristig) eine bessere Entwicklung in den mathematischen Basiskompetenzen als die Schülerinnen und Schüler der Kontrollgruppe mit einem sonderpädagogischen Unterstützungsbedarf im Bereich Lernen.

Prä-Post-Vergleich:

Mit $g \in \{EG_{SPU}, KG_{SPU}\}$ und $t \in \{t_1, t_2\}$ gilt

$\boldsymbol{H_0^{7a}}$: $\forall g,t: \mu_{g,t} = \mu_g + \mu_t - \mu$
$\vee\ \mu_{EG_{SPU},t_2} - \mu_{EG_{SPU},t_1} \not> \mu_{KG_{SPU},t_2} - \mu_{KG_{SPU},t_1}$

$\boldsymbol{H_1^{7a}}$: $\exists g,t: \mu_{g,t} \neq \mu_g + \mu_t - \mu$
$\wedge\ \mu_{EG_{SPU},t_2} - \mu_{EG_{SPU},t_1} > \mu_{KG_{SPU},t_2} - \mu_{KG_{SPU},t_1}$

Post-Follow-Up-I-Follow-Up-II-Vergleich:

Mit $g \in \{EG_{SPU}, KG_{SPU}\}$ und $t \in \{t_2, t_3, t_4\}$ gilt

$\boldsymbol{H_0^{7b}}$: $\forall g,t: \mu_{g,t} = \mu_g + \mu_t - \mu$
$\vee\ \mu_{EG_{SPU},t_3} - \mu_{EG_{SPU},t_2} \not> \mu_{KG_{SPU},t_3} - \mu_{KG_{SPU},t_2}$
$\vee\ \mu_{EG_{SPU},t_4} - \mu_{EG_{SPU},t_3} \not> \mu_{KG_{SPU},t_4} - \mu_{KG_{SPU},t_3}$
$\wedge$
$(\exists t: \mu_{EG_{SPU},t} \not> \mu_{KG_{SPU},t}$
$\vee\ H_0^{7a})$

$\boldsymbol{H_1^{7b}}$: $(\exists g,t: \mu_{g,t} = \mu_g + \mu_t - \mu$
$\wedge\ \mu_{EG_{SPU},t_3} - \mu_{EG_{SPU},t_2} > \mu_{KG_{SPU},t_3} - \mu_{KG_{SPU},t_2}$
$\wedge\ \mu_{EG_{SPU},t_4} - \mu_{EG_{SPU},t_3} > \mu_{KG_{SPU},t_4} - \mu_{KG_{SPU},t_3})$
$\vee$
$(\forall t: \mu_{EG_{SPU},t} > \mu_{KG_{SPU},t}$
$\wedge\ H_1^{7a})$

Anmerkung: EG = Experimentalgruppe, KG = Kontrollgruppe, t = Messzeitpunkt, SPU = Sonderpädagogischer Unterstützungsbedarf im Bereich Lernen

7.2.3 Hypothesenblock III: Einfluss der Lernfreude und der Leistungsheterogenität der eigenen Klasse auf die Wirksamkeit

Im folgenden Hypothesenblock III wird die Lernfreude und die Leistungsheterogenität der eigenen Klasse als vorhersagende Variablen der kurzfristigen Wirksamkeit (Punktezuwachs im HRT von der Prä- zur Posterhebung) herangezogen und begründet.

Motivation und Lernfreude haben einen großen Einfluss auf den Lernerfolg von Schülerinnen und Schülern (Matthes, 2006). „Kinder mit [einer hohen] Lernfreude haben [...] in aller Regel eine positivere Schuleinstellung, sie sind eher anstrengungsbereit und Neuem gegenüber aufgeschlossen“ (Rauer & Schuck, 2003, S. 15). Je älter die Schülerinnen und Schüler sind, umso höher sind die Zusammenhänge der Lernfreude mit der schulischen Leistung (Rauer & Schuck, 2003, S. 15). Bedeutsam ist in der Frage der Lernfreude, dass sich die Lerninhalte und Aufgaben der Fördermaßnahme an die Lebenswelt der Schülerinnen und Schüler anschließen, da die Freude am Lernen sonst weichen kann (Matthes, 2006). Die Fördermaßnahme versucht durch die inhaltliche Verknüpfung mit einem Sportereignis einen Bezug zur Lebenswelt der Schülerinnen und Schüler aufzubauen und die zu erlernenden Inhalte durch wechselnde Methoden attraktiv zu gestalten.

Daneben zeigen zahlreiche Studien wie zum Beispiel das Dortmunder Projekt Grundschulübergang (Van Ophuysen & Wendt, 2009), die KESS-Studie (Bos & Gröhlich, 2010) oder die LAU-Studie (Lehmann, Husfeldt, et al., 2001), dass der mathematische Kompetenzerwerb in der Sekundarstufe I stark von dem Leistungsniveau der eigenen Klasse abhängt. Eine leistungsstarke Klasse und gar eine leistungsstarke Schule kann sich positiv auf die Leistungsentwicklung der einzelnen Schülerinnen und Schüler auswirken (sog. Kompositionseffekt) (Opdenakker & Van Damme, 2007). Das innere Leistungsspektrum unterscheidet sich dabei von einer Lerngruppe zur nächsten.

Es stellt sich die Frage, ob die Lernfreude der Schülerinnen und Schüler und das Leistungsspektrum der Klasse als Prädiktoren für die kurzfristige Wirksamkeit der Fördermaßnahme (Prä-Post) identifiziert werden können. Die Analyse der Vorhersagekraft verspricht eine differenzierte Aufklärung der Wirksamkeit der Maßnahme.

Hypothese 8: Unterschiede in der Lernfreude der Schülerinnen und Schüler sowie Unterschiede im mathematischen Leistungsspektrum der einbezogenen Klassen haben eine signifikante Vorhersagekraft für das Ausmaß der kurzfristigen Wirksamkeit der Fördermaßnahme.

$H_0^{8a}: P \neq b * W + a$

$H_1^{8a}: P = b * W + a$

$H_0^{8b}: P \neq b * P_{LS} + a$

$H_1^{8b}: P = b * P_{LS} + a$

Anmerkungen: P = Leistungszuwachs von Prä zu Post, W = Lernfreude, P_{ls} = Leistungsspektrum in der Klasse

7.2.4 Hypothesenblock IV: Implementationsqualität der Maßnahme

Der Implementationsqualität von Fördermaßnahmen und -konzepten wird in der Interventions- und Evaluationsforschung eine bedeutsame Rolle zugeschrieben. Gräsel und Parchmann (2004) fordern, dass die Erhebung der Implementationsqualität ein verpflichtender Bestandteil von Evaluationsforschung wird, um die Praxistauglichkeit von Konzepten und Maßnahmen zu prüfen. „Erst durch das Vorhandensein von Informationen über das Ausmaß der Implementation wird es möglich, Aussagen über die Wirksamkeit und den Erfolg einer Maßnahme zu treffen" (Petermann, 2015b, S. 122). Der Hypothesenblock IV widmet sich dem Thema der Implementation*squalität* und beinhaltet dabei sowohl die Perspektive der Schülerinnen und Schüler sowie die der Interventionsleiterinnen und -leiter.

7.2.4.1 Beurteilung der Schülerinnen und Schüler

Aufgrund der didaktischen Überlegungen, der zielgruppenspezifischen Methodik und den positiven Resultaten der Pilotstudie (siehe Tabelle 28) wird davon ausgegangen, dass die Schülerinnen und Schüler die Implementationsvariablen „Akzeptanz" und „Angemessenheit" nach Petermann (2015b) grundsätzlich positiv beurteilen (vgl. das Vorgehen bei Schell, 2011).

Hypothese 9: Für die Schülerinnen und Schüler der Experimentalgruppe ist die Fördermaßnahme zufriedenstellend (acceptability). Sie stimmen den Items zur *Akzeptanz* der Maßnahme durchschnittlich „ziemlich" zu (entspricht dem Skalenwert 2 von 4).

$H_0^9: \mu_{AK} > 2$ $\qquad$ $H_1^9: \mu_{AK} \leq 2$

Hypothese 10: Die Schülerinnen und Schüler der Experimentalgruppe empfinden die Maßnahme als für sie passend (appropriateness). Sie stimmen den Items zur *Angemessenheit* der Maßnahme durchschnittlich „ziemlich" zu (entspricht dem Skalenwert 2 von 4).

$H_0^{10}: \mu_{AN} > 2$ $\qquad$ $H_1^{10}: \mu_{AN} \leq 2$

7.2.4.2 Beurteilung der Interventionsleiterinnen und -leiter

Die durchführenden Interventionsleiterinnen und -leiter erhalten durch die Projektleitung eine umfassende Einführung in die Diagnostik der Erhebungsinstrumente und in die Umsetzung der Maßnahme. Zudem werden sie zu Kriterien guten Classroom-Managements geschult. Auch regelmäßige Supervisionssitzungen haben zum Ziel, die erfolgreiche Durchführung der Maßnahme zu unterstützen. Es wird insgesamt davon ausgegangen, dass die Interventionsleiterinnen und -leiter die Implementationsvariablen „Akzeptanz", „Übernahme", „Angemessenheit", „Machbarkeit" und „Wiedergabetreue" nach Petermann (2015b) grundsätzlich positiv beurteilen (vgl. das Vorgehen bei Schell, 2011).

Hypothese 11: Die Interventionsleiterinnen und -leiter sind zufrieden mit der Maßnahme (acceptability). Sie stimmen den Items zur *Akzeptanz* der Maßnahme durchschnittlich „ziemlich" zu (entspricht dem Skalenwert 2 von 4).

$H_0^{11}: \mu_{AK} > 2$ $\quad$ $H_1^{11}: \mu_{AK} \leq 2$

Hypothese 12: Die Interventionsleiterinnen -leiter stehen einer Implementation der Fördermaßnahme in die alltägliche schulische Praxis positiv gegenüber (adoption). Sie stimmen den Items zur *Übernahme* der Maßnahme durchschnittlich „ziemlich" zu (entspricht dem Skalenwert 2 von 4).

$H_0^{12}: \mu_{Ü} > 2$ $\quad$ $H_1^{12}: \mu_{Ü} \leq 2$

Hypothese 13: Die Interventionsleiterinnen -leiter empfinden die Maßnahme als passend und als mit dem schulischen Alltag kompatibel (appropriateness). Sie stimmen den Items zur *Angemessenheit* der Maßnahme durchschnittlich „ziemlich" zu (entspricht dem Skalenwert 2 von 4).

$H_0^{13}: \mu_{AN} > 2$ $\quad$ $H_1^{13}: \mu_{AN} \leq 2$

Hypothese 14: Die Interventionsleiterinnen -leiter empfinden die Maßnahme als im geplanten schulischen Setting umsetzbar (feasibility). Sie stimmen den Aussagen zur *Machbarkeit* der Maßnahme durchschnittlich „ziemlich" zu (entspricht dem Skalenwert 2 von 4).

$H_0^{14}: \mu_{M} > 2$ $\quad$ $H_1^{14}: \mu_{M} \leq 2$

Hypothese 15: Die Interventionsleiterinnen -leiter setzen die Maßnahme entsprechend des Manuals um (fidelity). Sie stimmen den Aussagen zur *Wiedergabetreue* der Maßnahme durchschnittlich „ziemlich" zu (entspricht dem Skalenwert 2 von 4).

$H_0^{15}: \mu_{W} > 2$ $\quad$ $H_1^{15}: \mu_{W} \leq 2$

8 Methodik und Evaluation

In Fragen der wissenschaftlichen Verortung der Konzeption und Evaluation von Interventionen und Förderprogrammen spielt sowohl die Interventions- als auch die Evaluationsforschung eine Rolle. Die Interventionsforschung

> „[widmet sich] der Entwicklung und Überprüfung von Theorien, die auf Techniken der zielgerichteten Veränderung von Sachverhalten abzielen. [...] Die Evaluationsforschung schließlich befasst sich mit der wissenschaftlich fundierten Bewertung von Sachverhalten und insbesondere von Interventionsmaßnahmen hinsichtlich verschiedener Bewertungskriterien, etwa Effektivität, Effizienz, Akzeptanz oder Nachhaltigkeit“ (Döring & Bortz, 2016a, S. 977).

Diese eher eng definierte Unterscheidung ist jedoch nicht unumstritten, da „die Entwicklung, Implementation und Evaluation pädagogischer Programme oft in einer Hand liegen“ (Mittag & Bieg, 2010, S. 31). Mittag und Bieg (2010, S. 35) bevorzugen den eher weiten Begriff „pädagogische Interventionsforschung“ und fassen darunter drei Schritte:

1. Evaluation der Programmkonzeption vor der Durchführung,
2. Evaluation der Programmdurchführung während der Durchführung,
3. Evaluation der Programmwirksamkeit nach der Durchführung.

Die Evaluation der Programmkonzeption vor der Durchführung (1.) betrachtet zum Beispiel die grundlegende Konzeption und Konstruktion des Förderkonzepts. Während der Durchführung des Programms erfolgt dann eine begleitende, formative Evaluation (2.) und nach der Durchführung eine summative Evaluation des Förderkonzepts (3.).

Der methodische Zugang im Rahmen der vorliegenden Forschungsarbeit ist nach Döring und Bortz (2016a, S. 977) eher dem Bereich der „Evaluationsforschung“ zuzuordnen und umfasst die „Evaluation der Programmwirksamkeit nach der Durchführung“ (Mittag & Bieg, 2010, S. 35). Die formative Evaluation der vorliegenden Programmkonzeption und der -durchführung erfolgt im Rahmen einer vorangestellten Pilotstudie (siehe Kapitel 6.2). Die Ergebnisse der Pilotuntersuchung sind im Rahmen einer akademischen Abschlussarbeit positiv evaluiert (siehe F. Schmidt, 2014).

Die im Folgenden dargestellten Aspekte der Forschungsmethodik sind also im Kontext der summativen Evaluation der Wirksamkeit der entwickelten Fördermaßnahme zu verorten.

8.1 Forschungsdesign

Das Ziel der vierwöchigen Maßnahme ist es, die mathematischen Basiskompetenzen von Schülerinnen und Schülern in der Jahrgangsstufe fünf am Anfang der Sekundarstufe I zu fördern. Die Umsetzung der Maßnahme erfolgt im schulischen Setting – „in einer für die Versuchspersonen gewohnten Umgebung [...], da die Versuchspersonen in ihrer Alltagsumgebung verbleiben“ (Schnell, Hill & Esser, 2011, S. 218). Es handelt sich bei der Umsetzung um ein Feldexperiment. Im Rahmen von Feldexperimenten müssen oftmals Einbußen an interner Validität gemacht werden, da es nur schwer bis kaum möglich ist, alle möglichen Störfaktoren zu kontrollieren (Schnell et al., 2011, S. 218).

Die Umsetzung der Maßnahme im Rahmen der Studie erfolgt mittels eines Zwei-Gruppen-Designs. Dabei durchläuft eine „Experimentalgruppe“ zu Beginn des fünften Schuljahres die entwickelte vierwöchige Fördermaßnahme anstelle des regulären Mathematikunterrichts. Eine „Kontrollgruppe“, die „*in jeder Hinsicht äquivalent zur Experimentalgruppe* ist“ (Rost, 2013, S. 140; Herv. i. Orig.) dient dem Vergleich der Leistungsentwicklung und genießt im selben Zeitraum ausschließlich den regulären Mathematikunterricht. Dieses Design ermöglicht es, bei einer Messwerterhöhung in der Experimentalgruppe im Anschluss an die Förderung auf eine potenzielle Wirksamkeit der Maßnahme zu schließen (Schnell et al., 2011, S. 204). Aufgrund einer vorhandenen Kontrollgruppe handelt es sich um ein starkes Design in Form eines randomisierten Zwei-Gruppen-Plans mit Vortest, Behandlung und drei Nachtests (Rost, 2013, S. 141). Nach Bortz und Döring (2016b, S. 748) ist das beschriebene Design auch als Cross-Sequential-Design (Quersequenzmethode) zu bezeichnen.

Es werden im Vorfeld der Förderung (Prätest; t_1), vier Wochen später im direkten Anschluss (Posttest; t_2), nach vier Monaten zum Schulhalbjahr (Follow-Up-I; t_3) und nach zehn Monaten zum Schuljahresende (Follow-Up-II; t_4) akademische und soziale Kompetenzen der Schülerinnen und Schüler beider Untersuchungsgruppen erhoben. Abbildung 44 gibt einen groben schematischen Überblick über das Forschungsdesign.

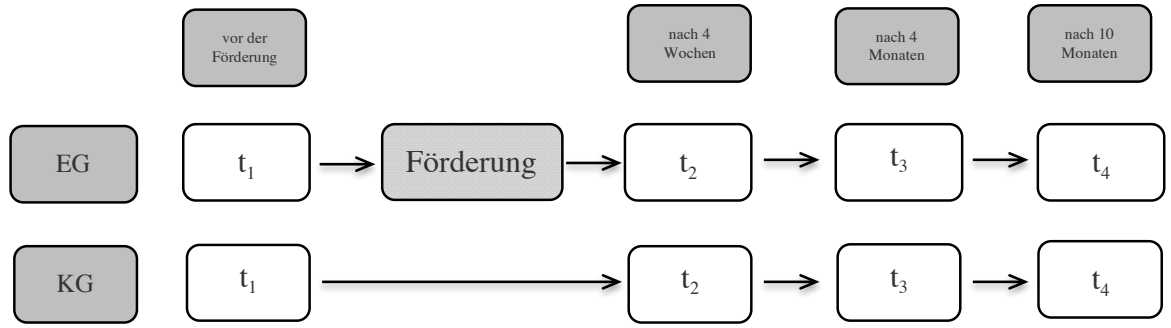


ABBILDUNG 44 SCHEMA DES FORSCHUNGSDESIGNS

In Tabelle 33 sind die Messvariablen, die „veränderliche[n] Beobachtungsgrößen“ (Hussy, Schreier & Echterhoff, 2010, S. 37) des Forschungsdesign im Detail aufgeführt (siehe Kapitel 8.7). Der erste tiefgestellte Buchstabe im Index steht hier für die Gruppe (E = Experimentalgruppe; K = Kontrollgruppe) und die zweite tiefgestellte Zahl für den Messzeitpunkt 1 bis 4. Die Variablen sind in der unten stehenden Legende erläutert. Das X steht stellvertretend für die Intervention, das Treatment (Schnell et al., 2011, S. 203).

TABELLE 33 DETAILLIERTER ÜBERBLICK ÜBER DAS FORSCHUNGSDESIGN

	Prätest	**Intervention**	**Posttest**	**Follow-Up I**	**Follow-Up II**
	Schuljahresbeginn	1./2. – 4./5. Schulwoche	Herbstferien	Schuljahreshalbjahr	Schuljahresende
	September 2014	Sept. & Oktober 2014	Oktober 2014	Januar 2015	Juli 2015
Experimentalgruppe	$\overrightarrow{P_{E,1}}$ $\overrightarrow{Q_{E,1}}$ $\overrightarrow{R_{E,1}}$	X	$\overrightarrow{P_{E,2}}$ $\overrightarrow{S_{E,2}}$ $\overrightarrow{T_{E,2}}$ $\overrightarrow{U_{E,2}}$ $\overrightarrow{V_{E,2}}$ $\overrightarrow{W_{E,2}}$	$\overrightarrow{P_{E,3}}$	$\overrightarrow{P_{E,4}}$
Kontrollgruppe	$\overrightarrow{P_{K,1}}$ $\overrightarrow{Q_{K,1}}$ $\overrightarrow{R_{K,1}}$	–	$\overrightarrow{P_{K,2}}$ $\overrightarrow{S_{K,2}}$ $\overrightarrow{T_{K,2}}$	$\overrightarrow{P_{K,3}}$	$\overrightarrow{P_{K,4}}$

Legende:

$\overrightarrow{P}$ = Mathematische Kompetenzen
$\overrightarrow{Q}$ = Kognitive Fähigkeiten
$\overrightarrow{R}$ = Lesekompetenz
$\overrightarrow{S}$ = Sozialverhalten
$\overrightarrow{T}$ = Lernverhalten
$\overrightarrow{U}$ = Implementationsqualität Schüler/Innen
$\overrightarrow{V}$ = Implementationsqualität Interventionsleiter/Innen
$\overrightarrow{W}$ = Lernfreude

Die hier vorliegende Untersuchung stellt einen Teil eines umfassenden Projekts dar, sodass z. T. noch weitere Variablen erhoben sind, die hier jedoch nicht von Relevanz sind und deswegen nicht aufgeführt werden.

OPERATIONALISIERUNG DER VARIABLEN

Tabelle 34 gibt einen zusammenfassenden Überblick über die Operationalisierung der Hypothesen mit den entsprechend relevanten Variablen aus dem Forschungsdesign. Durch eine Operationalisierung von Hypothesen, werden diese „präzisiert" und damit „empirisch prüfbar" (Sedlmeier & Renkewitz, 2013, S. 43). Die in Tabelle 34 genannten Erhebungsinstrumente werden im folgenden Kapitel 8.2 im Detail vorgestellt.

TABELLE 34 OPERATIONALISIERUNG DER VARIABLEN DER HYPOTHESEN

H*	**Variable**		**Operationalisierung**
H_1 – H_8	$\vec{P}$	Mathematische Kompetenz	Outcome in Form des summierten Rohwerts in der Subskala Rechenoperationen im Heidelberger Rechentest 1–4 (Haffner et al., 2005) mit den Subtests Addition, Subtraktion, Multiplikation, Division, Ergänzungsaufgaben und Größer-Kleiner-Vergleich.
H_3	$\vec{Q}$	Kognitive Fähigkeiten	Intelligenzquotient (grundlegende geistige Leistungsfähigkeit) ermittelt durch den CFT 20-R (Teil 1) (Weiß, 2006) mit den Subtests Reihenfortsetzen, Klassifikationen, Matrizen und Topologien
H_4	$\vec{R}$	Lesekompetenz	Basale Lesefertigkeit (Lesegeschwindigkeit) in Form des Lesequotienten im Salzburger Lesescreening 1–4 (Mayringer & Wimmer, 2003)

H*	Variable		Operationalisierung	
H_5	$\vec{S}$	Sozialverhalten	Durchschnittlicher T-Wert in der Subskala Sozialverhalten in der Lehrereinschätzliste für Sozial- und Lernverhalten (Petermann & Petermann, 2013) mit den Subtests Anstrengungsbereitschaft, Ausdauer, Konzentration, Selbstständigkeit beim Lernen und Sorgfalt beim Lernen	
H_6	$\vec{T}$	Lernverhalten	Durchschnittlicher T-Wert in der Subskala Lernverhalten in der Lehrereinschätzliste für Sozial- und Lernverhalten (Petermann & Petermann, 2013) mit den Subtests Kooperation, Selbstwahrnehmung, Selbstkontrolle, Einfühlungsvermögen, Hilfsbereitschaft, angemessene Selbstbehauptung und Sozialkontakt	
H_9 H_{10}	$\vec{U}$	Implementationsqualität aus Sicht der Schülerinnen und Schüler	Mittlere Beurteilung der Skalen im Implementationsfragebogen	
			$\vec{U}_1$	Akzeptanz
			$\vec{U}_2$	Angemessenheit
H_{11} H_{12} H_{13} H_{14} H_{15}	$\vec{V}$	Implementationsqualität aus Sicht der Interventionsleiterinnen und -leiter	Mittlere Beurteilung der Skalen im Implementationsfragebogen	
			$\vec{V}_1$	Akzeptanz
			$\vec{V}_2$	Angemessenheit
			$\vec{V}_3$	Übernahme
			$\vec{V}_4$	Machbarkeit
			$\vec{V}_5$	Wiedergabetreue
H_{8a}	$\vec{W}$	Lernfreude	Rohwert in der Subskala „Lernfreude“ im Fragebogen zur Erhebung emotionaler und sozialer Schulerfahrungen (Rauer & Schuck, 2003)	

Anmerkung: *H = Hypothese; die Variable und die Operationalisierung beziehen sich auf diese Hypothese/n

Zur Ermittlung der kurz- und langfristigen Wirksamkeit werden zu allen vier Messzeitpunkten die mathematischen Basiskompetenzen der Schülerinnen und Schüler durch die Subskala Rechenoperationen des Heidelberger Rechentest 1–4 (Haffner et al., 2005) erhoben ($\vec{P}$). Diese Erhebung wird ebenso zur Bestimmung der differenziellen Wirksamkeit (siehe Tabelle 32), zur Bestimmung des Vorwissens (Hypothese 2) und für die Ermittlung des mathematischen Leistungsspektrums der Klasse (Hypothese 8b) genutzt.

Um die differenzielle Wirksamkeit der Maßnahme für bestimmte Subgruppen ermitteln zu können, werden im Vorfeld der Maßnahme die kognitiven Fähigkeiten ($\vec{Q}$) mittels des Culture Fair Intelligence Tests 20-R (Weiß, 2006) (Hypothese 3) und die Lesekompetenz ($\vec{R}$) mit dem Salzburger Lese-Screening 1–4 (Mayringer & Wimmer, 2003) (Hypothese 4) erhoben. Da erfolgreiche mathematische Entwicklung eng mit dem Sozial- und Lernverhalten der Schülerinnen und Schüler zusammenhängt (Neukäter & David, 1996; Rolus-Borgward, 1997; Schröder & Wittrock, 2002), werden das Sozial- und Lernverhalten durch die Lehrereinschätzliste für das Sozial- und Lernverhalten im Anschluss an die Fördermaßnahme erfragt ($\vec{S}$ & $\vec{T}$) (Hypothesen

5 & 6). Zur Evaluation der Implementationsqualität der Maßnahme (Petermann, 2015b) werden zwei eigens entwickelte Implementationsfragebögen für die Schülerinnen und Schüler ($\vec{U}$) (Hypothesen 9 & 10) und für die Interventionsleiterinnen und -leiter eingesetzt ($\vec{V}$) (Hypothesen 11–15). Die Schülerinnen und Schüler der Experimentalgruppe werden zudem zu ihrer Lernfreude befragt ($\vec{W}$) (Hypothese 8a).

Neben der Lernfreude werden auch Personenmerkmale der Schülerinnen und Schüler erhoben. Dazu gehören zum Beispiel das Geschlecht, das Alter, die Muttersprache und Aspekte des sozioökonomischen Hintergrunds.

8.2 Erhebungsinstrumente

Im Forschungsdesign ist bereits aufgezeigt, zu welchen Zeitpunkten welche Variablen mit den entsprechenden Erhebungsinstrumenten erhoben werden. Im folgenden Kapitel wird ein Überblick über die Erhebungsinstrumente und ihre Güte gegeben. Das Kapitel wird unterteilt in die Instrumente, die von den Schülerinnen und Schülern bearbeitet werden und in diejenigen, die von den Lehrkräften bzw. den Interventionsleiterinnen und -leitern ausgefüllt werden.

8.2.1 Schülererhebung

Es folgt die Vorstellung der Erhebungsinstrumente, die von den Schülerinnen und Schülern selbst bearbeitet werden.

8.2.1.1 Heidelberger Rechentest (HRT 1–4)

Der Heidelberger Rechentest 1–4 (Haffner et al., 2005) ist ein Test zur Erfassung der mathematischen Basiskompetenzen und der numerisch-logischen sowie räumlich-visuellen mathematischen Fähigkeiten von Schülerinnen und Schülern. Er kann sowohl in der Grundschule als auch zu Beginn der Sekundarstufe I eingesetzt werden. Er ist als Gruppenverfahren einsetzbar und verfügt insgesamt über elf Untertests. „Der HRT soll praxisorientiert die Anwendung in Schulklassen erlauben und einen raschen Überblick über das Leistungsprofil einer ganzen Klasse sowie einzelner Schülerinnen und Schüler (Beurteilung einer Rechenschwäche aber auch besonderer mathematischer Begabungen) ermöglichen“ (Haffner et al., 2005, S. 13). Sechs Untertests umfassen die sog. „Rechenoperationen“ (Addition, Subtraktion, Multiplikation, Division, Ergänzungsaufgaben, Größer-Kleiner-Aufgaben), fünf weitere beinhalten die numerisch-logischen und räumlich-visuellen Fähigkeiten (Zahlenfolgen, Längenschätzen, Würfelzählen, Mengenzählen, Zahlenverbinden).

Die folgende Tabelle 35 zeigt, welche der Untertests im Rahmen der Studie eingesetzt werden. Es handelt sich hierbei um die sechs Untertests aus dem Bereich „Rechenoperationen“.

TABELLE 35 ÜBERBLICK ÜBER DIE UNTERTESTS „RECHENOPERATIONEN“ DES HRT 1–4

Untertest	Kürzel	Messinhalt	Anzahl Items	Zeit (Min.)
Addition	RA	Grundrechnen-Plus	40	2
Subtraktion	RS	Grundrechnen-Minus	40	2
Multiplikation	RM	Grundrechnen-Mal	40	2
Division	RD	Grundrechnen-Geteilt	40	2
Ergänzungs-aufgaben	EG	Rechenleistung bei variablen Gleichungsaufgaben	40	2
Größer-Kleiner-Aufgaben	GKV	Größenvergleiche, Über-blicksrechnen, Ungleichungen	40	2

Der HRT 1–4 ist ein Speed-Test, sodass es eine Zeitbegrenzung bei der Bearbeitung der Subtests gibt. Die reine Bearbeitungsdauer der Untertests umfasst 20,5 Minuten. „Dahinter steht die Annahme, dass Kinder mit ineffektiven, umständlichen oder unsicheren Lösungsstrategien weniger Aufgaben pro Zeiteinheit lösen können als geschickte, kompetente und geübte Rechner, d. h., dass die Anzahl gelöster Aufgaben pro Untertest ein guter Indikator/Schätzwert für die [...] mathematische Fähigkeit/Kompetenz darstellt“ (Haffner et al., 2005, S. 13). Es werden sowohl T-Werte als auch Prozentränge zum Vergleich der Ergebnisse mit der Normstichprobe angeboten. Das Instrument entspricht den hohen Ansprüchen wissenschaftlicher Güte, die der Tabelle 13 entnommen werden können. Der HRT 1–4 findet im Rahmen der Studie nach sorgfältiger Prüfung Anwendung.

8.2.1.2 Salzburger Lese-Screening (SLS 1–4)

Das Salzburger Lese-Screening 1–4 (SLS) (Mayringer & Wimmer, 2003) ist ein Lesetest, der halbjährlich ab der Mitte des zweiten Schuljahres bis zum Ende der vierten Klasse normiert ist. Das SLS besteht aus einer Liste mit 72 Sätzen, die entweder falsch oder richtig sind (zum Beispiel „Tee kann man trinken“; „In der Wüste regnet es oft“). Die Schülerinnen und Schüler haben drei Minuten Zeit, die Sätze zu lesen und müssen sie jeweils als richtig oder falsch markieren. Das Verfahren misst die basale Lesefertigkeit in Form der Lesegeschwindigkeit. Der Test kann als Gruppentest eingesetzt werden, sodass auch die Einschätzung einer Klassenleistung möglich ist. Durch die Anzahl korrekt beurteilter Sätze kann anhand der Normtabellen ein Lesequotient (LQ) ermittelt werden und ein Leserückstand somit auch in Klassenjahren bestimmt werden. Die Abstufung wird dabei an die der Intelligenz angelehnt (Mittelwert 100) und findet sich in Tabelle 36.

TABELLE 36 EINORDNUNG DER LQ-WERTE (MAYRINGER & WIMMER, 2003, S. 22)

Lesequotient (LQ)	Leistung
LQ ≤ 69	sehr schwach
70 ≥ LQ < 80	schwach
80 ≥ LQ < 90	unterdurchschnittlich
90 ≥ LQ < 110	durchschnittlich
110 ≥ LQ < 120	überdurchschnittlich
120 ≥ LQ < 130	gut
LQ ≥ 130	sehr gut

Das Salzburger Lese-Screening 1–4 weist für die vierte Klassenstufe eine Paralleltestreliabilität von $r = .91$ auf (2. Klassenstufe $r = .92$, 3. Klassenstufe $r = .90$). Die Kriteriumsvalidität (Kriterium: Lesezeit beim Lautlesen im Salzburger Lese- und Rechtschreibtest) beträgt für die vierte Klassenstufe $r = .76$ (zweite Klassenstufe $r = .81$, dritte Klassenstufe $r = .81$). Auch die Augenscheinvalidität gilt als erfüllt (Mayringer & Wimmer, 2003, S. 14 ff.). Das Salzburger Lese-Screening genügt insgesamt den Ansprüchen an die wissenschaftliche Güte von Testverfahren.

Es liegt auch das Salzburger Lese-Screening 5–8 (Auer, Gruber, Wimmer & Mayringer, 2005) vor, doch ist dieses erst ab dem *Ende* der fünften Klasse normiert, sodass es sich für den Einsatz zu Beginn der Klasse fünf – und somit für den Einsatz in der vorliegenden Untersuchung – nicht eignet. Für den Einsatz unmittelbar zu Beginn der Klasse fünf im Rahmen der vorliegenden Untersuchung erfolgt eine Orientierung an der Normierung des SLS 1–4 für das Ende der vierten Klasse.

8.2.1.3 Culture Fair Intelligence Test (CFT 20-R)

Der Culture Fair Intelligence Test 20-R (Weiß, 2006), der als Grundintelligenztest Skala 2-Revision (CFT 20-R) bekannt ist, kann bei Kindern und Jugendlichen im Alter zwischen 8;5 und 19 Jahren sowie bei Erwachsenen zwischen 20 bis 60 Jahren eingesetzt werden. Der CFT 20-R misst die sog. „Grundintelligenz" – die „grundlegende[...] geistige[...] Leistungsfähigkeit" (Weiß, 2006, S. 11) und gilt als sprach- und kulturfreier Test, sodass Personen mit ungenügend ausgeprägten Kulturtechniken oder einem geringen Sprachvermögen nicht benachteiligt sind. Der Test besteht aus zwei Teilen, die jeweils aus vier inhaltlich gleichen Subtests bestehen und nacheinander durchgeführt werden:

1. Reihen fortsetzen
2. Klassifikationen
3. Matrizen
4. Topologien.

Es geht in den Subtests folglich darum, figurale Beziehungen zu erkennen und durch formallogisches Denken zu verarbeiten. Die Durchführung des ersten Testteils dauert ca. 20 Minuten. Er besteht aus 56 Items und kann in Form einer Gruppentestung durchgeführt werden. Die Durchführung beider Testteile führt zu einer höheren Reliabilität und Validität.

Bezüglich der Gütekriterien gilt die Objektivität durch eine manualisierte Instruktion für die Durchführung als gegeben. Die Reliabilität (im Manual nicht näher bestimmt) für den ersten Testteil liegt bei $r = .92$ und die Korrelation zwischen den beiden Testteilen bei $r = .82$. Bezüglich der externen Validität heißt es: „Resümierend kann [...] gefolgert werden, dass der [...] CFT 20-R eine wesentlich gerechtere Beurteilung der intellektuellen Leistungskapazität von Kindern aus ungünstigerem sozialen, regionalen und schulischen Milieu ermöglicht als konventionelle, insbesondere von verbalen Kenntnissen abhängige Intelligenztests oder schulische Leistungstests“ (Weiß, 2006, S. 17). Auch die Konstruktvalidität und die prognostische Validität sind für das Instrument detailliert nachgewiesen (Weiß, 2006, S. 16). Das Ergebnis des CFT 20-R ist neben einem Prozentrang der Intelligenzquotient (IQ) mit dem Mittelwert 100 und der Standardabweichung von 15.

Laut Autor ist es möglich, nur den ersten Testteil durchzuführen, was im Rahmen der vorliegenden Untersuchung genutzt wird. Zudem kann bei Probanden mit Migrationshintergrund, Sprachdefiziten und besonderen psychischen Auffälligkeiten die Testzeit pro Subtest um eine Minute verlängert werden, wovon hier ebenso Gebrauch gemacht wird.

Folgende Tabelle 37 gibt einen Überblick über IQ-Wertebereiche und die Intelligenzniveaus (nach der ICD-10; vgl. Deutsche Gesellschaft für Kinder- und Jugendpsychiatrie, 2011, S. 34).

TABELLE 37 IQ-WERTEBEREICHE UND INTELLIGENZNIVEAUS

Wertebereich	Intelligenzniveau
IQ < 20	Schwerste Intelligenzminderung
20 ≥ IQ < 35	Schwere Intelligenzminderung
35 ≥ IQ < 50	Mittelgradige Intelligenzminderung
50 ≥ IQ < 70	Leichte Intelligenzminderung
70 ≥ IQ < 85	Niedrige Intelligenz
85 ≥ IQ < 115	Durchschnittliche Intelligenz
115 ≥ IQ < 129	Hohe Intelligenz
IQ > 129	Sehr hohe Intelligenz

8.2.1.4 Fragebogen zur Erfassung emotionaler und sozialer Schulerfahrungen (FEESS 3–4)

Zur Erhebung der Lernfreude der Schülerinnen und Schüler der Experimentalgruppe wird die Subskala „Lernfreude“ des Fragebogens zur Erfassung emotionaler und sozialer Schulerfahrungen (Rauer & Schuck, 2003) eingesetzt. Der FEESS 3–4 ist ein Gruppentestverfahren und kann ab dem zweiten Halbjahr der dritten Klasse eingesetzt werden. Er ist bis zum Ende der vierten Klasse normiert. Das Verfahren dient der Erfassung emotionaler und sozialer Erfahrungen von Schulkindern und besteht aus zwei Teilfragebögen. In dem ersten Teilfragebogen (SIKS) werden das Selbstkonzept der Fähigkeit, die erlebte soziale Integration und das Klassenklima erfasst. In dem zweiten Teilfragebogen (SALGA) geht es um die Schuleinstellung, die Lernfreude, die Anstrengungsbereitschaft und um das Gefühl des Angenommenseins.

„Die Skala Lernfreude (LF) erfasst das Ausmaß, in dem ein Kind die alltäglichen schulischen Aufgaben mit einem positiven Gefühl verbindet, mit froher Erwartungshaltung sowohl an die

Aufgaben im Unterricht als auch die Schulfächer insgesamt herangeht. Es handelt sich um keine fachspezifische, sondern um eine generalisierte Lernfreude im schulischen Alltag“ (Rauer & Schuck, 2003, S. 14 f.). Die Skala umfasst 13 Aussagen, die mit „stimmt gar nicht“, „stimmt kaum“, „stimmt ziemlich“ bis „stimmt genau“ beantwortet werden sollen.

Der FEESS 3-4 gilt als objektiv, da er genaue Angaben zur Durchführung, Auswertung und Interpretation gibt (Rauer & Schuck, 2003, S. 50). Die innere Konsistenz der Skala Lernfreude liegt bei Cronbachs $\alpha = .86$ und die Retestreliabilität dieser Skala nach vier Wochen (in der Jahrgangsstufe vier) bei $r = .71$ bzw. nach einem Jahr (in der Jahrgangsstufe drei) bei $r = .62$. Die beiden Teilfragebögen werden durch eine Faktorenanalyse bestätigt, die Interkorrelationen der Skalen sind alle hochsignifikant (Rauer & Schuck, 2003, S. 55). Für die Skala Lernfreude wird zuerst für jeden Schüler manuell ein Rohwert ermittelt, der dann mittels Normtabellen in einen Prozentrang und T-Wert transformiert werden kann. Der FEESS 3–4 bietet sich als Instrument, das den hohen Ansprüchen an die wissenschaftliche Güte erfüllt, zum Einsatz im Rahmen der vorliegenden Untersuchung an. Es werden hier die erreichten Rohwerte fokussiert, sodass kein Vergleich mit den Normwerten notwendig ist.

8.2.1.5 Implementationsfragebogen

Im Anschluss an die durchgeführte Fördermaßnahme füllen die Schülerinnen und Schüler einen Implementationsfragebogen aus, der von der Projektleitung entwickelt wird. Dieser Implementationsfragebogen soll u. a. Hinweise darauf geben, inwiefern die Schülerinnen und Schüler die eingesetzten Methoden als lernförderlich wahrnehmen und auch wie viel Freude sie an der Fördermaßnahme insgesamt haben. „Implementation“ wird dabei verstanden als die „*Diskrepanz* zwischen der Planung und Verwirklichung von Konzepten“ (Petermann, 2015b, S. 122 Herv. i. Orig.). Erhoben wird im Rahmen der vorliegenden Untersuchung die Implementations*qualität* (siehe Kapitel 4.4).

Der Fragebogen für die Schülerinnen und Schüler ist ein Paper-and-Pencil-Bogen mit 35 geschlossenen Ausagen und zwei offenen Fragen. Die 35 positiv gerichteten Aussagen werden anhand einer 4-stufigen Likertskala, die an der des Fragebogens zur Erfassung emotionaler und sozialer Schulerfahrungen (Rauer & Schuck, 2003) angelehnt ist („stimmt gar nicht“, „stimmt kaum“, „stimmt ziemlich“, „stimmt genau“), beurteilt. Zur Übung der Bearbeitung der Skala werden zu Beginn zwei Übungssätze vorangestellt, die ebenso dem Fragebogen zur Erfassung emotionaler und sozialer Schulerfahrungen (Rauer & Schuck, 2003) entnommen sind: „Ich kann gut mit dem Fahrrad fahren“, „Ich kann Disco-Musik nicht leiden“.

Die beiden offenen Fragen lauten: „Was hat dir an den Olympischen Rechenspielen besonders Spaß gemacht?“ und „Was würdest du gerne an den Olympischen Rechenspielen ändern?“

Die 35 Aussagen beziehen sich auf zwei Implementationsvariablen nach Petermann (2015b, S. 123 ff.):

> 1. „Akzeptanz“: Inwiefern ist die Maßnahme für die Schülerinnen und Schüler bezüglich verschiedener Aspekte zufriedenstellend?
>
> 2. „Angemessenheit“: Inwiefern beurteilen die Schülerinnen und Schüler die eingesetzten Methoden als lernförderlich?

Tabelle 38 gibt einen Überblick über die Konstruktion des Fragebogens. Der gesamte Fragebogen ist im Detail im Anhang 3 einzusehen.

TABELLE 38 KONSTRUKTION DES IMPLEMENTATIONSFRAGEBOGENS FÜR DIE SCHÜLERINNEN UND SCHÜLER

Variable	Operationalisierung	Anzahl Items	Beispielitems	Abk.
Akzeptanz	• Einschätzung des eigenen Lernerfolgs	5	• „Nach den Olympischen Rechenspielen fällt mir das Plusrechnen leichter.“ • „Durch die Olympischen Rechenspiele hatte ich Freude am Rechnen.“ • „Die Olympischen Rechenspiele haben mir gefallen.“	AK_{LE}
	• Zufriedenheit mit den Methoden	9		AK_{ZF}
	• Zufriedenheit insgesamt	2		AK_{ZI}
	• Geben einer Schulnote	1	• „Ich gebe den Olympischen Rechenspielen die Note:“	AK_{Note}
Angemessenheit	• Lernförderlicher Einsatz der Arbeitsblätter	6	• „Ich wusste, wie ich die Aufgaben auf den Arbeitsblättern lösen konnte.“ • „Die Erklärungen der Rechenwege an der Tafel haben mir beim Lösen von Aufgaben geholfen.“ • „Ich habe auf dem Stadionplan erkannt, ob ich mich in den Rechensprints verbessert oder verschlechtert habe.“	AN_{AB}
	• Lernförderlicher Einsatz des Advance Organizers	2		AN_{AO}
	• Lernförderlicher Einsatz der Direkten Instruktion	2		AN_{DI}
	• Lernförderlicher Einsatz des formativen Leistungsfeedbacks	2		AN_{CBM}
	• Lernförderlicher Einsatz der peer-mediierten Lernformen	2		AN_{PM}
	• Lernförderlicher Einsatz des Verhaltensfeedbacks	4		AN_{VF}

Die interne Konsistenz über die 34 Items (ausgeschlossen der Frage nach der Notenvergabe) liegt bei Cronbachs $\alpha = .924$.

8.2.2 InterventionsleiterInnen- und Lehrererhebung

Zur Erhebung des Sozial- und Lernverhaltens wird die Lehrereinschätzliste für Sozial- und Lernverhalten (Petermann & Petermann, 2013) eingesetzt. Diese wird von den hospitierenden Mathematiklehrkräften der teilnehmenden Klassen ausgefüllt. Die durchführenden Interventionsleiterinnen und -leiter bearbeiten, wie die Schülerinnen und Schüler, im Anschluss an die Fördermaßnahme einen Fragebogen zur Implementationsqualität. Beide Instrumente werden im Folgenden dargestellt.

8.2.2.1 Lehrereinschätzliste für Sozial- und Lernverhalten (LSL)

Die Lehrereinschätzliste für Sozial- und Lernverhalten (LSL) (Petermann & Petermann, 2013) ist ein Ratingverfahren und kann bei Schülerinnen und Schülern im Alter zwischen 6 und 19 Jahren eingesetzt werden. Die Lehrereinschätzliste besteht aus dem Teilbereich „Sozialverhalten", der aus sechs Skalen besteht, und aus dem Teilbereich „Lernverhalten", der aus vier Skalen besteht. Den Skalen sind jeweils fünf Aussagen zu dem jeweiligen Schüler bzw. der Schülerin zugeordnet, die durch die Lehrkraft auf einer vierstufigen Skala beantwortet werden (zum Beispiel „Schließt Kompromisse" oder „Nimmt eigene Gefühle wahr"). Die LSL wird folglich für je einen Schüler bzw. eine Schülerin ausgefüllt und soll sich auf die Beobachtungen im Zeitraum der letzten vier Wochen beziehen.

TABELLE 39 ÜBERBLICK ÜBER DIE SKALEN DER LEHREREINSCHÄTZLISTE FÜR SOZIAL- UND LERNVERHALTEN (PETERMANN & PETERMANN, 2013, S. 7)

Sozialverhalten	Lernverhalten
Kooperation	Anstrengungsbereitschaft/Ausdauer
Selbstwahrnehmung	Konzentration
Selbstkontrolle	Selbstständigkeit beim Lernen
Einfühlungsvermögen/Hilfsbereitschaft	Sorgfalt beim Lernen
Angemessene Selbstbehauptung	
Sozialkontakt	

Die Bearbeitungsdauer erstreckt sich auf ca. fünf Minuten pro Bogen. Für jede Skala der beiden Teilbereiche werden auf Grundlage der Rohwerte Prozentränge und T-Werte ermittelt. Ein Prozentrang von 10 oder kleiner weist auf eine starke Verhaltensabweichung hin, ein Prozentrang zwischen 10 und 20 zeigt ein risikobehaftetes Verhalten an (Petermann & Petermann, 2013, S. 23). Die Reliabilität für die einzelnen Skalen liegt bei $.82 \leq \alpha \leq .95$. Im Rahmen einer Validitätsprüfung können hohe Korrelationen zwischen der Deutsch- und Mathematiknote und dem durch die LSL erhobenen Sozial- und Lernverhalten gefunden werden. Zudem bestätigt eine Faktorenanalyse die Einteilung der Skalen in das Sozial- und Lernverhalten (Petermann & Petermann, 2013, S. 7).

Mittlerweile ist auch eine Schülereinschätzliste für das Sozial- und Lernverhalten (Petermann & Petermann, 2014) veröffentlicht, die jedoch zum Zeitpunkt der Umsetzung der vorliegenden Untersuchung noch nicht zugänglich ist.

8.2.2.2 Implementationsfragebogen

Der Implementationsfragebogen für die Interventionsleiterinnen und -leiter umfasst die Skalen Akzeptanz, Übernahme(-bereitschaft), Angemessenheit, Machbarkeit und Wiedergabetreue. Diese sind zentrale Variablen der Implementations*qualität* (Petermann, 2015b, S. 123). Auch hier wird unter Implementation die „*Diskrepanz* zwischen der Planung und Verwirklichung von Konzepten“ (Petermann, 2015b, S. 122 Herv. i. Orig.) verstanden.

Der Fragebogen besteht aus 58 geschlossenen Aussagen, die mit einer vierstufigen Likertskala von „stimmt gar nicht“, „stimmt kaum“, „stimmt ziemlich“ bis „stimmt genau“ beantwortet werden. Tabelle 40 gibt einen groben Überblick über die fünf Implementationsvariablen, die Operationalisierung, die Anzahl der Items und Beispielitems. Der gesamte Fragebogen ist im Anhang 1 einzusehen.

TABELLE 40 KONSTRUKTION DES IMPLEMENTATIONSFRAGEBOGENS FÜR DIE INTERVENTIONSLEITERINNEN UND -LEITER

Variable	Operationalisierung	Anzahl Items	Beispielitems	Abk.
Akzeptanz	• Zufriedenheit	5	• „Die Intervention hat meine Erwartungen insgesamt erfüllt.“	AK_{ZI}
	• Geben einer Schulnote	1	• „Ich gebe den Olympischen Rechenspielen die Note:…“	AK_{Note}
Übernahme	• Erwägung erneuter Durchführung	2	• „Ich würde die Olympischen Rechenspiele weiterempfehlen.“	Ü
Angemessenheit	• Passung für Zielgruppe	9	• „Die Olympischen Rechenspiele sind für die Zielgruppe (SuS an Ober- oder Gesamtschulen) angemessen.“ • „Die Materialien sind für leistungsschwache Schülerinnen und Schüler lernförderlich differenziert.“ • „Die Strategieerklärungen an der Tafel helfen den Schülerinnen und Schülern beim Lösen von Aufgaben.“ • „Die Arbeitsblätter helfen den Schülerinnen und Schülern beim Lösen von Aufgaben.“ • „Der Advance Organizer verdeutlicht den Schülerinnen und Schülern das Thema der Stunde.“ • „Die Stempel sind als positive Verstärker angemessen.“	AN_{ZG}
	• Passung der Inhalte	1		AN_{PI}
	• Passung der Methoden	1		AN_{PME}
	• Passung der Zeitplanung	4		AN_{ZP}
	• Lernförderlicher Einsatz des Advance Organizers	2		AN_{AO}
	• Lernförderlicher Einsatz der Arbeitsblätter	1		AN_{AB}
	• Lernförderlicher Einsatz der Direkten Instruktion	2		AN_{DI}
	• Lernförderlicher Einsatz der Spiele	1		AN_{SP}
	• Lernförderlicher Einsatz der Peer-Mediierten Methoden	3		AN_{PM}
	• Lernförderlicher Einsatz des formativen Leistungsfeedbacks	2		AN_{CBM}
	• Lernförderlicher Einsatz des Verhaltensfeedbacks	6		AN_{VF}

Variable	Operationalisierung	Anzahl Items	Beispielitems	Abk.
Machbarkeit	• Umsetzbarkeit mit den Ressourcen im Feld Schule	6	• „Das Durchführungsmanual ist einfach zu handhaben.“ • „Die Materialien sind eindeutig gekennzeichnet.“ • „Meine Schule verfügt über die notwendigen Ressourcen (Materialien, Ausstattung) zu einer erfolgreichen Umsetzung.“	M_{SR}
	• Eignung des Manuals	5		M_{M}
	• Umsetzbarkeit mit den persönlichen Ressourcen	1		M_{PR}
Wiedergabetreue	• Einhalten der Vorgaben des Manuals	6	• „Ich habe die Formulierungen der direkten Instruktionen übernommen.“	W

Einige Aussagen erfordern eine zusätzliche Präzisierung der Antwort (zum Beispiel „Die Zeitplanung in den Stunden ist zu kurz.“ → falls „stimmt genau“ oder „stimmt ziemlich“: „In welchen Stunden?). Zudem beinhaltet auch der Implementationsbogen für die Interventionsleiterinnen und -leiter offene Fragen:

> 1. Welche Materialien haben sich bei der Durchführung als besonders hilfreich erwiesen? Warum?
>
> 2. Welche Schwierigkeiten zeigten sich bei der inhaltlichen und konzeptionellen Durchführung der Olympischen Rechenspiele? Warum?
>
> 3. Welche weiteren Anregungen bzw. Verbesserungsvorschläge haben Sie?

Die interne Konsistenz über die 57 Items (ausgeschlossen der Frage nach der Notenvergabe) liegt bei Cronbachs $\alpha = .849$.

8.3 Stichprobenrekrutierung

Im Frühjahr 2014 werden alle relevanten Ober- und (Integrierten) Gesamtschulen im Stadtgebiet Oldenburg zur Rekrutierung von Schülerinnen und Schülern zur Teilnahme an der Studie einbezogen ($N_{gesamt} = 8$). Eine Auflistung der Oldenburger Schulen ist über eine Internetseite der Stadt Oldenburg (www.oldenburg.de/microsites/schule/schulen-in-oldenburg.html) zugänglich. Die anvisierte Stichprobe ist folglich als Klumpenstichprobe (cluster sample) zu bezeichnen. Die Zielpopulation sind Ober- und Gesamtschülerinnen und Schüler in der fünften Klasse. Die Klumpen der Zielpopulation sind die einzelnen Schulen (Döring & Bortz, 2016c, S. 314).

Die Schulen werden von der Projektleitung per Losverfahren randomisiert der Kontroll- und Experimentalgruppe zugeordnet. Nach der Randomisierung werden die Schulen mit einem Informationsbrief bzgl. der Studie angeschrieben und zur Teilnahme beworben. Nach ca. ein bis zwei Wochen erfolgt eine telefonische Rückfrage, ob sich die jeweilige Schule eine Teilnahme an der Studie vorstellen kann. Nachdem alle Ab- und Rücksprachen getroffen sind, erklären sich sieben der acht Schulen zur Teilnahme bereit (87,5 %). Davon fallen drei auf die Experimentalgruppe und vier auf die Kontrollgruppe. Eine Schule der Experimentalgruppe erklärt sich

lediglich mit zwei Klassen zur Teilnahme an der Studie bereit, in allen anderen teilnehmenden Schulen werden alle Klassen in der fünften Jahrgangsstufe in die Studie einbezogen.

Die Experimentalgruppe besteht abschließend aus ingsgesamt 11 Klassen, die Kontrollgruppe aus 15 Klassen. Die Durchführung der Studie inkl. der Erhebungen in den rekrutierten Schulen wird durch die Niedersächsische Landesschulbehörde genehmigt.

8.4 Untersuchungsdurchführung

Der Entwicklungsprozess der Maßnahme im Rahmen einer Projektgruppe ist bereits dargestellt (siehe Kapitel 6.2). Die Beschreibung der Untersuchungsdurchführung beginnt im Folgenden mit der Pilotstudie, die der Hauptstudie vorangestellt war.

PILOTSTUDIE

Die Rekrutierung von Schülerinnen und Schülern zur Teilnahme an der Studie beginnt im Frühjahr 2014 (für das konkrete Vorgehen siehe Kapitel 8.3). Nach Rücksprache mit der Schulleitung erklärt sich eine Integrierte Gesamtschule in Oldenburg neben der Teilnahme an der Hauptstudie für die Durchführung der vorangestellten Pilotstudie bereit. Die Pilotstudie (siehe Tabelle 41) erfolgt vor den Sommerferien im Juni und Juli 2014 in einer fünften Klasse dieser Schule mit 30 Schülerinnen und Schülern. Zwei Studentinnen der Projektgruppe führen die Pilotstudie gemeinsam durch. Da die Durchführung der Hauptstudie für den Herbst 2014 zu Beginn der fünften Jahrgangsstufe geplant ist, ist somit nur eine Pilotierung in einer Lerngruppe möglich, die sich im Zeitraum der Durchführung *am Ende* der fünften Jahrgangsstufe befindet. Während der Pilotuntersuchung füllen die beiden Interventionsleiterinnen nach jeder Stunde einen Evaluationsbogen aus, mit dem die Stunde zur weiteren Überarbeitung methodisch und inhaltlich reflektiert wird (formative Evaluation).

Im August 2014 findet im Rahmen eines Projekttages die Reflexion und die Evaluation der Pilotstudie auf Grundlage der formativen Evaluation der Pilotuntersuchung durch die Interventionsleiterinnen statt (Mittag & Bieg, 2010, S. 34). Die benannten Probleme und Fallstricke der einzelnen Stunden werden in der Folge bearbeitet und behoben. Zeitgleich zur Durchführung der Pilotstudie werden auf verschiedenen Wegen (in Seminaren und Vorlesungen) Masterstudierende des Lehramts Sonderpädagogik zur Umsetzung der Maßnahme und Bachelorstudierende desselben Studiengangs zur Durchführung der Diagnostik in der Kontrollgruppe im Rahmen der Hauptstudie beworben.

TABELLE 41 ZEITLICHER ABLAUF DER PILOTSTUDIE

Datum/Zeitraum	Schritte im Projekt
16.06.2014 – 25.07.2014	Durchführung der Pilotstudie in einer fünften Klasse (n = 30) am Ende des Schuljahres
11.08.2014	Reflexion und Evaluation der Umsetzung der Pilotstudie
August 2014	Überarbeitung der Unterrichtsstunden und der Materialien in der Projektgruppe

HAUPTSTUDIE

Im September und Oktober des Jahres 2014 findet nach intensiver Überarbeitung der Fördermaßnahme die Hauptstudie statt. Die Umsetzung der Maßnahme in der Experimentalgruppe führen 13 Studentinnen des Studiengangs Master of Education Sonderpädagogik durch. Neun dieser Studentinnen sind für die Durchführung neu hinzugekommen, vier sind bereits an der Entwicklung der Maßnahme beteiligt. Zwölf weitere Studierende des Bachelorstudiengangs Sonderpädagogik führen die Erhebungen bei den Schülerinnen und Schülern der Kontrollgruppe durch. Diese Studierenden sind ebenso neu in das Projekt hinzugekommen und sind noch nicht mit der Studie vertraut. Zudem unterstützen zwei studentische Hilfskräfte die Durchführung der Studie.

Bevor die Umsetzung der Hauptstudie beginnt, werden die Interventionsleiterinnen und -leiter intensiv bezüglich der Maßnahme und den Erhebungsinstrumenten geschult. Zudem erhalten die Interventionsleiterinnen und -leiter eine Einführung in das Thema Klassenführung, um in der Umsetzung der Maßnahme und der Diagnostik eine lernförderliche Atmosphäre in den Klassen zu schaffen und in präventiven sowie interventiven Strategien im Umgang mit Verhaltensstörungen geschult zu sein.

Zur Umsetzung der Maßnahme erhalten die Interventionsleiterinnen und -leiter ein gebundenes Materialmanual, in dem alle Kopiervorlagen zu finden sind, ein gebundenes Durchführungsmanual mit tabellarischen Stundenabläufen, Hinweisen zur Durchführung und ausformulierten Instruktionen. Zudem erhalten sie den Stempel für das Feedbacksystem (siehe Abbildung 43) und den Advance Organizer (blanko) in DIN-A0-Format (siehe Abbildung 39). Für die Diagnostik bekommen die Interventionsleiterinnen und -leiter, wie auch die Studierenden in der Kontrollgruppe, ein gebundenes Manual, in dem die Instruktionen und wichtige Hinweise für die Erhebungen zu finden sind. Sofern es möglich ist, arbeiten die Interventionsleiterinnen und -leiter und die Studierenden in der Kontrollgruppe in Tandems.

Die Durchführung der Studie wird durch Supervisionssitzungen durch eine wissenschaftliche Mitarbeiterin begleitet. Die Projektleitung wird im Oktober 2014 durch zwei Doktorandinnen und einen Doktoranden erweitert.

Nach der Durchführung der Maßnahme findet im Oktober 2014 ein Reflexionstreffen statt, um die Implementation der Maßnahme im Rahmen der Hauptstudie zu reflektieren. Dabei geht es sowohl um Stärken als auch um Schwächen des Konzepts.

Im Januar 2015 und im Juli 2015 erfolgten je ein Follow-Up-Messzeitpunkt (I und II), der von den jeweiligen Interventionsleiterinnen und -leitern und den Bachelorstudierenden durchgeführt wird. Es werden dafür keine weiteren Studierenden für das Projekt rekrutiert. Die Studierenden in beiden Untersuchungsgruppen haben die Gelegenheit, ihre Abschlussarbeiten im Rahmen des Projekts zu verfassen und erhalten keine finanzielle Vergütung. Zur Unterstützung in der Dateneingabe und in der statistischen Auswertung wird eine Einführung in das Statistikprogramm SPSS angeboten.

Ein Abschluss des Projekts erfolgt im Juli 2015 sowohl für die mitarbeitenden Studierenden als auch für die teilnehmenden Lehrkräfte in den Räumlichkeiten der Universität.

TABELLE 42 ZEITLICHER ABLAUF DER HAUPTSTUDIE

Datum/Zeitraum	Schritte im Projekt
Frühjahr 2014	Rekrutierung von Schulen zur Teilnahme an der Studie
	Pilotierung der Lernverlaufsdiagnostik
Juli 2014	Anwerben von Studierenden für die Umsetzung der Hauptstudie
01.–02.09.2014	Einführung der Interventionsleiterinnen und -leiter in die Maßnahme und in die Diagnostik (Experimentalgruppe)
08.09.2014	Einführung der Bachelorstudierenden in die Diagnostik (Kontrollgruppe)
September & Oktober 2014	Durchführung der Hauptstudie (inkl. Supervisionssitzungen)
06.10.2014	Reflexion der Hauptstudie in der Projektgruppe
05.11.2014	Interventionsleiterinnen und -leiter (Experimentalgruppe): Einführung in SPSS
07.11.2014	Studierende der Kontrollgruppe: Einführung in SPSS
Januar 2015	Follow-Up-I
Juli 2015	Follow-Up-II
08.07.2015	Projektabschluss mit den Interventionsleiterinnen und -leitern
15.07.2015	Projektabschluss mit den Lehrkräften
16.07.2015	Projektabschluss mit den Studierenden der Kontrollgruppe

Insgesamt arbeiten an der Entwicklung und Umsetzung des Projekts auf universitärer Seite 17 Studentinnen des Masters of Education Sonderpädagogik, 12 Studierende des Bachelor of Arts Sonderpädagogik, zwei studentische Hilfskräfte, zwei Doktorandinnen, ein Doktorand und eine wissenschaftliche Mitarbeiterin sowie Prof. Dr. Clemens Hillenbrand als Lehrstuhlinhaber und Schirmherr des Projekts.

8.5 Forschungsethik

Bevor die Auswertungsstrategie der Daten aufgezeigt wird, soll an dieser Stelle rechtschaffen dargelegt werden, wie im Rahmen der Studie forschungsethische Aspekte berücksichtigt werden, denn „weder Menschen noch Tiere dürfen für wiss. Zwecke (d. h. v. a. durch die verwendeten Forschungsdesigns und Datenerhebungsverfahren) in unethischer Weise missbraucht oder geschädigt werden“ (Döring, 2014a). In der Literatur sind verschiedene Aspekte aufzufinden, die im Rahmen von Forschungsethik in der Wissenschaft berücksichtigt werden sollten (vgl. Döring, 2014a; Hussy et al., 2010). Es folgt hier eine Orientierung an den vier Richtlinien nach Hussy und Kollegen (2010, S. 43 f.).

PSYCHISCHE UND PHYSISCHE UNVERSEHRTHEIT DER TEILNEHMERINNEN UND TEILNEHMER

Die Teilnehmerinnen und Teilnehmer in der Studie bearbeiten Testbögen, die für ihre Altersgruppe konzipiert sind, sodass diese als angemessen angesehen werden. Es wird kein anstößiges oder provozierendes Untersuchungsmaterial verwendet. Die methodisch-didaktische und in-

haltliche Aufbereitung der Maßnahme beruht auf wissenschaftlichen Befunden zur Lernförderung im Bereich Mathematik. Die Erhebungen werden anonym durchgeführt, indem jeder Schüler und jede Schülerin eine depersonalisierte Kennziffer erhält. Die Interventionsleiterinnen und -leiter werden im Vorfeld der Studie intensiv geschult bzgl. der Durchführung der Diagnostik und der Fördermaßnahme. Sie erhalten auch Strategien und Hinweise für ein gutes Classroom-Management, um eine lernförderliche Atmosphäre zu schaffen. Entscheidet sich ein Schüler bzw. eine Schülerin dazu, nicht an der Erhebung teilzunehmen, wird dies akzeptiert und lediglich für die Aufklärung der Drop-Out-Quoten festgehalten.

TRANSPARENZ FÜR DIE VERSUCHSPERSONEN

Die Lehrkräfte werden vor, während und nach der Studie von der Projektleitung via Informationsschreiben über die (nächsten) Projektschritte informiert. Im Rahmen eines Präsentationsnachmittags (siehe Tabelle 42) werden die Ergebnisse am Ende der Studie von der Projektleitung präsentiert und ein Dankeschön an die Lehrkräfte ausgesprochen. Die Schülerinnen und Schüler werden vor jeder Testdurchführung von den Studierenden über die Ziele der Studie informiert. Es wird gemeinsam mit den Schülerinnen und Schülern rekapituliert, weshalb sie an der Studie teilnehmen und welche wissenschaftlichen Erkenntnisse daraus gewonnen werden können.

VERMEIDUNG VON TÄUSCHUNG

Im Rahmen der Studie werden keine Täuschungen vorgenommen. Den Lehrkräften und den Schülerinnen und Schülern werden die Ziele der Studie unter Berücksichtigung der Gruppenzugehörigkeit (Kontroll- vs. Experimentalgruppe) transparent gemacht und erläutert.

FREIWILLIGKEIT DER TEILNAHME

Die Teilnahme an der Studie ist für alle Beteiligten freiwillig. Es wird im Vorfeld die Genehmigung für die Studie seitens der Niedersächsischen Landesschulbehörde eingeholt (gemäß dem Erlass zu Umfragen und Erhebungen in Schulen). In dem Antrag werden u. a. Angaben über die Freiwilligkeit an der Teilnahme, Aspekte der Anonymisierung, Beachtung datenschutzrechtlicher Bestimmungen, der Umgang mit Daten bzw. Verarbeitung im wissenschaftlichen Kontext und der Zeitpunkt der endgültigen Vernichtung der Daten erläutert. Die Eltern der Schülerinnen und Schüler erhalten zudem eine Einverständniserklärung, die sie im Vorfeld der Studie ausfüllen. Es wid niemand zur Teilnahme an der Studie gezwungen.

8.6 Auswertungsstrategien

Im Folgenden werden die Auswertungsstrategien entsprechend der Fragestellungen und Hypothesen vorgestellt. Das folgende Kapitel 8.6.1 ermöglicht dabei einen Überblick über die eingesetzten Verfahren in der jeweiligen Hypothesenprüfung. Die genannten statistischen Verfahren werden dann im darauffolgenden Kapitel 8.6.2 im Detail erläutert.

8.6.1 Überblick über die Auswertungsstrategien der Hypothesen

Folgende Tabelle 43 zeigt, welche statistischen Verfahren in der Beschreibung der Stichprobe bzw. in der Prüfung der Hypothesen zur Anwendung kommen.

TABELLE 43 AUSWERTUNGSSTRATEGIEN ZU DEN HYPOTHESEN

Bezug	Auswertungsstrategie
Beschreibung der Stichprobe	• χ^2-Test • t-test bei unabhängigen Stichproben inkl. Prüfung der Voraussetzungen • ggf. Mann-Whitney-U-Test (bei Verletzung der Voraussetzungen)
H_1–H_7	• Zweifaktorielle Varianzanalyse mit Messwiederholung inkl. Prüfung der Voraussetzungen • Effektstärke (d) nach Cohen (1988) • Teststärke/Power • Detailanalyse: t-tests inkl. effektstärkenbasierter sukzessiver α-Abwärtsadjustierung
H_8	• Multiple lineare Regression inkl. Prüfung der Voraussetzungen
H_9–H_{15}	• t-test bei einer Stichprobe inkl. Prüfung der Voraussetzungen • ggf. Wilcoxon-Test (bei Verletzung der Vorausetzungen)

8.6.2 Beschreibung der Strategien und Verfahren

Die im Folgenden vorgestellten Strategien und Verfahren werden mittels des Programms „Statistical Package for the Social Sciences“ (SPSS) in seiner 23. Version und des Programms G*Power durchgeführt.

8.6.2.1 χ^2-Test

„Aufgabe der χ^2-Verfahren ist es, die Häufigkeitsverteilung zu analysieren, die aus der Einteilung der Versuchsobjekte in verschiedene Kategorien entstanden ist“ (Rasch, Friese, Hofmann & Naumann, 2006a, S. 172). Grundlage des Verfahrens sind Nominaldaten, aus denen (genuin) Häufigkeiten ermittelt werden können (Rasch et al., 2006a, S. 171). Das am häufigste verwendete Verfahren ist der „χ^2-Test auf Unabhängigkeit“ bzw. der „zweidimensionale χ^2-Test“, der zum Einsatz kommt, wenn geprüft werden soll, ob „zwischen [...] zwei untersuchten Merkmalen ein Zusammenhang besteht“ (Sedlmeier & Renkewitz, 2013, S. 553).

Da der χ^2-Test zu den nonparametrischen Verfahren gehört, gehen mit ihm weniger Voraussetzungen einher. Damit unterscheidet sich das Verfahren maßgeblich zum Beispiel vom parametrischen t-Test oder der Varianzanalyse. Voraussetzungen für den χ^2-Test sind (nach Sedlmeier & Renkewitz, 2013, S. 564):

1. Jede Person muss eindeutig einer bestimmten Merkmalsausprägung bzw. Merkmalskombination zugeordnet werden können (d. h. eine Person darf nicht zeitgleich mehreren Merkmalsausprägungen zugeordnet werden)
2. Mehrere Beobachtungen müssen voneinander unabhängig sein.

Durch χ^2-Tests wird im Rahmen dieser Studie untersucht, ob zum Beispiel Merkmale wie „Geschlecht“, „gesprochene Sprache zu Hause“ oder die „besuchte Schulform“ in der Experimental- und Kontrollgruppe gleich verteilt sind, d. h. gleich häufig vorkommen. Die Ergebnisse können Hinweise auf die Qualität der Rekrutierung der Stichprobe geben.

8.6.2.2 t-Test

Der t-Test zählt zu den parametrischen Verfahren (Rasch, Friese, Hofmann & Naumann, 2006b, S. 44) und prüft für zwei Gruppen (Rasch et al., 2006b, S. 45), „ob sich [die] Populationsmittelwerte signifikant unterscheiden“ (Sedlmeier & Renkewitz, 2013, S. 396). Der t-Test arbeitet dabei mit der Streuung und dem arithmetischen Mittel der Werte (Rasch et al., 2006b, S. 44). Die Mittelwertdifferenz zwischen den beiden Gruppen ist dann signifikant, „wenn die ermittelte Wahrscheinlichkeit des t-Werts kleiner ist als der angenommene α-Fehler bzw. der empirische t-Wert größer ist als der Betrag des kritischen t-Werts“ (Rasch et al., 2006b, S. 60). Zentral unterschieden wird dabei der t-Test für unabhängige Stichproben (zum Beispiel Jungen und Mädchen) von dem t-Test für abhängige Stichproben (zum Beispiel eine Stichprobe zu zwei Messzeitpunkten). Zudem gibt es die Möglichkeit, einen t-Test bei nur einer Stichprobe durchzuführen und den berechneten Mittelwert gegen einen festgesetzten Erwartungswert zu testen. Letzteres erfolgt nach ähnlichem Prinzip wie der t-Test für abhängige Stichproben.

Nach Sendlmeier und Renkewitz (2013, S. 399 ff.) sind vor der Durchführung eines t-Tests folgende Voraussetzungen zu prüfen:

- intervallskalierte Messwerte
- Normalverteilung der Populationswerte
- gleiche Varianzen in den Populationen

„Allerdings sind t-Tests relativ robust gegenüber Verletzungen der Annahmen über die Verteilung der Populationswerte, es sei denn, die Stichproben sind sehr klein“ (Sedlmeier & Renkewitz, 2013, S. 400). Die Stichproben sollten dabei ungefähr die gleiche Größe haben und mehr als 30 Probanden beinhalten (Rasch et al., 2006b, S. 59). Bezüglich der Signifikanzschwelle wird ein α von 0.05 angenommen.

Sind die Voraussetzung zum Einsatz der parametrischen t-Tests nicht gegeben, kann auf non-parametrische Verfahren zurückgegriffen werden (siehe Tabelle 44).

TABELLE 44 PARAMETRISCHE UND ZUGEHÖRIGE NON-PARAMETRISCHEN VERFAHREN

Parametrisch	Non-parametrisch
t-Test bei unabhängigen Stichproben	Mann-Whitney-U-Test
t-Test bei abhängigen Stichproben	Wilcoxon-Test
t-Test bei einer Stichprobe	Wilcoxon-Test

Für den t-Test bei unabhängigen Stichproben bietet sich alternativ der Mann-Whitney-U-Test an, bei dem t-Test für abhängige Stichproben und bei dem t-Test bei einer Stichprobe der Wilcoxon-Test.

8.6.2.3 Varianzanalyse mit Messwiederholung

Eine Varianzanalyse beantwortet grundsätzlich die Frage „ob zwischen [...] Stichprobenmittelwerten Unterschiede bestehen, die nicht mehr plausibel durch eine zufällige Abweichung der Stichprobenmittelwerte vom *gleichen* Populationsmittelwert erklärt werden können. Dies geschieht [...] durch eine Betrachtung von Varianzen“ (Sedlmeier & Renkewitz, 2013, S. 422 f.; Herv. i. Orig.). Im Vergleich zum t-Test dient die Varianzanalyse dem Vergleich von mehr als zwei unabhängigen Stichproben (Zöfel, 2003, S. 129). In wissenschaftlichen Experimenten kann vorgesehen sein, einige Faktoren wiederholt zu untersuchen (z. B. mathematische Kompetenz), andere Faktoren hingegen nicht (z. B. das Geschlecht). Die Varianzanalyse stellt somit eine Verallgemeinerung des t-tests dar (Rasch et al., 2006a, S. 1). Der Vorteil der Varianzanalyse ist dabei die Vermeidung der α-Fehler-Kumulierung, die bei der Durchführung mehrerer t-Tests auftreten würde (Sedlmeier & Renkewitz, 2013, S. 421).

Liegt ein Design vor, in dem mehrere (zum Beispiel: zwei) Faktoren die Messwerte von Personen systematisch beeinflussen können, wird von einer sog. „Zweifaktoriellen Varianzanalyse“ gesprochen, wird dabei ein Faktor wiederholt gemessen, von einer „Zweifaktoriellen Varianzanalyse mit Messwiederholung“ (Bühner & Ziegler, 2009, S. 482).

Folgende Fragestellungen werden mit der zweifaktoriellen Varianzanalyse mit Messwiederholung – mit Bezug auf die vorliegende Untersuchung – beantwortet (nach Bühner & Ziegler, 2009, S. 485):

1. Erzeugt der Faktor A (Gruppenfaktor; *hier: Zugehörigkeit zu den Untersuchungsgruppen*) Unterschiede in den Mittelwerten der AV (*hier: mathematische Basiskompetenz*)?
2. Erzeugt der Faktor B (Messwiederholungsfaktor, *hier: Messzeitpunkte t_1 bis t_4*) Unterschiede in den Mittelwerten der AV (*hier: mathematische Basiskompetenz*)?
3. Gibt es eine wechselseitige Beeinflussung der beiden Faktoren (*hier: Gruppenzugehörigkeit und Messwiederholung*)?

Die erste Frage bezieht sich auf die sog. Zwischensubjekteffekte, die zweite Frage auf die Innersubjekteffekte. Gemeinsam werden sie als Haupteffekte bezeichnet. Die dritte Frage bezieht sie auf die sog. Interaktionseffekte und betrachtet die Wechselwirkung der Faktoren.

Zur Anwendung der zweifaktoriellen Varianzanalyse mit Messwiederholung gelten folgende Voraussetzungen (Bühner & Ziegler, 2009, S. 518 f.):

INTERVALLSKALENNIVEAU UND NORMALVERTEILUNG

Die abhängige Variable muss intervallskaliert sein. Zudem muss die Normalverteilung der abhängigen Variable in allen Teilstichproben (zum Beispiel nach Gruppenzugehörigkeit in allen Messzeitpunkten) auf Normalverteilung geprüft werden. Zur Überprüfung eignet sich die Analyse von Histogrammen.

SPHÄRIZITÄT

Die Varianzen der paarweisen Differenzen zwischen den Messzeitpunkten müssen homogen sein. Zur Überprüfung eignet sich der Mauchly-W-Test. Sollte die Sphärizität verletzt sein, kann mit der Greenhouse-Geisser- oder mit der Huynh-Feld-Korrektur fortgefahren werden.

Die Sphärizitätsanalyse ist jedoch nur bei mehr als zwei einbezogenen Messzeitpunkten relevant, da sonst nur eine Differenzwertverteilung mit einer gewissen Varianz vorliegt, die nicht mit einer anderen Varianz verglichen werden könnte.

BALANCIERTHEIT DES DESIGNS

Als letzte Voraussetzung gilt die Balanciertheit des Designs: Für jede Person muss zu jedem Messzeitpunkt ein Messwert vorliegen. Personen mit fehlenden Messwerten werden in der zweifaktoriellen Varianzanalyse mit Messwiederholung nicht berücksichtigt.

Die zweifaktorielle Varianzanalyse mit Messwiederholung wird zur Verifizierung bzw. Falsifizierung der Hypothesen im Hypothesenblock I und II eingesetzt. Die abhängige Variable ist die mathematische Basiskompetenz in Form des summierten Rohwerts der Skala Rechenoperationen des Heidelberger Rechentests 1–4, da es Hinweise dazu gibt, dass „Rohwerte im unteren Leistungsbereich besser differenzieren als standardisierte Werte, wodurch Rohwerte individuelle Leistungsverbesserungen besser darstellen können“ (Ise, Engel & Schulte-Körne, 2012, S. 134). Die unabhängigen Variablen bilden die Messwiederholung und die Zugehörigkeit zu den Untersuchungsgruppen. Da eine unabhängige Variable einen Gruppenfaktor darstellt und die andere einen Messwiederholungsfaktor, sprechen Bühner und Ziegler (2009, S. 483) von einem „gemischten Design“. Bezüglich der Signifikanzschwelle wird hier ein α von 0.05 angenommen.

8.6.2.4 Effektstärke

Effektstärken geben bei Mittelwertdifferenzen an, wie *stark* der beobachtete Effekt ist. Es wird analysiert, wie viele Standardabweichungen zwei Mittelwerte auseinander liegen. Mit Hilfe der Effektstärke können Aussagen zu der praktischen Bedeutsamkeit (Größe und Richtung) eines Mittelwertunterschieds gemacht werden (Bühner & Ziegler, 2009, S. 177 f.). Dabei unterscheiden sich die Klassifikationen der Effektstärke (groß, mittel, klein) je nach Autorengruppe (Zierer, 2014, S. 117).

Die Tabelle 45 zeigt die im pädagogisch-psychologischen Kontext gängige Klassifikation der Effektstärke d nach Cohen (1988). Im Vergleich stellt die Klassifikation eine eher strenge Einordnung der Effektstärke dar (zur Übersicht siehe Zierer, 2014, S. 117). Da in Varianzanalysen das η^2 als Angabe der Effektstärke ausgegeben wird, wird diese Klassifikation ebenso aufgeführt. Die Seite www.psychometrica.de ermöglicht eine Umrechnung des η^2 in Cohen's d.

TABELLE 45 KLASSIFIKATION VON COHEN'S D UND η2 (LENHARD & LENHARD, 2016B)

Cohen's d	η^2	Effektgröße
$d < 0.2$	$\eta^2 < 0.01$	kein Effekt
$0.2 \leq d < 0.5$	$0.01 \leq \eta^2 < 0.06$	kleiner Effekt
$0.5 \leq d < 0.8$	$0.06 \leq \eta^2 < 0.14$	mittlerer Effekt
$0.8 \leq d$	$0.14 \leq \eta^2$	großer Effekt

Hattie und Kollegen (2013, S. 21) ergänzen, „die Effektstärke von 0.40 legt einen Punkt fest, ab dem Effekte einer Innovation die Lernleistung derart verbessern, dass wir in der realen Welt Unterschiede beobachten können". Eine Effektstärke von 1.0 bedeutet dabei, dass sich die Leistungen des Lernenden um eine Standardabweichung verbessert hat. Eine Verbesserung um eine Standardabweichung kann mit einem Lernfortschritt von zwei bis drei Jahren verglichen werden (Hattie et al., 2013, S. 9 f.). Hattie und Kollegen (2013, S. 11) setzen die Cut-Off-Werte etwas niedriger an und nehmen schon ab einer Effektstärke von $d = 0.2$ einen kleinen, ab $d = 0.4$ einen mittleren und ab $d = 0.6$ einen großen Effekt an.

8.6.2.5 α-Fehler-Adjustierung

Der α-Fehler bezeichnet die Wahrscheinlichkeit, die Alternativhypothese anzunehmen, obwohl in Wirklichkeit die Nullhypothese gilt. Konventionell wird das α-Fehler-Niveau bei 5 % angelegt. Wenn jedoch mehrere voneinander unabhängige Signifikanztests durchgeführt werden, wächst die Irrtumswahrscheinlichkeit. Das Gesamt-α-Niveau steigt, was als α-Fehlerkumulierung bezeichnet wird (Rasch et al., 2006a, S. 3). „Die Wahrscheinlichkeit, dass die H_0 durch einen von mehreren Tests fälschlicherweise zurückgewiesen wird, steigt also mit der Anzahl der durchgeführten Tests dramatisch an" (Rasch et al., 2006a, S. 4).

Rost (2013, S. 255) schlägt als Vorgehen zur α-Fehler-Adjustierung die „effektstärkenbasierte sukzessive α-Abwärtsadjustierung" vor. Dieses Vorgehen bietet sich an, wenn unterschiedliche Versuchspersonenzahlen oder unterschiedliche inferenzstatistische Verfahren zu Grunde liegen. Die signifikanten Ergebnisse werden nach der Effektstärke absteigend sortiert und das akzeptierte α-Niveau Schritt für Schritt verringert: Beim größten Effekt wird keine Adjustierung vorgenommen, beim zweitgrößten Effekt wird das α durch 2 dividiert, beim drittgrößten Effekt α durch 3 etc.

Folgende Tabelle 46 verdeutlicht das Vorgehen.

TABELLE 46 EFFEKTSTÄRKENBASIERTE SUKZESSIVE α-ABWÄRTSADJUSTIERUNG

Rang nach Effektgröße	**α-Fehler-Adjustierung**
1	$\frac{\alpha}{1} = 0.0500$
2	$\frac{\alpha}{1} = 0.0250$
3	$\frac{\alpha}{1} = 0.0166$

Ist das auf Grundlage der Daten ermittelte α gleich oder kleiner dem adjustierten α, so ist auch nach der α-Fehler-Adjustierung von einem signifikanten Effekt auszugehen.

8.6.2.6 Poweranalyse

Ergibt ein t-Test ein nicht signifikantes Ergebnis, besteht dennoch eine Wahrscheinlichkeit dafür, dass die endgültige Entscheidung für die H_0 nicht richtig ist. Diese Wahrscheinlichkeit wird als β-Fehler bezeichnet. Die Wahrscheinlichkeit für den β-Fehler sollte bei 10 % oder weniger liegen, damit die Nullhypothese zuverlässig angenommen werden kann (Rasch et al., 2006b, S. 79). Ist die β-Fehler-Wahrscheinlichkeit höher, kann weder die Alternativ- noch die Nullhypothese angenommen werden.

Mit Hilfe des β-Fehlers kann die Teststärke bzw. die sog. Power des t-Tests ermittelt werden. „Die Teststärke ist die Wahrscheinlichkeit, die H_1 anzunehmen, wenn sie auch in Wirklichkeit gilt" (Rasch et al., 2006b, S. 79). Die Teststärke wird durch 1-β ermittelt und stellt folglich die Gegenwahrscheinlichkeit zum β-Fehler dar. 1-β sollte nach Rasch und Kollegen (2006b, S. 80) mindestens 90 % betragen, Rost (2013, S. 247) bezeichnet eine Teststärke von 80 % als „angemessen". Mit der Teststärkebestimmung a priori kann im Vorfeld einer Studie ein Stichprobenumfang berechnet werden. Die Teststärkenbestimmung a posteriori ist dann notwendig, wenn unerwartet ein nicht signifikantes Ergebnis auftritt. Die Analyse der Power ist dabei nicht nur für einen t-Test sondern für fast alle statistischen Verfahren anwendbar, somit auch für Varianzanalysen (Rasch et al., 2006b, S. 109).

Die Abbildung 45 visualisiert die Verortung der Poweranalyse (Teststärke) a posteriori (grau hinterlegt).

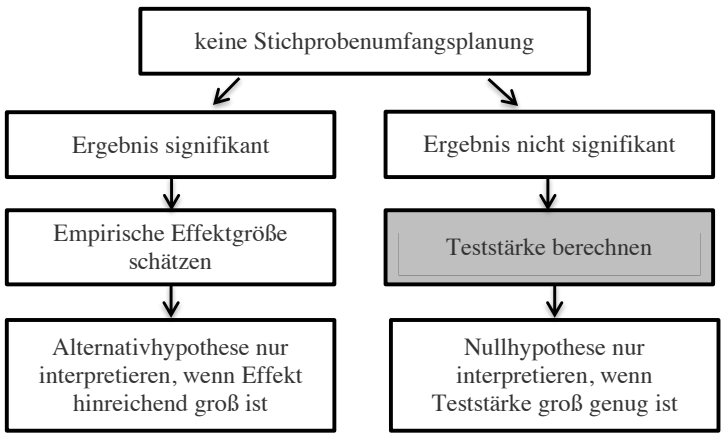


ABBILDUNG 45 ENTSCHEIDUNGSBAUM FÜR T-TESTS (RASCH ET AL., 2006B, S. 109)

8.6.2.7 (Multiple lineare) Regression

Die Regressionsanalyse gehört zu den sog. strukturprüfenden Verfahren und dient grundsätzlich dazu, einen möglichen funktionalen Zusammenhang zwischen zwei Variablen zu ermitteln (Rasch et al., 2006b, S. 146). Das x ist dabei die unabhängige Variable (Prädiktor) und das y die gesuchte abhängige Variable (Kriterium). Im Gegensatz zu einer Korrelation, die einen möglichen Zusammenhang zwischen zwei Variablen aufdeckt, gibt die Regressionsanalyse Aufschlüsse über die Art und Bezifferung dieses Zusammenhangs (Bühner & Ziegler, 2009, S. 586). „Linear" bedeutet, dass sich Änderungen in der abhängigen Variable und der unabhängigen Variable proportional zueinander verhalten. Der Quotient $\frac{\Delta y}{\Delta x_i}$ ist konstant. Die multiple

lineare Regression unterscheidet sich von der einfachen linearen Regression dahingehend, dass sie mehrere Prädiktoren nutzt, um die Ausprägung eines Kriterium zu bemessen und hervorzusagen (Rasch et al., 2006b, S. 146).

Ein Ergebnis der Regressionsanalyse ist die Regressionsgleichung. Mit n Prädiktoren und deren Koeffizienten β_i lautet sie allgemein:

$$\hat{y} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n$$

$\hat{y}$ ist die abhängige Variable und wird wegen ihres „hypothetischen" Werts mit einem Dach dargestellt, x stellt die unabhängige Variable dar.

Ein zentraler Wert der Regressionsanalyse ist das Bestimmtheitsmaß R^2, das die Bewertung der Güte des Regressionsmodels vornimmt. R^2 steht in direktem Zusammenhang mit dem multiplen Korrelationskoeffizienten und ist damit ein Maß der Varianzaufklärung. R^2 kann einen Wert zwischen 0 und 1 annehmen. Der Wert 1 zeigt an, dass sich das Regressionsmodell gänzlich aus den erklärenden Variablen herleiten lässt. Das korrigierte R^2 ermöglicht es, Regressionsmodelle mit unterschiedlichen Anzahlen an erklärenden Variablen miteinander zu vergleichen.

Voraussetzungen für die Regressionsanalyse sind (nach Rasch et al., 2006b, S. 163):

- Das Kriterium muss intervallskaliert und normalverteilt sein.
- Der Prädiktor muss entweder intervallskaliert und normalverteilt oder dichotom nominalskaliert sein.
- Die Einzelwerte verschiedener Versuchspersonen müssen voneinander unabhängig zustande gekommen sein.
- Der Zusammenhang der Variablen muss theoretisch linear sein (Durbin-Watson-Test).

Zudem muss die Varianz der Residuen des Modells über den gesamten Wertebereich der Prädiktoren homogen sein (Homoskedastizität). Entsprechend ist dies ein Kriterium, das *nach* der Ermittlung des Modells betrachtet wird und entsprechend zur Bewertung der Anwendbarkeit des Modells angewandt wird.

8.7 Operationalisierung der Wirksamkeit

Nachdem die Hypothesen, das Design und auch die statistischen Verfahren vorgestellt sind, erfolgt hier die Operationalisierung der kurz- und langfristigen Wirksamkeit der Fördermaßnahme (vgl. Hagenauer, 2010, S. 244 f.).

Für die kurz- und langfristige Wirksamkeit (Hypothesen 1–7) ist die Varianzanalyse mit Messwiederholung das zentrale statistische Verfahren (siehe Tabelle 34).

1. Kurzfristige Wirksamkeit der Maßnahme:

Die Schülerinnen und Schüler entwickeln sich im Prä-Post-Vergleich in wechselseitiger Abhängigkeit von der Gruppenzugehörigkeit und der Zeit (Interaktionseffekt). Dabei zeigt die Experimentalgruppe eine bessere Entwicklung als die Kontrollgruppe.

Folgende Abbildung 46 zeigt die Zusammensetzung der drei zentralen Effekte einer Varianzanalyse (Innersubjekteffekt, Zwischensubjekteffekt, Interaktionseffekt). Für die kurzfristige Wirksamkeit wird ein signifikanter Interaktionseffekt vorausgesetzt. Der Innersubjekteffekt und der Zwischensubjekteffekt sind zur Beantwortung der Hypothesen 1–7 *nicht* relevant und deswegen nicht fett markiert.

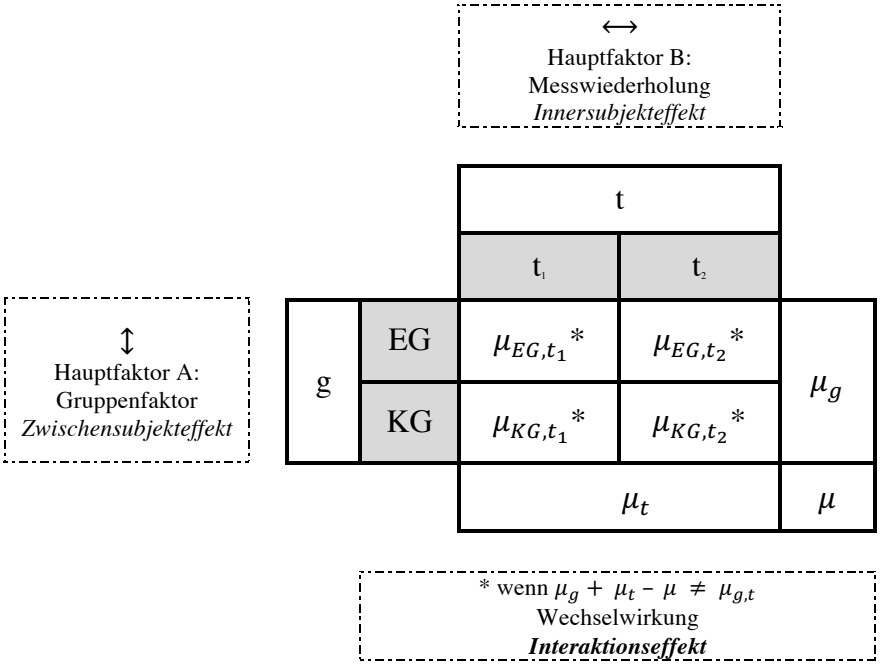


ABBILDUNG 46 STATISTISCHE FUNDIERUNG DER KURZFRISTIGEN WIRKSAMKEIT

2. Langfristige Wirksamkeit der Maßnahme:

a) Die Schülerinnen und Schüler entwickeln sich im Post-Follow-Up-I-Follow-Up-II-Vergleich signifikant in wechselseitiger Abhängigkeit von der Gruppenzugehörigkeit und der Zeit (Interaktionseffekt). Dabei zeigt die Experimentalgruppe im Post-Follow-Up-I- und im Follow-Up-I-Follow-Up-II-Vergleich keine schlechtere Entwicklung als die Kontrollgruppe.

b) Liegt eine kurzfristige, jedoch keine langfristige Wirksamkeit nach a) vor, spricht ein signifikanter Zwischensubjekteffekt (Haupteffekt) im Post-Follow-Up-I-Follow-Up-II-Vergleich zum Vorteil der Experimentalgruppe für einen persistenten Entwicklungsanschub und somit für die langfristige Wirksamkeit der Förderung.

Die Abbildung 47 zeigt die Zusammensetzung der drei zentralen Effekte der Varianzanalyse (Innersubjekteffekt, Zwischensubjekteffekt, Interaktionseffekt). Für die langfristige Wirksamkeit sind der Interaktionseffekt und der Zwischensubjekteffekt relevant. Der Innersubjekteffekt wird zur Beantwortung der Hypothesen 1–7 *nicht* benötigt und ist deswegen nicht fett markiert.

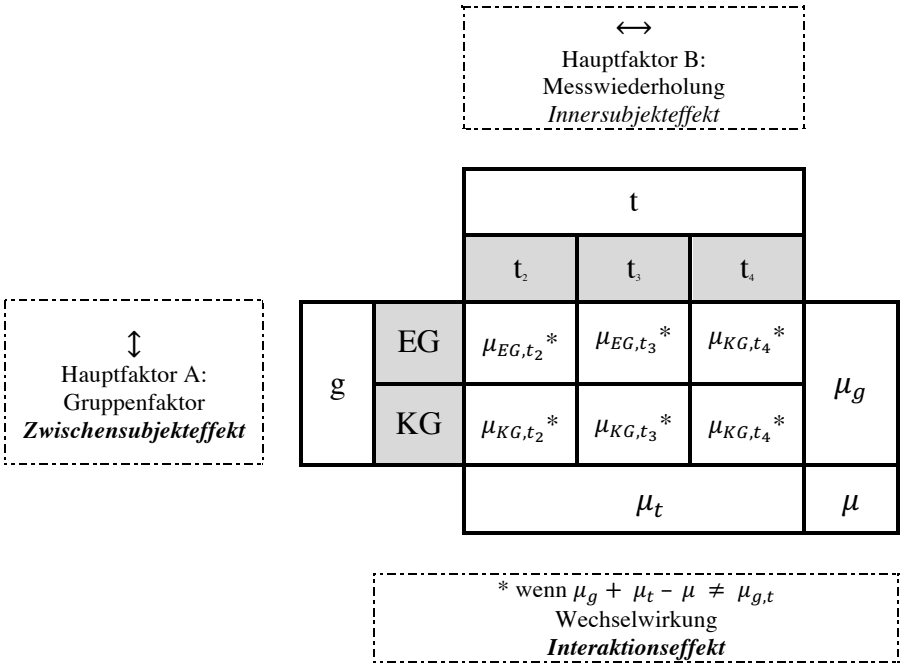


ABBILDUNG 47 STATISTISCHE FUNDIERUNG DER LANGFRISTIGEN WIRKSAMKEIT

8.8 Datenpflege und -aufbereitung

Die Daten von t_1 bis t_4 werden von möglichst denselben geschulten Studierenden erhoben. Bis auf den HRT 1–4 werden die Daten mit einem Fragebogenscanner eingelesen, um potenzielle Fehler in der Auswertung gering zu halten. Der HRT 1–4 wird nach dem Messzeitpunkt t_1 unter Anleitung der Projektleitung von den Studierenden auf Itemebene in eine vorgegebene SPSS-Maske eingetragen. In den folgenden drei Messzeitpunkten t_2– t_4 pflegen die Studierenden die Ergebnisse mit Hilfe des Leitfadens selbstständig in die SPSS-Maske ein, die von der Projektleitung vorbereitet wird. Die Studierenden führen zu allen Messzeitpunkten eine Liste zu den Drop-Outs, in der eingetragen wird, weshalb Schülerinnen und Schüler nicht anwesend waren und ob es bei bestimmten Schülerinnen und Schülern trotz Anwesenheit ggf. motivationale Schwierigkeiten gab (siehe Kapitel 9.2).

Im Anschluss an den letzten Messzeitpunkt erfolgt durch die Projektleitung ein Abgleich der eingetragenen HRT-Daten mit den archivierten HRT-Testbögen. Es werden über einen Zeitraum von vier Tagen ca. 80 % aller Bögen stichprobenweise nachgeprüft.

Auffällig ist in der Datenprüfung, dass bei einigen Schülerinnen und Schülern vereinzelt bestimmte Subtests nicht bearbeitet sind. Mögliche Gründe dafür können sein, dass die Schülerinnen und Schüler in der Erhebung den Start zur Bearbeitung des Subtests verpassen, dass sie abgelenkt sind, die Konzentration nachlässt oder sie eine mangelnde Motivation bezüglich einer spezifischen Grundrechenart aufweisen.

Sofern ein Wert fehlt, der Schüler oder die Schülerin jedoch in den anderen fünf Subtests *und* in den anderen drei Messzeitpunkten keine Unregelmäßigkeiten in der Bearbeitung des HRT

zeigt, wird eine Imputation durchgeführt, damit der Testbogen im Rahmen der Ergebnisauswertung weiterhin einbezogen werden kann und nicht wegen des einen fehlenden Subtests listenweise herausfällt (vgl. Kapitel 8.6).

Die Imputationstechnik sieht dabei vor, den Mittelwert der T-Tests der verbleibenden fünf Subtests des Messzeitpunkts zu bilden und daraus den – von den den Autoren des HRT 1–4 in Normtabellen bereitgestellten – „T-Wert der durchschnittlichen T-Werte“ zu ermitteln (mT1-Wert). Mit diesem T-Wert der durchschnittlichen T-Werte kann ein Rückschluss auf den vermeintlichen Rohwert in dem fehlenden Subtest getroffen werden. Durch diese Form der Imputation des fehlenden Werts verändert sich der mittlere T-Wert des *gesamten* Tests im Nachhinein *nicht*.

Diese Imputationstechnik wird dann genutzt, wenn ein oder maximal zwei Subtests fehlen. Dass zwei Subtests fehlen, ist über alle vier Messzeitpunkte insgesamt zehnmal der Fall (0,03 %). In einem einzigen Fall fehlen drei Subtests. Für diesen Test werden keine Werte imputiert.

Folgende Tabelle 47 gibt einen Überblick über die absolute Anzahl der imputierten Subtests bezogen auf den jeweiligen Messzeitpunkt.

TABELLE 47 ÜBERBLICK ÜBER IMPUTIERTE WERTE FÜR EINZELNE SUBTESTS IM HRT 1–4 (T_1–T_4)

Messzeitpunkt	Imputierter Subtest	Experimental-gruppe	Kontroll-gruppe	Gesamt-gruppe
Prä (t_1)	Addition	–	–	–
	Subtraktion	–	–	–
	Multiplikation	–	–	–
	Division	–	–	–
	Ergänzungsaufgaben	1	–	1
	Größer-Kleiner-Vergleich	1	1	2
Post (t_2)	Addition	–	–	–
	Subtraktion	–	1	1
	Multiplikation	–	–	–
	Division	2	1	3
	Ergänzungsaufgaben	2	2	4
	Größer-Kleiner-Vergleich	1	3	4
FollowUp-I (t_3)	Addition	–	–	–
	Subtraktion	–	–	–
	Multiplikation	2	0	2
	Division	8	3	11
	Ergänzungsaufgaben	7	2	9
	Größer-Kleiner-Vergleich	2	1	3
Follow-Up-II (t_4)	Addition	–	–	–
	Subtraktion	–	1	1
	Multiplikation	1	–	1
	Division	3	3	6
	Ergänzungsaufgaben	3	2	5
	Größer-Kleiner-Vergleich	–	3	3

Insgesamt zeigt sich, dass hauptsächlich in den Subtests Division, Ergänzungsaufgaben und Größer-Kleiner-Vergleich vereinzelt Subtests fehlen. Dabei steigt die Anzahl der fehlenden Subtests zeitlich im Laufe der Messzeitpunkte.

9 Ergebnisse

9.1 Beschreibung der Stichprobe

Im Folgenden wird die durch die Rekrutierung der Schulen gewonnene Stichprobe beschrieben. Die Tabelle 48 zeigt die Ergebnisse der Stichprobenbeschreibung im Detail. Dabei sind die dargestellten Merkmale nicht immer zum gleichen Messzeitpunkt erhoben, was der Index anzeigt. Die Altersangaben sind hier auf eine Stelle gerundet.

Insgesamt besteht die anfänglich rekrutierte Stichprobe aus 649 Schülerinnen und Schülern. Davon haben 13 Schülerinnen und Schüler keine elterliche Erlaubnis, an der Studie teilzunehmen – elf davon in der Experimentalgruppe und zwei in der Kontrollgruppe. Die Stichprobe reduziert sich deshalb auf 636 Schülerinnen und Schüler.

In der Gesamtstichprobe befinden sich 348 Schüler und 288 Schülerinnen. Die Experimentalgruppe besteht aus 260 Schülerinnen und Schülern und die Kontrollgruppe aus 376 Schülerinnen und Schülern. Die Anzahl an Schülern überwiegt insgesamt mit 54,72 %. Die jeweilige Geschlechterverteilung innerhalb der Experimental- und Kontrollgruppe unterscheidet sich nicht signifikant voneinander ($\chi^2 = 0.366, p_{\chi^2} = 0.545$). Das Durchschnittsalter beträgt 10 Jahre und 7 Monate ($\varnothing$ = 10,85 Jahre; Min: 9;2 Jahre; Max: 13;1 Jahre). Auch die Altersverteilung innerhalb der Experimental- und Kontrollgruppe unterscheidet sich nicht signifikant voneinander ($\chi^2 = 37.597, p_{\chi^2} = 0.308$).

In beiden Untersuchungsgruppen befinden sich Schülerinnen und Schüler mit Unterstützungsbedarfen. Die Fallzahlen sind für inferenzstatistische Analysen zu klein. Die prozentuale Anzahl ist in den beiden Gruppen vergleichbar.

Insgesamt verteilen sich die Schülerinnen und Schüler auf drei Oberschulen und vier Integrierte Gesamtschulen. Die Schülerinnen und Schüler der Experimentalgruppe besuchen davon zwei Oberschulen und eine Integrierte Gesamtschule, diejenigen der Kontrollgruppe entsprechend zwei Oberschulen und zwei Integrierte Gesamtschulen (siehe Kapitel 8.3). Die Verteilung der Schülerinnen und Schüler auf die beiden Schulformen innerhalb der Untersuchungsgruppen unterscheidet sich signifikant voneinander ($\chi^2 = 37.694, p_{\chi^2} = 0.000$). In der Kontrollgruppe besuchen demnach mehr Schülerinnen und Schüler eine Integrierte Gesamtschule (76,06 %) als in der Experimentalgruppe (52,69 %).

Die Experimentalgruppe besteht insgesamt aus elf Klassen mit einer durchschnittlichen Klassengröße von 23,6 Schülerinnen und Schülern. Die Kontrollgruppe besteht aus insgesamt 15 Klassen mit einer durchschnittlichen Schülerstärke von 25 Schülerinnen und Schülern pro Klasse. Die durchschnittliche Klassengröße in der Experimental- und Kontrollgruppe unterscheidet sich nicht signifikant voneinander ($p_t = 0.426$).

Die beiden Versuchsgruppen unterscheiden sich signifikant bezüglich der gesprochenen Sprache zu Hause ($\chi^2 = 12.940, p_{\chi^2} = 0.002$). In der Experimentalgruppe sprechen demnach weniger Schülerinnen und Schüler ausschließlich Deutsch zu Hause (51,54 %) als in der Kontrollgruppe (63,03 %). Ebenso unterscheiden sich die beiden Gruppen bezüglich des Geburtslands der Schülerinnen und Schüler ($\chi^2 = 8.936, p_{\chi^2} = 0.003$). In der Experimentalgruppe sind 81,54 %

der Schülerinnen und Schüler in Deutschland geboren, in der Kontrollgruppe 84,57 %. Einen weiteren Unterschied gibt es bei der Anzahl der Bücher, die sich zu Hause befinden ($\chi^2 = 32.312$, $p_{\chi^2} = 0.000$). In der Kontrollgruppe ist der Anteil an Schülerinnen und Schülern mit eher vielen Büchern zu Hause größer, in der Experimentalgruppe findet sich der Anteil derjenigen mit eher wenigen Büchern.

In der folgenden Tabelle 48 sind die Signifikanzen der Merkmale, die sich zwischen den beiden Untersuchungsgruppen signifikant voneinander unterscheiden, fett markiert.

TABELLE 48 SOZIODEMOGRAFISCHE BESCHREIBUNG DER STICHPROBE INKLUSIVE FEHLENDER WERTE UND χ^2-TESTS

Merkmal	Ausprägung	Experimental-gruppe	Kontroll-gruppe	Gesamt-gruppe	χ^2	df	p_{χ^2}
Geschlecht[1]	männlich weiblich fehlende Angabe *gesamt*	146 (56,15 %) 114 (43,85 %) 0 (0,00 %) *260 (100 %)*	202 (53,72 %) 174 (46,28 %) 0 (0,00 %) *376 (100 %)*	348 (54,72 %) 288 (45,28 %) 0 (0,00 %) *636 (100 %)*	0.366	1	0.545
Alter[1]	9,0–9,9 10,0–10,9 11,0–11,9 12,0–12,9 13,0–13,9 fehlende Angabe *Ø Alter*	3 (1,15 %) 151 (58,08 %) 93 (35,77 %) 12 (4,62 %) 1 (0,038 %) 0 (0,00 %) *10,89 Jahre*	5 (1,33 %) 233 (61,97 %) 116 (30,85 %) 17 (4,52 %) 2 (0,53 %) 3 (0,80 %) *10,81 Jahre*	8 (1,26 %) 384 (60,38 %) 209 (32,86 %) 29 (4,56 %) 3 (0,47 %) 3 (0,47 %) *10,85 Jahre*	37.597	34	0.308
Sonderpäd. Unterstützungs-bedarfe[1, 4]	Lernen Emotional-soziale Entwicklung Geistige Entwicklung Körperlich-motorische Entwicklung Hören Kein Unterstützungsbedarf fehlende Angabe	9 (3,46 %) 3 (1,15 %) 0 (0,00 %) 1 (0,38 %) 1 (0,38 %) 218 (83,85 %) 28 (10,77 %)	21 (5,59 %) 7 (1,86 %) 2 (0,53 %) 0 (0,00 %) 1 (0,27 %) 317 (84,31 %) 28 (7,45 %)	30 (4,72 %) 10 (1,57 %) 2 (0,31 %) 1 (0,16 %) 2 (0,31 %) 535 (84,12 %) *56 (8,81 %)*	–	–	–
Besuchte Schulform[1]	OBS IGS	123 (47,31 %) 137 (52,69 %)	90 (23,94 %) 286 (76,06 %)	213 (33,49 %) 423 (66,51 %)	37.694	1	**0.000**
Klassengröße[1]	Anzahl der Schülerinnen und Schüler pro Klasse	Ø 23,6	Ø 25,0		–	–	0.426 *(hier: p_t)*
Sprache zu Hause[2]	Deutsch Deutsch und eine weitere Sprache nicht Deutsch fehlende Angabe	134 (51,54 %) 90 (34,62 %) 11 (4,23 %) 25 (9,62 %)	237 (63,03 %) 84 (22,34 %) 10 (2,66 %) 45 (11,97 %)	371 (58,33 %) 174 (27,36 %) 21 (3,30 %) 70 (11,01 %)	12.940	2	**0.002**

Merkmal	Ausprägung	Experimental-gruppe	Kontroll-gruppe	Gesamt-gruppe	χ^2	df	p_{χ^2}
Geburtsland ist Deutschland$_2$	ja nein fehlende Angabe	212 (81,54 %) 23 (8,85 %) 25 (9,62 %)	318 (84,57%) 12 (3,19 %) 46 (12,23 %)	530 (83,33 %) 35 (5,50 %) 71 (11,16 %)	8.936	1	**0.003**
Anzahl Bücher zu Hause$_2$	Keine oder nur sehr wenige Genug, um ein Regalbrett zu füllen Genug, um ein Regal zu füllen Genug, um drei Regale zu füllen Viel mehr als drei Regale fehlende Angabe	24 (9,23 %) 61 (23,46 %) 70 (26,92 %) 45 (17,31 %) 30 (11,54 %) 30 (11,54 %)	20 (5,32 %) 52 (13,83 %) 82 (21,81 %) 65 (17,29 %) 105 (27,93 %) 52 (18,83 %)	44 (6,92 %) 113 (17,77 %) 152 (23,90 %) 110 (17,30 %) 135 (21,23 %) 82 (12,89 %)	32.312	4	**0.000**

Anmerkungen:
$_1$ = erhoben zum Prä-Messzeitpunkt
$_2$ = erhoben zum Post-Messzeitpunkt
$_3$ = erhoben zum Follow-Up-I-Messzeitpunkt
$_4$ = erhoben zum Follow-Up-II-Messzeitpunkt

Neben den soziodemografischen Angaben wird die Stichprobe bzgl. ihrer Kompetenzen im Bereich der Kulturtechniken und ihres Sozial- und Lernverhaltens vorgestellt. Der Vergleich dieser Entwicklungsbereiche ist für die weitere Ergebnisdarstellung von Bedeutung, da überprüft wird, inwiefern die Experimental- und Kontrollgruppe eine vergleichbare Ausganssituation im Rahmen der Studie vorweisen und ob die Randomisierung dementsprechend gelungen ist.

Die mathematischen Basiskompetenzen werden mit der Subskala Rechenoperationen Heidelberger-Rechentest 1–4 erhoben. Zur Beschreibung der Stichprobe wird der standardisierte mT1-Wert herangezogen, der den durchschnittlichen T-Wert über die sechs Subtests der Subskala Rechenoperationen angibt. Die Experimentalgruppe erreicht zum Prä-Messzeitpunkt einen durchschnittlichen T-Wert von 43,96 (SD: 7,65), die Kontrollgruppe einen durchschnittlichen T-Wert von 43,48 (SD: 7,72). Der Kolmogorov-Smirnov-Test (im Index abgekürzt mit K-S) ergibt, dass die Werte sowohl in der Kontrollgruppe ($p_{K\text{-}S} = 0.00$) als auch in der Experimentalgruppe ($p_{K\text{-}S} = 0.046$) nicht normalverteilt sind. Der Mann-Whitney-U-Test weist auf keine signifikanten Unterschiede in den mathematischen Basiskompetenzen zwischen der Experimental- und Kontrollgruppe hin ($p_U = 0.254$). Dies wird durch den t-Test bei unabhängigen Stichproben bestätigt ($p_t = 0.445$).

Bezüglich der kognitiven Fähigkeiten weist die Experimentalgruppe einen durchschnittlichen IQ von 95,60 (SD: 14,53) und die Kontrollgruppe von 96,46 (SD: 14,78) auf. Der durchschnittliche IQ der gesamten Stichprobe liegt bei 96,10 (SD: 14,66). Sowohl die IQ-Werte in der Experimentalgruppe als auch in der Kontrollgruppe weisen eine Normalverteilung auf; es liegt kein signifikanter Unterschied zwischen den IQ-Werten zwischen den Gruppen vor ($p_t = 0.482$).

Bezüglich der Lesekompetenz zeigt sowohl die Experimental- als auch die Kontrollgruppe durchschnittliche Lesequotienten ($LQ_{EG} = 84{,}94$, SD: 15,38; $LQ_{KG} = 87{,}29$, SD: 17,10). Die Werte sind in der Kontrollgruppe jedoch nicht normalverteilt ($p_{K\text{-}S} = 0.001$), sodass neben dem t-test für unabhängige Stichproben auch der Mann-Whitney-U-Test zur Absicherung eingesetzt wird. Beide Tests zeigen keine signifikanten Unterschiede in der Lesekompetenz der Untersuchungsgruppen ($p_t = 0.080$, $p_U = 0.152$).

Bezüglich des Sozial- und Lernverhaltens lässt sich zusammenfassen, dass keine Subskala der LSL in der Kontroll- oder in der Experimentalgruppe normalverteilt ist ($p_{K\text{-}S} < 0.05$). Analysen auf Unterschiede in der Ausprägung zwischen den Gruppen mit dem Mann-Whitney-U-Test und dem t-Test für abhängige Stichproben ergeben keine signifikanten Unterschiede für die Skalen Kooperation ($p_U = 0.555$, $p_t = 0.275$), Selbstwahrnehmung ($p_U = 0.296$, $p_t = 0.351$), Selbstkontrolle ($p_U = 0.835$, $p_t = 0.914$), Einfühlungsvermögen ($p_U = 0.808$, $p_t = 0.595$), Angemessene Selbstbehauptung ($p_U = 0.627$, $p_t = 0.606$), Sozialkontakt ($p_U = 0.511$, $p_t = 0.278$), Anstrengungsbereitschaft ($p_U = 0.735$, $p_t = 0.538$), Konzentration ($p_U = 0.196$, $p_t = 0.241$), beim Lernen ($p_U = 0.465$, $p_t = 0.241$) und Sorgfalt beim Lernen ($p_U = 0.560$, $p_t = 0.718$).

TABELLE 49 TESTUNG DER UNTERSUCHUNGSGRUPPEN AUF VORTESTUNTERSCHIEDE

Merkmal		Ausprägung	Experimentalgruppe	Kontrollgruppe	Gesamtgruppe	df	p_t	p_u
Mathematische Basiskompetenz[1]		T-Wert < 40 T-Wert 41–60 T-Wert > 60 T-Wert fehlend *Ø T-Wert*	75 (28,85 %) 168 (64,62 %) 8 (3,08 %) 9 (3,46 %) *43,96*	116 (30,85 %) 232 (61,70 %) 11 (2,93 %) 17 (4,52 %) *43,48*	191 (30,03 %) 400 (62,89 %) 19 (2,99 %) 26 (4,09 %) *43,72*	608	0.445	0.254
Kognitive Fähigkeiten[1]		IQ < 85 IQ 85–100 IQ 101–115 IQ > 115 IQ fehlend *Ø IQ*	52 (20,00 %) 100 (38,46 %) 78 (30,00 %) 22 (8,46 %) 8 (3,08 %) *95,60*	66 (17,55 %) 142 (37,77 %) 105 (27,93 %) 33 (8,78 %) 30 (7,98 %) *96,46*	118 (18,55 %) 242 (38,05 %) 183 (28,77 %) 55 (8,65 %) 83 (13,05 %) *96,10*	546.16	0.482	–
Lesekompetenz[1]		LQ < 90 LQ 90–109 LQ ≥ 110 fehlend *Ø LQ*	156 (60,00 %) 78 (30,00 %) 15 (5,77 %) 11 (4,23 %) *84,94*	198 (52,66 %) 112 (29,79 %) 34 (9,04 %) 32 (8,51 %) *87,29*	354 (55,66 %) 190 (29,87 %) 50 (7,86 %) 43 (6,76 %) *86,12*	564.00	0.080	0.152
Sozialverhalten[2]	Kooperation	T-Wert < 40 T-Wert 40–60 T-Wert > 60 T-Wert fehlend *Ø T-Wert*	30 (11,54 %) 199 (76,54 %) 0 (0,00 %) 31 (11,92 %) *49,91*	31 (8,24 %) 149 (39,63 %) 0 (0,00 %) 196 (52,13 %) *49,01*	61 (9,59 %) 348 (54,72 %) 0 (0,00 %) 227 (35,69 %) *49,46*	340,13	0.275	0.555
	Selbst-wahrnehmung	T-Wert < 40 T-Wert 40–60 T-Wert > 60 T-Wert fehlend *Ø T-Wert*	26 (10,00 %) 168 (64,62 %) 35 (13,46 %) 31 (11,92 %) *50,71*	19 (5,05 %) 124 (32,98 %) 37 (9,84 %) 196 (52,13 %) *51,52*	45 (7,08 %) 292 (45,91 %) 72 (11,32 %) 227 (35,69 %) *50,62*	407	0.351	0.296

Merkmal		Ausprägung	Experimentalgruppe	Kontrollgruppe	Gesamtgruppe	df	p_t	p_u
Sozialverhalten$_2$	Selbstkontrolle	T-Wert < 40 T-Wert 40–60 T-Wert > 60 T-Wert fehlend *Ø T-Wert*	25 (9,62 %) 173 (66,54 %) 31 (11,92 %) 31 (11,92 %) *50,42*	21 (5,59 %) 134 (38,03 %) 25 (6,65 %) 196 (52,13 %) *50,52*	46 (7,23 %) 307 (48,27 %) 56 (8,81 %) 227 (35,69 %) *50,47*	407	0.914	0.835
	Einfühlungsvermögen	T-Wert < 40 T-Wert 40–60 T-Wert > 60 T-Wert fehlend *Ø T-Wert*	20 (7,69 %) 209 (80,38 %) 0 (0,00 %) 31 (11,92 %) *49,61*	37 (9,84 %) 143 (38,03 %) 0 (0,00 %) 196 (52,13 %) *49,12*	57 (8,96 %) 352 (55,35 %) 0 (0,00 %) 227 (35,69 %) *49,37*	407	0.595	0.808
	Angemessene Selbstbehauptung	T-Wert < 40 T-Wert 40–60 T-Wert > 60 T-Wert fehlend *Ø T-Wert*	23 (8,85 %) 206 (79,23 %) 0 (0,00 %) 31 (11,92 %) *49,66*	18 (4,79 %) 162 (43,09 %) 0 (0,00 %) 196 (52,13 %) *51,09*	41 (6,45 %) 368 (57,86 %) 0 (0,00 %) 227 (35,69 %) *50,34*	407	0.606	0.627
	Sozialkontakt	T-Wert < 40 T-Wert 40–60 T-Wert > 60 T-Wert fehlend *Ø T-Wert*	22 (8,46 %) 180 (69,23 %) 27 (10,38 %) 31 (11,92 %) *50,18*	20 (5,32 %) 132 (35,11 %) 28 (7,45 %) 196 (52,13 %) *50,19*	42 (6,60 %) 312 (49,06 %) 55 (8,65 %) 227 (35,69 %) *50,19*	407	0.278	0.511
Lernverhalten$_2$	Anstrengungsbereitschaft	T-Wert < 40 T-Wert 40–60 T-Wert > 60 T-Wert fehlend *Ø T-Wert*	18 (6,92 %) 188 (72,31 %) 23 (8,85 %) 31 (11,92 %) *50,73*	25 (6,65 %) 130 (34,57 %) 25 (6,65 %) 196 (52,13 %) *50,19*	43 (6,76 %) 318 (50,00 %) 47 (7,39 %) 227 (35,69 %) *50,46*	406	0.538	0.735
	Konzentration	T-Wert < 40 T-Wert 40–60 T-Wert > 60 T-Wert fehlend *Ø T-Wert*	19 (7,31 %) 187 (71,92 %) 23 (8,85 %) 31 (11,92 %) *50,04*	16 (4,26 %) 135 (35,90 %) 29 (7,71 %) 196 (52,13 %) *50,91*	35 (5,50 %) 322 (50,63 %) 52 (8,18 %) 227 (35,69 %) *50,48*	407	0.241	0.196

Merkmal		Ausprägung	Experimentalgruppe	Kontrollgruppe	Gesamtgruppe	df	p_t	p_u
Lern- verhalten$_2$	Selbstständig-keit	T-Wert < 40 T-Wert 40–60 T-Wert > 60 T-Wert fehlend *Ø T-Wert*	46 (17,69 %) 183 (70,38 %) 0 (0,00 %) 31 (11,92 %) *47,53*	45 (11,97 %) 135 (35,90 %) 0 (0.00 %) 196 (52,13 %) *46,42*	91 (14,31 %) 318 (50,00 %) 0 (0,00 %) 227 (35,69 %) *46,98*	407	0.312	0.465
	Sorgfalt beim Lernen	T-Wert < 40 T-Wert 40–60 T-Wert > 60 T-Wert fehlend *Ø T-Wert*	54 (20,77 %) 175 (67,31 %) 0 (0,00 %) 31 (11,92 %) *46,07*	27 (7,18 %) 153 (40,69 %) 0 (0,00 %) 196 (52,13 %) *46,29*	81 (12,74 %) 328 (51,57 %) 0 (0,00 %) 227 (35,69 %) *46,18*	402,1	0.718	0.560

Anmerkungen: 1 = erhoben zum Prä-Messzeitpunkt; 2 = erhoben zum Post-Messzeitpunkt; 3 = erhoben zum Follow-Up-I-Messzeitpunkt; 4 = erhoben zum Follow-Up-II-Messzeitpunkt; p_t = t-test bei unabhängigen Stichproben; p_u = Mann-Whitney-U-Test

9.2 Darstellung des Drop-Outs

DROP-OUTS IN DER SCHÜLERERHEBUNG

Folgende Tabelle 50 gibt Aufschluss über die Drop-Out-Quoten über die vier Messzeitpunkte. Aufgeführt sind diejenigen Schülerinnen und Schüler, die an dem jeweiligen Messzeitpunkt nicht anwesend sind. Zudem sind die Variablen angegeben, die an dem Messzeitpunkt erhoben werden und somit bei den Schülerinnen und Schülern bei dem Messzeitpunkt fehlen (vgl. Tabelle 33). Die Angaben in Tabelle 50 beziehen sich auf die Instrumente, die von den Schülerinnen und Schülern bearbeitet werden.

Die Drop-Outs werden von den Studierenden zu jedem Messzeitpunkt dokumentiert. Die angegebenen Zahlen stellen die absoluten Häufigkeiten der Drop-Outs dar, jeweils bezogen auf die angegebene Untersuchungsgruppe und dem spezifischen Messzeitpunkt. Einen Hinweis auf den prozentualen Anteil bezüglich der Stichprobe zu t_i ermöglicht der Wert mit *.

Nicht einbezogen sind diejenigen Schülerinnen und Schüler, die keine Erlaubnis zur Teilnahme an der Studie haben ($n_{EG} = 11$; $n_{KG} = 2$).

TABELLE 50 AUFSCHLÜSSELUNG FEHLENDER SCHÜLERINNEN UND SCHÜLER (DROP-OUTS) ZU DEN VIER MESSZEITPUNKTEN PRÄ, POST, FOLLOW-UP-I UND FOLLOW-UP-II

Messzeitpunkt	Betrifft die Variablen	Grund	Experimental-gruppe	Kontroll-gruppe	Gesamt-gruppe
Prä (t_1)	$\overrightarrow{P}$ $\overrightarrow{Q}$ $\overrightarrow{R}$	Krankheit Verzug andere *gesamt*	5 – 1 *6 (2,3 %*)*	11 – 4 *15 (4,0 %*)*	16 – 5 21 (3,3 %*)
Post (t_2)	$\overrightarrow{P}$ $\overrightarrow{U}$ $\overrightarrow{V}$ $\overrightarrow{W}$	Krankheit Verzug andere *gesamt*	14 – 4 *18 (6,9 %*)*	15 – 4 *19 (5,1 %*)*	29 – 8 37 (5,8 %*)
Follow-Up-I (t_3)	$\overrightarrow{P}$	Krankheit Verzug andere *gesamt*	11 3 10 *24 (9,2 %*)*	27 3 7 *37 (9,8 %*)*	38 6 17 61 (9,6 %*)
Follow-Up-II (t_4)	$\overrightarrow{P}$	Krankheit Verzug andere *gesamt*	6 5 11 *22 (8,5 %*)*	27 8 8 *43 (11,4 %*)*	33 13 19 65 (10,2 %*)

Insgesamt ist zu erkennen, dass die Drop-Out-Quoten bezüglich der Schülerinstrumente zwischen der Experimental- und Kontrollgruppe zu den vier Messzeitpunkten *nicht* massiv voneinander abweichen. Die Drop-Out-Quoten steigen von Messzeitpunkt zu Messzeitpunkt. Am Ende des fünften Schuljahres (t_4) beträgt die Drop-Out-Quote in Bezug auf die ursprünglich vorhandene Stichprobe insgesamt 10,2 %.

DROP-OUTS IN DER LEHRERERHEBUNG

Die Lehrkräfte werden im Rahmen der Studie im Post-Messzeitpunkt darum gebeten, eine Einschätzung des Lern- und Sozialverhaltens der Schülerinnen und Schüler mit der Lehrer-einschätzliste für das Sozial- und Lernverhalten auf Grundlage ihrer Beobachtungen im Zeitraum der vergangenen vier Wochen zu tätigen und zudem eine Auskunft über vorliegende sonderpädagogische Unterstützungsbedarfe zu geben. Auffällig ist dabei die hohe Drop-Out-Quote bzgl. der Lehrereinschätzliste zum Post-Messzeitpunkt, wie in Tabelle 49 deutlich wird. Sieben Lehrkräfte in der Kontrollgruppe füllen die Lehrereinschätzlisten im Post-Messzeitpunkt nicht aus, sodass dort insgesamt 52,13 % der Werte fehlen. Die Fallzahl kommt durch eine Lehrkraft zustande, die zu Beginn des Schuljahres spontan die Klasse wechselte und deswegen keine Einschätzung abgeben konnte. Zudem fühlen sich sechs Lehrkräfte einer Schule nicht in der Lage, das Lern- und Sozialverhalten der Schülerinnen und Schüler nach vier Wochen einschätzen zu können, sodass von insgesamt sieben Klassen keine LSL-Daten vorliegen. Bezüglich der Angaben zu den Unterstützungsbedarfen liegen abschließend die Informationen von 91,19 % der Schülerinnen und Schüler vor (siehe Tabelle 48).

Das folgende Kapitel 9.3 umfasst die Darstellung der Ergebnisse des ersten Hypothesenblocks.

9.3 Hypothesenblock I

Der erste Hypothesenblock (Hypothese 1a und 1b) beinhaltet Annahmen zur kurz- und langfristigen Wirksamkeit der Maßnahme bezogen auf die Gesamtgruppe. Die kurzfristige Wirksamkeit wird durch einen Prä-Post-Vergleich und die langfristige Wirksamkeit durch einen Post-Follow-Up-I-Follow-Up-II-Vergleich vorgenommen. Die inferenzstatistischen Analysen erfolgen durch zweifaktorielle Varianzanalysen mit Messwiederholung.

Die Überprüfung auf Normalverteilung der Rohwerte der Subskala Rechenoperationen des HRT 1–4 mittels des Kolmogorov-Smirnov-Anpassungstests zeigt, dass alle Messzeitpunkte in der Experimentalgruppe und – bis auf den letzten Follow-Up-II-Messzeitpunkt ($p_{K\text{-}S_FU\text{-}II} = 0.045$) – auch die Messzeitpunkte in der Kontrollgruppe normalverteilt sind (für Details siehe Anhang 5).

ERGEBNISSE ZUR HYPOTHESE 1A

Die Hypothese 1a nimmt an, dass die Schülerinnen und Schüler der Experimentalgruppe im Prä-Post-Vergleich eine bessere Entwicklung in den mathematischen Basiskompetenzen (gemessen durch die Subskala Rechenoperationen des HRT 1–4) zeigen als die Schülerinnen und Schüler der Kontrollgruppe.

Abbildung 48 zeigt, dass die Schülerinnen und Schüler der Experimentalgruppe in der Prä-Testung im Durchschnitt 131,92 Aufgaben richtig lösen, in der Kontrollgruppe sind es 129,66 Aufgaben. Die Experimentalgruppe löst in der Posttestung durchschnittlich 11,94 richtige Aufgaben mehr, die Kontrollgruppe 7,16.

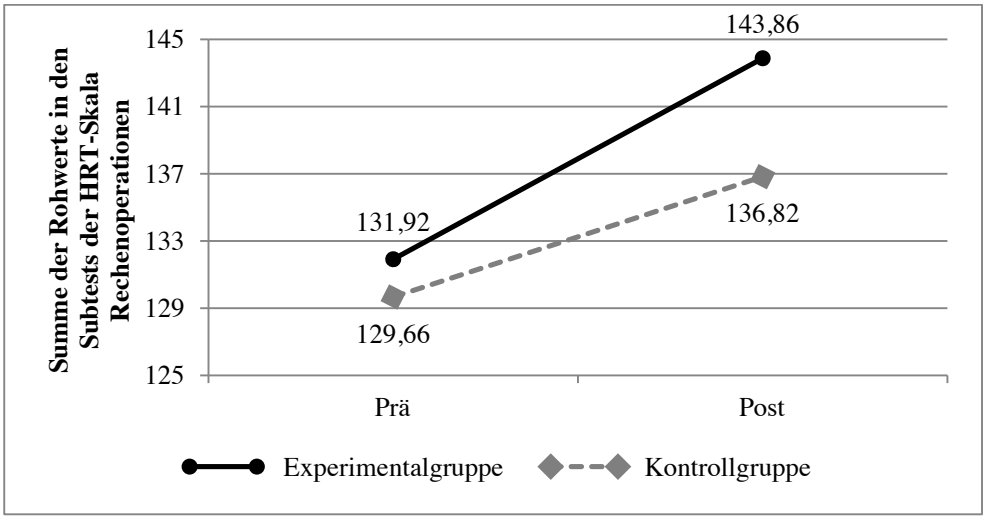


ABBILDUNG 48 PRÄ-POST-ENTWICKLUNG DER GESAMTGRUPPE

Da nur zwei Messzeitpunkte einbezogen werden, sind keine Angaben zur Sphärizität notwendig. Bezüglich der drei zentralen Effekte, die die zweifaktorielle Varianzanalyse ergibt, zeigt sich Folgendes:

Der Haupteffekt „Messwiederholung" (Innersubjekteffekt) zeigt an, dass sich die Mittelwerte der Rohwerte beider Gruppen *zusammen* zwischen den beiden Messzeitpunkten signifikant unterscheiden ($F(1) = 254.331$; $p = 0.000$; $\eta^2 = 0.307$). Beide Gruppen *zusammen* weisen also eine signifikante Verbesserung ihrer mathematischen Basiskompetenzen zwischen der Prä- und Post-Erhebung auf.

Der zweite Haupteffekt „Gruppenzugehörigkeit" (Zwischensubjekteffekt) ist knapp nicht signifikant ($F(1) = 3.595$; $p = 0.058$; $\eta^2 = 0.006$). Die Mittelwerte *zwischen* den beiden Gruppen unterscheiden sich in den beiden Messzeitpunkten nicht signifikant voneinander.

Der Interaktionseffekt ist hingegen signifikant ($F(1) = 15.796$; $p = 0.000$; $\eta^2 = 0.027$). Die Entwicklung der mathematischen Basiskompetenzen von der Prä- zur Posttestung unterscheidet sich mit einem kleinen Effekt ($d = 0.333$) signifikant in wechselseitiger Abhängigkeit von der Gruppenzugehörigkeit und dem Messzeitpunkt. Die Schülerinnen und Schüler der Experimentalgruppe entwickeln sich im Prä-Post-Vergleich signifikant besser als diejenigen der Kontrollgruppe.

Es liegt eine – nach der hier zugrunde gelegten Definition – kurzfristige Wirksamkeit der Fördermaßnahme vor. Die zentralen Werte zu den durchschnittlich erreichten Rohwerten in den beiden Messzeitpunkten finden sich deskriptiv in Tabelle 51.

TABELLE 51 ZENTRALE KENNWERTE ZUR HYPOTHESE 1A – KURZFRISTIGE WIRKSAMKEIT IN DER GESAMTGRUPPE

	Experimentalgruppe			**Kontrollgruppe**				
t	N	M	SD	N	M	SD	*p*	η^2 (*d*)
Prä	233	131,92	28,75	344	129,66	28,82	0.000	0.027 (0.333)
Post		143,86	31,37		136,82	30,10		

Anmerkungen:
Zweifaktorielle Varianzanalyse mit Messwiederholung auf einem Faktor (Faktor 1: Messwiederholung, Faktor 2: Gruppe); Ergebnisse des Interaktionseffekts (*p*, η^2, *d*)

Ergebnisse zur Hypothese 1b

Die Hypothese 1b nimmt an, dass die Schülerinnen und Schüler der Experimentalgruppe im Zeitraum zwischen der Post-, Follow-Up-I- und Follow-Up-II Erhebung eine bessere Entwicklung in den mathematischen Basiskompetenzen (gemessen durch die Subskala Rechenoperationen des HRT 1–4) zeigen als die Schülerinnen und Schüler der Kontrollgruppe. Da die Varianzanalyse grundsätzlich nur die Fälle einschließt, für die zu jedem Messzeitpunkt Werte vorhanden sind, können sich die Post-Werte hier im Vergleich zur Hypothese 1a unterscheiden.

Die Voraussetzung der Sphärizität (Homogenität der Varianzen der paarweisen Differenzen zwischen den Messzeitpunkten) ist nicht gegeben ($p_{\text{Mauchly}} = 0.006$), sodass im Folgenden der Huynh-Feldt-Korrekturwert ($p = 0.985$) genutzt wird.

Die Schülerinnen und Schüler der Experimentalgruppe lösen im Follow-Up-I-Messzeitpunkt durchschnittlich 3,59 Aufgaben weniger als im Post-Messzeitpunkt. Im Follow-Up-II hingegen steigen die Leistungen um weitere 10,27 richtig gelöste Aufgaben wieder an. Die Schülerinnen und Schüler der Kontrollgruppe verbessern sich zum Follow-Up-I-Messzeitpunkt durchschnittlich um 4,98 richtig gelöste Aufgaben und zum Follow-Up-II-Messzeitpunkt erneut um 4,81 Aufgaben (Abbildung 49).

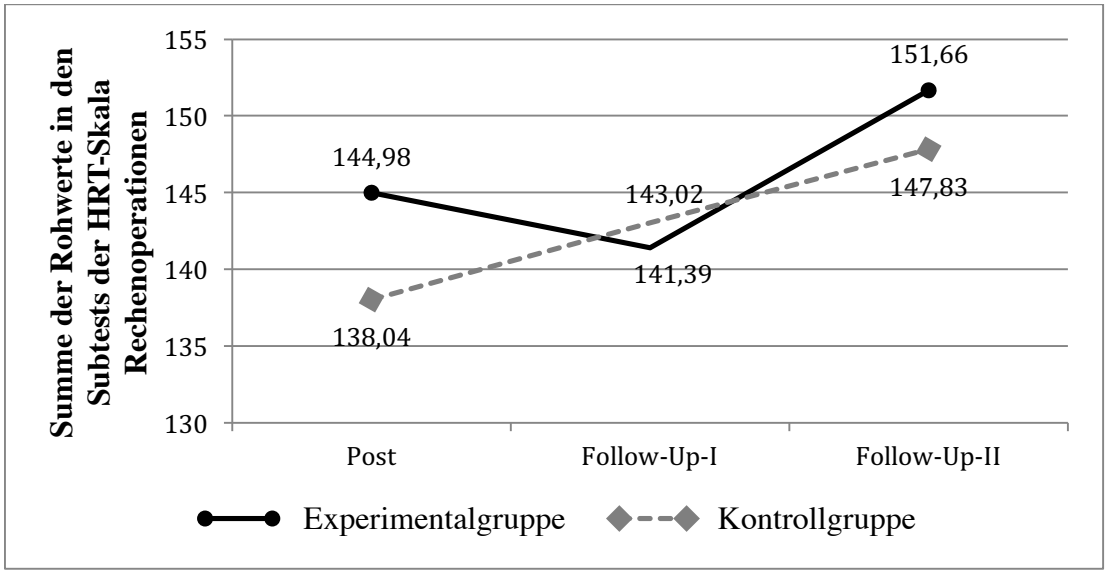


Abbildung 49 Post-Follow-Up-I-Follow-Up-II-Entwicklung der Gesamtgruppe

Bezüglich der drei zentralen Effekte, die die zweifaktorielle Varianzanalyse ergibt, zeigt sich Folgendes:

Der Haupteffekt „Messwiederholung“ (Innersubjekteffekt) zeigt an, dass sich die Mittelwerte der Rohwerte beider Gruppen *zusammen* zwischen den drei Messzeitpunkten signifikant unterscheiden ($F(1.970) = 73{,}383$; $p = 0.000$; $\eta^2 = 0.134$). Für beide Gruppen *zusammen* zeigen sich folglich signifikante Unterschiede in den mathematischen Basiskompetenzen über den Zeitraum von der Post-, der Follow-Up-I- zur Follow-Up-II-Erhebung, die Messzeitpunktbedingung hat einen Einfluss auf die Entwicklung der mathematischen Basiskompetenzen der Schülerinnen und Schüler.

Der zweite Haupteffekt „Gruppenzugehörigkeit“ (Zwischensubjekteffekt) ist nicht signifikant ($F(1) = 1.266$; p = 0.261; $\eta^2 = 0.003$). Die Mittelwerte der beiden Gruppen innerhalb der drei Messzeitpunkte unterscheiden sich nicht signifikant voneinander.

Der Interaktionseffekt hingegen ist signifikant (F(2) = 16.508; $p = 0.000$; $\eta^2 = 0.034$). Die langfristige Entwicklung der mathematischen Basiskompetenzen von der Post-, über die Follow-Up-I- zur Follow-Up-II-Erhebung unterscheidet sich signifikant in Abhängigkeit von der Gruppenzugehörigkeit.

Da die visuelle Analyse (siehe Abbildung 49) einen Rückgang der Rohwerte in der Experimentalgruppe vom Post- zum Follow-Up-I-Messzeitpunkt anzeigt, werden Detailanalysen vorgenommen, um den Interaktionseffekt differenziert aufklären zu können. Mittels t-Tests bei unabhängigen Stichproben zeigt sich ein signifikanter Rückgang der Leistungen in der Experimentalgruppe zwischen dem Post- und Follow-Up-I-Messzeitpunkt im Vergleich zur Kontrollgruppe ($p_t = 0.000$; $d_{korr} = –0.281$). Doch auch die weitere Entwicklung vom Follow-Up-I- zum Follow-Up-II-Messzeitpunkt unterscheidet sich signifikant zwischen den beiden Gruppen, nun zum Vorteil der Experimentalgruppe ($p_t = 0.000$; $d_{korr} = 0.179$). Betrachtet man die Entwicklung vom Post- zum Follow-Up-II-Messzeitpunkt, zeigt die Experimentalgruppe mit einem unbedeutenden Effekt einen leichten Rückgang in ihrer mathematischen Entwicklung im Vergleich zur Kontrollgruppe ($p_t = 0.048$; $d = –0.103$). Die mathematischen Basiskompetenzen innerhalb des Follow-Up-II unterscheiden sich nicht signifikant voneinander ($p_t = 0.076$, $d = 0.125$). Die Ergebnisse der α-Abwärtsadjustierung zeigen, dass die errechneten Signifikanzwerte der Detailanalyse auch nach der α-Abwärtsadjustierung weiterhin als signifikant angesehen werden können (siehe Tabelle 65).

Es liegt – nach der hier zugrunde gelegten Definition – keine langfristige Wirksamkeit der Fördermaßnahme vor. Die zentralen Werte der Varianzanalyse sind in Tabelle 52 zusammengefasst und durch die Ergebnisse der Detailanalysen erweitert.

TABELLE 52 ZENTRALE KENNWERTE ZUR HYPOTHESE 1B – LANGFRISTIGE WIRKSAMKEIT IN DER GESAMTGRUPPE

	Experimentalgruppe			**Kontrollgruppe**						
t	N	M	SD	N	M	SD	p	η^2 (d)	d_{korr*}	d_{korr**}
Post	190	144,98	31,74	287	138,04	29,65	0.000	0.034 (0.375)	–	–
FU-I		141,39	30,79		143,02	29,54			–0.281	–
FU-II		151,66	30,62		147,83	30,68			–0.103	0.179

Anmerkungen:
Zweifaktorielle Varianzanalyse mit Messwiederholung auf einem Faktor (Faktor 1: Messwiederholung, Faktor 2: Gruppe); Ergebnisse des Interaktionseffekts (p, η^2, d) und der Detailanalysen
* Effektstärke bezogen auf Post-Messzeitpunkt
** Effektstärke bezogen auf Follow-Up-I-Messzeitpunkt

9.4 Hypothesenblock II

Der zweite Hypothesenblock (Hypothesen 2–7) tätigt Annahmen zur kurz- und langfristigen Wirksamkeit der Maßnahme bei Schülerinnen und Schüler, die ungünstige Voraussetzungen für den mathematischen Kompetenzerwerb aufweisen. Die kurzfristige Wirksamkeit wird durch Prä-Post-Analysen und die langfristige Wirksamkeit durch Post-Follow-Up-I-Follow-

Up-II-Analysen ermittelt. Die inferenzstatistischen Analysen erfolgen durch zweifaktorielle Varianzanalysen mit Messwiederholung.

9.4.1 Wirksamkeit bei Schülerinnen und Schülern mit einem geringen Vorwissen

Schülerinnen und Schüler mit einem geringen Vorwissen werden im Rahmen dieser Studie definiert über einen durchschnittlichen T-Wert in der Subskala Rechenoperationen des Heidelberger Rechentests 1–4 von kleiner als 40.

Die Überprüfung auf Normalverteilung der Rohwerte der Subskala Rechenoperationen des HRT 1–4 mittels des Kolmogorov-Smirnov-Anpassungstests zeigt, dass in der Experimentalgruppe lediglich die Werte im Prä-Messzeitpunkt nicht normalverteilt sind ($p_{K\text{-}S_Prä} = 0.002$). In der Kontrollgruppe sind die Werte im Prä- und Follow-Up-I-Messzeitpunkt nicht normalverteilt ($p_{K\text{-}S_Prä} = 0.019$, $p_{K\text{-}S_FU\text{-}I} = 0.042$) (für Details siehe Anhang 5).

ERGEBNISSE ZUR HYPOTHESE 2A

Die Hypothese 2a nimmt an, dass die Schülerinnen und Schüler mit einem geringen Vorwissen aus der Experimentalgruppe im Prä-Post-Vergleich eine bessere Entwicklung in den mathematischen Basiskompetenzen (gemessen durch die Subskala Rechenoperationen des HRT 1–4) zeigen als die Schülerinnen und Schüler mit einem geringen Vorwissen aus der Kontrollgruppe.

Abbildung 50 zeigt, dass die Schülerinnen und Schüler aus der Experimentalgruppe im Post-Messzeitpunkt durchschnittlich 15,8 Aufgaben mehr richtig lösen als im Prä-Messzeitpunkt, in der Kontrollgruppe sind es durchschnittlich 7,77 Aufgaben.

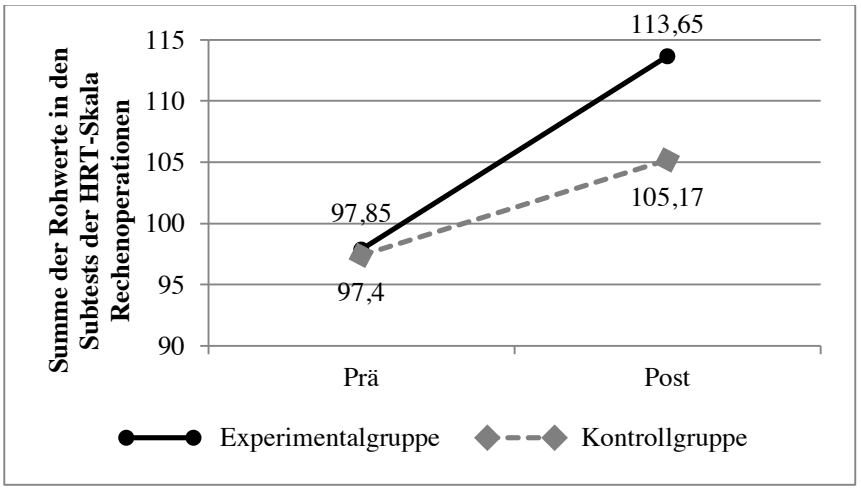


ABBILDUNG 50 PRÄ-POST-ENTWICKLUNG DER SCHÜLERINNEN UND SCHÜLER MIT EINEM GERINGEN VORWISSEN

Bezüglich der drei zentralen Effekte zeigt sich Folgendes:

Der Haupteffekt „Messwiederholung“ (Innersubjekteffekt) zeigt an, dass sich die Mittelwerte der Rohwerte beider Gruppen *zusammen* zwischen den beiden Messzeitpunkten signifikant unterscheiden ($F(1) = 108{,}731$; $p = 0.000$; $\eta^2 = 0.382$). Für beide Gruppen *zusammen* zeigt sich also eine signifikante Verbesserung in den mathematischen Kompetenzen zwischen der Prä- und Post-Erhebung.

Der zweite Haupteffekt „Gruppenzugehörigkeit“ (Zwischensubjekteffekt) ist hingegen nicht signifikant ($F(1) = 3.319$; $p = 0.070$; $\eta^2 = 0.019$).

Der Interaktionseffekt ist signifikant (F(1) = 12.608; $p = 0.000$; $\eta^2 = 0.067$). Die Entwicklung der mathematischen Basiskompetenzen von der Prä- zur Posttestung unterscheidet sich mit einem mittleren Effekt ($d = 0.536$) signifikant in Abhängigkeit von der Gruppenzugehörigkeit und dem Messzeitpunkt. Die Schülerinnen und Schüler mit einem geringen Vorwissen in der Experimentalgruppe entwickeln sich im Prä-Post-Vergleich signifikant besser als diejenigen mit einem geringen Vorwissen in der Kontrollgruppe.

Es liegt eine – nach der hier zugrunde gelegten Definition – kurzfristige Wirksamkeit der Fördermaßnahme bei Schülerinnen und Schülern mit einem geringen Vorwissen vor.

Die zentralen Werte der Analyse sind in Tabelle 53 zusammengefasst.

TABELLE 53 ZENTRALE KENNWERTE ZUR HYPOTHESE 2A – KURZFRISTIGE WIRKSAMKEIT BEI EINEM GERINGEN VORWISSEN

	Experimentalgruppe			**Kontrollgruppe**				
t	N	M	SD	N	M	SD	p	η^2 (d)
Prä	68	97,85	17,14	110	97,40	12,18	0.000	0.067 (0.536)
Post		113,65	24,94		105,17	16,67		

Anmerkungen:
Zweifaktorielle Varianzanalyse mit Messwiederholung auf einem Faktor (Faktor 1: Messwiederholung, Faktor 2: Gruppe); Ergebnisse des Interaktionseffekts (p, η^2, d)

ERGEBNISSE ZUR HYPOTHESE 2B

Die Hypothese 2b nimmt an, dass die Schülerinnen und Schüler mit einem geringen Vorwissen aus der Experimentalgruppe im Zeitraum zwischen der Post-, Follow-Up-I- und Follow-Up-II Erhebung eine bessere Entwicklung in den mathematischen Basiskompetenzen (gemessen durch die Subskala Rechenoperationen des HRT 1–4) zeigen als die Schülerinnen und Schüler mit einem geringen Vorwissen aus der Kontrollgruppe (langfristige Wirksamkeit). Da die Varianzanalyse grundsätzlich nur die Fälle einschließt, für die zu jedem Messzeitpunkt Werte vorhanden sind, können sich die Post-Werte hier im Vergleich zur Hypothese 2a unterscheiden.

Die Abbildung 51 zeigt, dass die Schülerinnen und Schüler der Experimentalgruppe im Follow-Up-I-Messzeitpunkt durchschnittlich 1,55 Aufgaben weniger richtig lösen als im Post-Messzeitpunkt. Im Follow-Up-II-Messzeitpunkt hingegen lösen wieder 10,12 Aufgaben mehr richtig als im FU-I-Messzeitpunkt. Die Schülerinnen und Schüler der Kontrollgruppe lösen im Follow-Up-I-Messzeitpunkt durchschnittlich 7,58 Aufgaben mehr richtig als im Post-Messzeitpunkt; im Follow-Up-II-Messzeitpunkt weitere 4,27 Aufgaben als im FU-I-Messzeitpunkt.

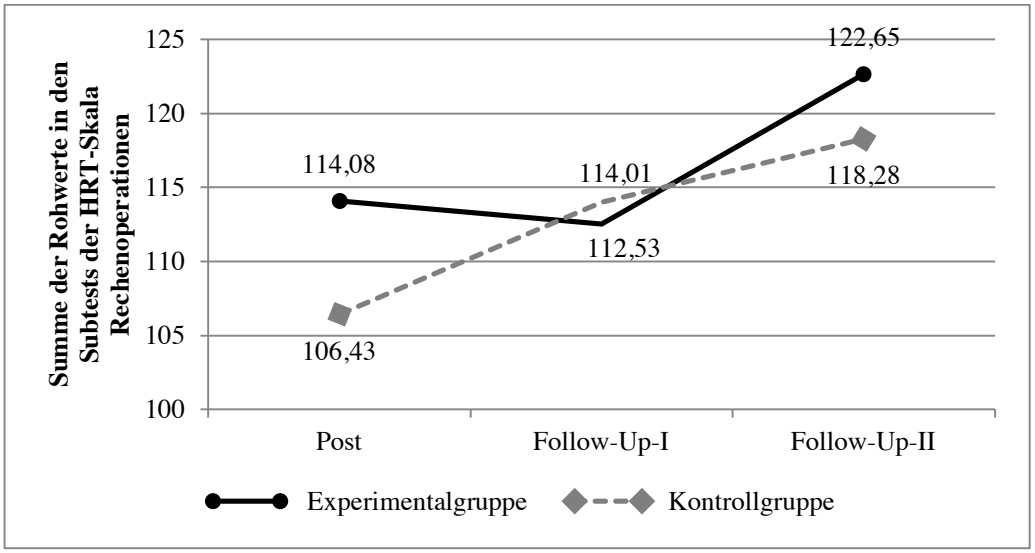


ABBILDUNG 51 POST-FOLLOW-UP-I-FOLLOW-UP-II-ENTWICKLUNG DER SCHÜLERINNEN UND SCHÜLER MIT EINEM GERINGEN VORWISSEN

Die Voraussetzung der Sphärizität (Homogenität der Varianzen der paarweisen Differenzen zwischen den Messzeitpunkten) ist nicht gegeben ($p_{\text{Mauchly}} = 0.007$), sodass im Folgenden der Huynh-Feldt-Korrekturwert ($p = 0.955$) genutzt wird.

Bezüglich der drei zentralen Effekte, die die zweifaktorielle Varianzanalyse ergibt, zeigt sich Folgendes:

Der Haupteffekt „Messwiederholung“ (Innersubjekteffekt) zeigt an, dass sich die Mittelwerte der Rohwerte beider Gruppen *zusammen* zwischen den drei Messzeitpunkten signifikant unterscheiden ($F(1.905) = 4.317$; $p = 0.016$; $\eta^2 = 0.031$). Für beide Gruppen *zusammen* zeigen sich also signifikante Unterschiede in den mathematischen Basiskompetenzen über den Zeitraum von der Post-, der Follow-Up-I- zur Follow-Up-II-Erhebung. Die Messzeitpunktbedingung hat einen positiven Einfluss auf die Entwicklung der mathematischen Basiskompetenzen der Schülerinnen und Schüler.

Der zweite Haupteffekt „Gruppenzugehörigkeit“ (Zwischensubjekteffekt) ist nicht signifikant ($F(1) = 1.040$; $p = 0.310$; $\eta^2 = 0.008$). Die Mittelwerte *zwischen* den beiden Gruppen unterscheiden sich zu den drei Messzeitpunkten nicht signifikant voneinander.

Der Interaktionseffekt ist signifikant (F(1.905) = 4,317; $p = 0.016$; $\eta^2 = 0.031$). Die Entwicklung der mathematischen Kompetenzen von der Post-, über die Follow-Up-I- zur Follow-Up-II-Erhebung unterscheidet sich je nach Gruppenzugehörigkeit und Messzeitpunkt.

Da die visuelle Analyse (siehe Abbildung 51) auf einen Rückgang der Rohwerte in der Experimentalgruppe vom Post- zum Follow-Up-I-Messzeitpunkt hinweist, werden Detailanalysen vorgenommen, um den Interaktionseffekt differenziert aufklären zu können. Mittels eines t-Tests bei unabhängigen Stichproben zeigt sich ein signifikanter Rückgang der Leistungen in der Experimentalgruppe zwischen dem Post- und Follow-Up-I-Messzeitpunkt mit einem kleinen Effekt ($p_t = 0.002$; $d_{\text{korr}} = -0.44$). Die weitere Entwicklung vom Follow-Up-I- zum Follow-Up-II-Messzeitpunkt unterscheidet sich nicht signifikant zwischen den beiden Gruppen ($p_t = 0.100$; $d_{\text{korr}} = 0.257$). Wird die direkte Entwicklung der Experimental- und der Kontrollgruppe vom Post- zum Follow-Up-II-Messzeitpunkt betrachtet, zeigen sich keine Entwick-

lungsunterschiede zwischen den Gruppen ($p_t = 0.281$; $d = –0.158$). Die Ergebnisse der α-Abwärtsadjustierung zeigen, dass die errechneten Signifikanzwerte der Detailanalyse auch nach der α-Abwärtsadjustierung weiterhin als signifikant angesehen werden können (siehe Tabelle 65).

Es liegt – nach der hier zugrunde gelegten Definition – keine langfristige Wirksamkeit der Fördermaßnahme bei Schülerinnen und Schülern mit einem geringen Vorwissen vor. Die zentralen Werte der Varianzanalyse sind in Tabelle 54 zusammengefasst und durch die Ergebnisse der Detailanalysen ergänzt.

TABELLE 54 ZENTRALE KENNWERTE ZUR HYPOTHESE 2B – LANGFRISTIGE WIRKSAMKEIT IN DER GESAMTGRUPPE

	Experimentalgruppe			**Kontrollgruppe**						
t	N	M	SD	N	M	SD	p	η^2 (d)	d_{korr*}	d_{korr**}
Post	50	114,08	26,76	85	106,43	16,12	0.016	0.031 (0.357)	–	–
FU-I		112,53	27,18		114,01	19,67			–0.440	–
FU-II		122,65	26,72		118,28	18,93			–0.158	0.257

Anmerkungen:
Zweifaktorielle Varianzanalyse mit Messwiederholung auf einem Faktor (Faktor 1: Messwiederholung, Faktor 2: Gruppe); Ergebnisse des Interaktionseffekts (p, η^2, d) und der Detailanalysen
* Effektstärke mit Bezug auf den Post-Messzeitpunkt
** Effektstärke mit Bezug auf den Follow-Up-I-Messzeitpunkt

9.4.2 Wirksamkeit bei Schülerinnen und Schülern mit geringen kognitiven Fähigkeiten

Im Rahmen dieser Studie werden Schülerinnen und Schüler mit geringen kognitiven Fähigkeiten definiert über einen Intelligenzquotienten im CFT-20-R (Weiß, 2006) von < 85 (siehe auch Deutsche Gesellschaft für Kinder- und Jugendpsychiatrie, 2011). Die Überprüfung auf Normalverteilung der Rohwerte der Subskala Rechenoperationen des HRT 1–4 mittels des Kolmogorov-Smirnov-Anpassungstests zeigt, dass in der Experimental- und Kontrollgruppe die Rohwerte in allen Messzeitpunkten normalverteilt sind (für Details siehe Anhang 5).

ERGEBNISSE ZUR HYPOTHESE 3A

Die Hypothese 3a nimmt an, dass die Schülerinnen und Schüler mit geringen kognitiven Fähigkeiten der Experimentalgruppe im Prä-Post-Vergleich eine bessere Entwicklung in den mathematischen Basiskompetenzen (gemessen durch die Subskala Rechenoperationen des HRT 1–4) zeigen als die Schülerinnen und Schüler mit geringen kognitiven Fähigkeiten aus der Kontrollgruppe.

Abbildung 52 zeigt, dass die Schülerinnen und Schüler der Experimentalgruppe im Post-Messzeitpunkt durchschnittlich 12,76 Aufgaben mehr richtig lösen als im Prä-Messzeitpunkt und die Schülerinnen und Schüler der Kontrollgruppe 5,3 Aufgaben. Die visuelle Analyse weist auf einen stärkeren Zuwachs in der Experimentalgruppe hin.

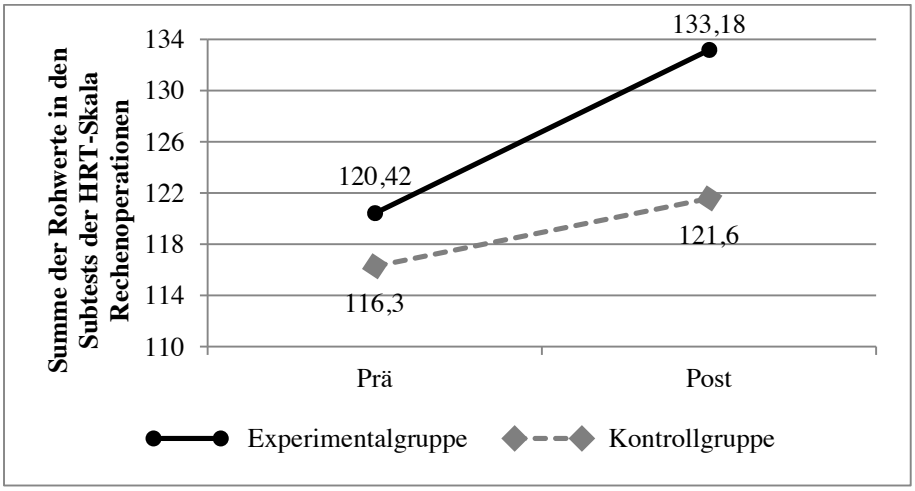


ABBILDUNG 52 PRÄ-POST-ENTWICKLUNG DER SCHÜLERINNEN UND SCHÜLER MIT GERINGEN KOGNITIVEN FÄHIGKEITEN

Bezüglich der drei zentralen Effekte der Varianzanalyse zeigt sich Folgendes:

Der Haupteffekt „Messwiederholung" (Innersubjekteffekt) zeigt an, dass sich die Mittelwerte der Rohwerte beider Gruppen *zusammen* zwischen den beiden Messzeitpunkten signifikant unterscheiden ($F(1) = 43.577$; $p = 0.000$; $\eta^2 = 0.293$). Für beide Gruppen *zusammen* zeigt sich also eine signifikante Verbesserung in den mathematischen Kompetenzen zwischen der Prä- und Post-Erhebung.

Der zweite Haupteffekt „Gruppenzugehörigkeit" (Zwischensubjekteffekt) hingegen ist nicht signifikant ($F(1) = 1.885$; $p = 0.173$; $\eta^2 = 0.018$).

Der Interaktionseffekt ist signifikant ($F(1) = 7.686$; $p = 0.007$; $\eta^2 = 0.068$). Die Entwicklung der mathematischen Basiskompetenzen von der Prä- zur Posttestung unterscheidet sich mit einem mittleren Effekt ($d = 0.540$) signifikant in wechselseitiger Abhängigkeit von der Gruppenzugehörigkeit und dem Messzeitpunkt. Die Schülerinnen und Schüler mit geringen kognitiven Fähigkeiten in der Experimentalgruppe entwickeln sich im Prä-Post-Vergleich signifikant besser als diejenigen in der Kontrollgruppe.

Es liegt eine – nach der hier zugrunde gelegten Definition – kurzfristige Wirksamkeit der Fördermaßnahme bei Schülerinnen und Schülern mit geringen kognitiven Fähigkeiten vor.

Die zentralen Kennwerte zur Hypothese 3a sind in der Tabelle 55 zusammengefasst.

TABELLE 55 ZENTRALE KENNWERTE ZUR HYPOTHESE 3A – KURZFRISTIGE WIRKSAMKEIT BEI GERINGEN KOGNITIVEN FÄHIGKEITEN

	Experimentalgruppe			**Kontrollgruppe**				
t	N	M	SD	N	M	SD	*p*	η^2 (*d*)
Prä	49	120,24	30,85	58	116,30	27,67	0.007	0.068 (0.540)
Post		133,18	29,10		121,60	29,34		

Anmerkungen:
Zweifaktorielle Varianzanalyse mit Messwiederholung auf einem Faktor (Faktor 1: Messwiederholung, Faktor 2: Gruppe); Ergebnisse des Interaktionseffekts (p, $\boldsymbol{\eta}^2$, d)

ERGEBNISSE ZUR HYPOTHESE 3B

Die Hypothese 3b nimmt an, dass die Schülerinnen und Schüler mit geringen kognitiven Fähigkeiten der Experimentalgruppe im Zeitraum zwischen der Post-, Follow-Up-I- und Follow-Up-II Erhebung eine bessere Entwicklung in den mathematischen Basiskompetenzen (gemessen durch die Subskala Rechenoperationen des HRT 1–4) zeigen als die Schülerinnen und Schüler mit geringen kognitiven Fähigkeiten in der Kontrollgruppe (langfristige Wirksamkeit). Da die Varianzanalyse grundsätzlich nur die Fälle einschließt, für die zu jedem Messzeitpunkt Werte vorhanden sind, können sich die Post-Werte hier im Vergleich zur Hypothese 3a unterscheiden.

Die Abbildung 53 zeigt, dass die Schülerinnen und Schüler der Experimentalgruppe im Follow-Up-I-Messzeitpunkt durchschnittlich 1,66 Aufgaben weniger richtig lösen als im Post-Messzeitpunkt, im Follow-Up-II-Messzeitpunkt jedoch wieder 7,65 Aufgaben mehr. Die Schülerinnen und Schüler der Kontrollgruppe lösen im Follow-Up-I-Messzeitpunkt 6,31 Aufgaben mehr richtig als im Post-Messzeitpunkt und weitere 4,12 Aufgaben im Follow-Up-II-Messzeitpunkt.

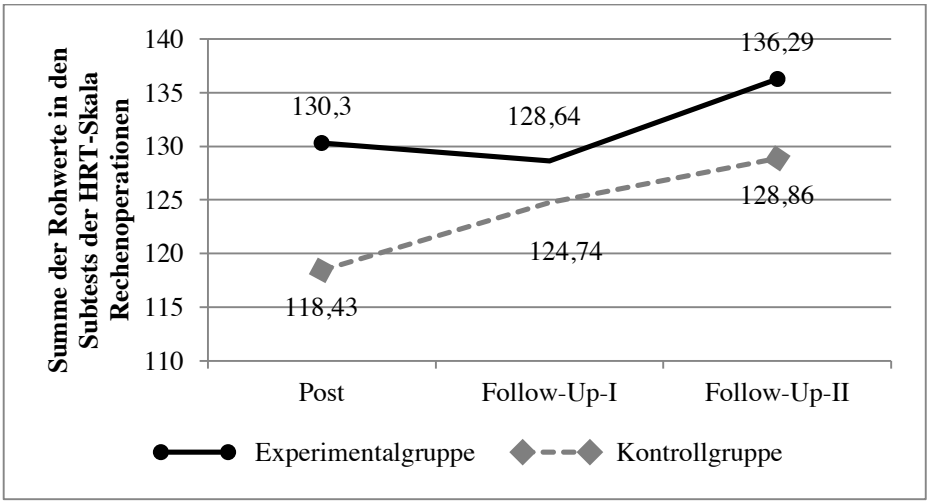


ABBILDUNG 53 POST-FOLLOW-UP-I-FOLLOW-UP-II-ENTWICKLUNG DER SCHÜLERINNEN UND SCHÜLER MIT GERINGEN KOGNITIVEN FÄHIGKEITEN

Die Voraussetzung der Sphärizität (Homogenität der Varianzen der paarweisen Differenzen zwischen den Messzeitpunkten) ist gegeben ($p_{\text{Mauchly}} = 0.779$).

Bezüglich der drei zentralen Effekte, die die zweifaktorielle Varianzanalyse ergibt, zeigt sich Folgendes:

Der Haupteffekt „Messwiederholung“ (Innersubjekteffekt) zeigt an, dass sich die Mittelwerte der Rohwerte beider Gruppen *zusammen* zwischen den drei Messzeitpunkten signifikant unterscheiden ($F(2) = 9{,}138$; $p = 0.000$; $\eta^2 = 0.104$). Für beide Gruppen *zusammen* zeigen sich also signifikante Unterschiede in den mathematischen Basiskompetenzen über den Zeitraum von der Post-, der Follow-Up-I- zur Follow-Up-II-Erhebung, die Messzeitpunktbedingung hat einen Einfluss auf die Entwicklung der mathematischen Basiskompetenzen der Schülerinnen und Schüler.

Der zweite Haupteffekt „Gruppenzugehörigkeit“ (Zwischensubjekteffekt) ist nicht signifikant ($F(1) = 2.020$; $p = 0.159$; $\eta^2 = 0.025$). Die Werte *zwischen* den beiden Gruppen unterscheiden sich zu den drei Messzeitpunkten nicht signifikant voneinander.

Der Interaktionseffekt ist hier nicht signifikant (F(2) = 2.037; $p = 0.134$; $\eta^2 = 0.025$). Die Entwicklung der mathematischen Kompetenzen von der Post-, über die Follow-Up-I- zur Follow-Up-II-Erhebung unterscheidet sich nicht in wechselseitiger Abhängigkeit vom Messzeitpunkt und der Gruppenzugehörigkeit.

Zentrale Kennwerte zu den Ergebnissen der Hypothese 3b zeigt die Tabelle 56.

TABELLE 56 ZENTRALE KENNWERTE ZUR HYPOTHESE 3B – LANGFRISTIGE WIRKSAMKEIT BEI GERINGEN KOGNITIVEN FÄHIGKEITEN

	Experimentalgruppe			**Kontrollgruppe**				
t	N	M	SD	N	M	SD	*p*	η^2 *(d)*
Post	38	130,30	31,09	43	118,43	25,97	0.134	0.025 (0.320)
FU-I		128,64	29,16		124,74	23,11		
FU-II		136,29	26,95		128,86	22,71		

Anmerkungen:
Zweifaktorielle Varianzanalyse mit Messwiederholung auf einem Faktor (Faktor 1: Messwiederholung, Faktor 2: Gruppe); Ergebnisse des Interaktionseffekts (*p*, η^2, *d*)

9.4.3 Wirksamkeit bei Schülerinnen und Schülern mit geringen Lesekompetenzen

Im Rahmen der vorliegenden Untersuchung werden Schülerinnen und Schüler mit geringen Lesekompetenzen definiert über einen Lesequotienten im SLS 1–4 (Mayringer & Wimmer, 2003) von kleiner als 90 (siehe Tabelle 36). Die Überprüfung auf Normalverteilung der Rohwerte der Subskala Rechenoperationen des HRT 1–4 mittels des Kolmogorov-Smirnov-Anpassungstests zeigt, dass in der Experimental und Kontrollgruppe die Rohwerte in allen Messzeitpunkten normalverteilt sind (für Details siehe Anhang 5).

ERGEBNISSE ZUR HYPOTHESE 4A

Die Hypothese 4a nimmt an, dass die Schülerinnen und Schüler mit einer geringen Lesekompetenz der Experimentalgruppe im Prä-Post-Vergleich eine bessere Entwicklung in den mathematischen Basiskompetenzen (gemessen durch die Subskala Rechenoperationen des HRT 1–4) zeigen als die Schülerinnen und Schüler mit einer geringen Lesekompetenz aus der Kontrollgruppe.

Abbildung 54 zeigt, dass die Schülerinnen und Schüler der Experimentalgruppe im Post-Messzeitpunkt durchschnittlich 13,53 Aufgaben mehr richtig lösen als im Prä-Messzeitpunkt, in der Kontrollgruppe sind es 7,40 Aufgaben. Die visuelle Analyse weist auf einen stärkeren Zuwachs in der Experimentalgruppe hin.

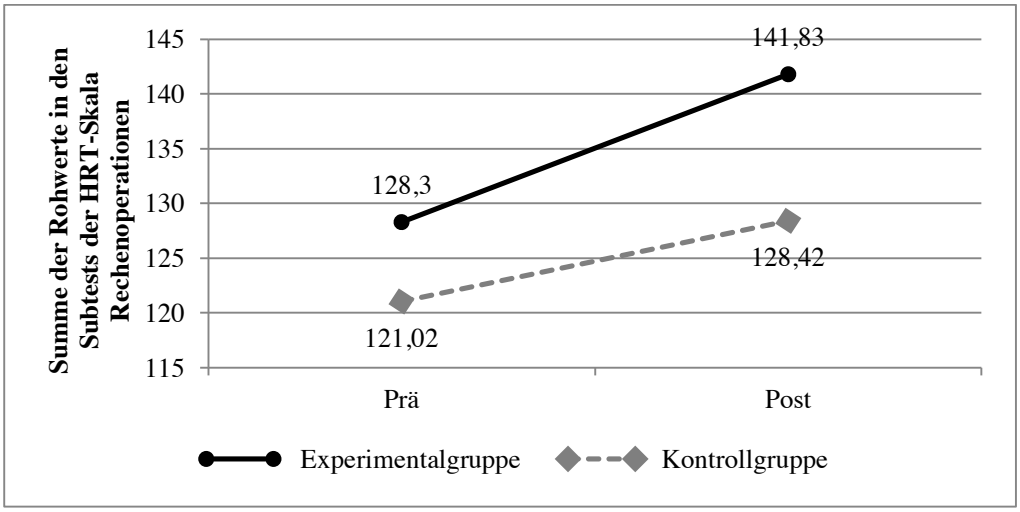


ABBILDUNG 54 PRÄ-POST-ENTWICKLUNG DER SCHÜLERINNEN UND SCHÜLER MIT GERINGEN LESEKOMPETENZEN

Da nur zwei Messzeitpunkte einbezogen werden, sind für die Hypothese 4a keine Angaben zur Sphärizität notwendig. Bezüglich der drei zentralen Effekte der Varianzanalyse zeigt sich Folgendes:

Der Haupteffekt „Messwiederholung“ (Innersubjekteffekt) zeigt an, dass sich die Mittelwerte der Rohwerte beider Gruppen *zusammen* zwischen den beiden Messzeitpunkten signifikant unterscheiden ($F(1) = 171.537$; $p = 0.000$; $\eta^2 = 0.343$). Für beide Gruppen *zusammen* zeigt sich eine signifikante Verbesserung in den mathematischen Basiskompetenzen zwischen der Prä- und Post-Erhebung.

Der zweite Haupteffekt „Gruppenzugehörigkeit“ (Zwischensubjekteffekt) ist ebenso signifikant ($F(1) = 11.446$; $p = 0.001$; $\eta^2 = 0.034$). Die Mittelwerte *zwischen* den beiden Gruppen unterscheiden sich in den beiden Messzeitpunkten signifikant voneinander.

Auch der Interaktionseffekt ist signifikant ($F(1) = 714.675$; $p = 0.000$; $\eta^2 = 0.043$). Die Entwicklung der mathematischen Kompetenzen von der Prä- zur Posttestung entwickelt sich mit einem kleinen Effekt signifikant in wechselseitiger Abhängigkeit von der Gruppenzugehörigkeit und dem Messzeitpunkt. Die Schülerinnen und Schüler mit geringen Lesekompetenzen in der Experimentalgruppe entwickeln sich im Prä-Post-Vergleich signifikant besser als diejenigen in der Kontrollgruppe.

Es liegt eine – nach der hier zugrunde gelegten Definition – kurzfristige Wirksamkeit der Fördermaßnahme bei Schülerinnen und Schülern mit einer geringen Lesekompetenz vor.

Die zentralen Werte der Hypothese 4a sind der Tabelle 57 zu entnehmen.

TABELLE 57 ZENTRALE KENNWERTE ZUR HYPOTHESE 4A – KURZFRISTIGE WIRKSAMKEIT BEI GERINGEN LESEKOMPETENZEN

	Experimentalgruppe			**Kontrollgruppe**				
t	N	M	SD	N	M	SD	p	η^2 *(d)*
Prä	141	128,30	28,56	189	121,02	26,61	0.000	0.043
Post		141,83	30,68		128,42	28,25		(0.424)

Anmerkungen:
Zweifaktorielle Varianzanalyse mit Messwiederholung auf einem Faktor (Faktor 1: Messwiederholung, Faktor 2: Gruppe); Ergebnisse des Interaktionseffekts (p, $\boldsymbol{\eta}^2$, d)

ERGEBNISSE ZUR HYPOTHESE 4B

Die Hypothese 4b nimmt an, dass die Schülerinnen und Schüler mit einer geringen Lesekompetenz der Experimentalgruppe im Zeitraum zwischen der Post-, Follow-Up-I- und Follow-Up-II Erhebung eine bessere Entwicklung in den mathematischen Basiskompetenzen (gemessen durch die Subskala Rechenoperationen des HRT 1–4) zeigen als die Schülerinnen und Schüler mit einer geringen Lesekompetenz in der Kontrollgruppe. Da die Varianzanalyse grundsätzlich nur die Fälle einschließt, für die zu jedem Messzeitpunkt Werte vorhanden sind, können sich die Post-Werte hier im Vergleich zur Hypothese 4a unterscheiden.

Die Abbildung 55 zeigt, dass die Schülerinnen und Schüler der Experimentalgruppe im Follow-Up-I-Messzeitpunkt durchschnittlich 4,25 Aufgaben weniger richtig lösen als im Post-Messzeitpunkt, im Follow-Up-II-Messzeitpunt hingegen wieder 10,11 Aufgaben mehr. Die Schülerinnen und Schüler der Kontrollgruppe lösen im Follow-Up-I-Messzeitpunkt durchschnittlich 6,01 Aufgaben mehr richtig als im Post-Messzeitpunkt und im Follow-Up-II-Messzeitpunkt weitere 3,82 Aufgaben.

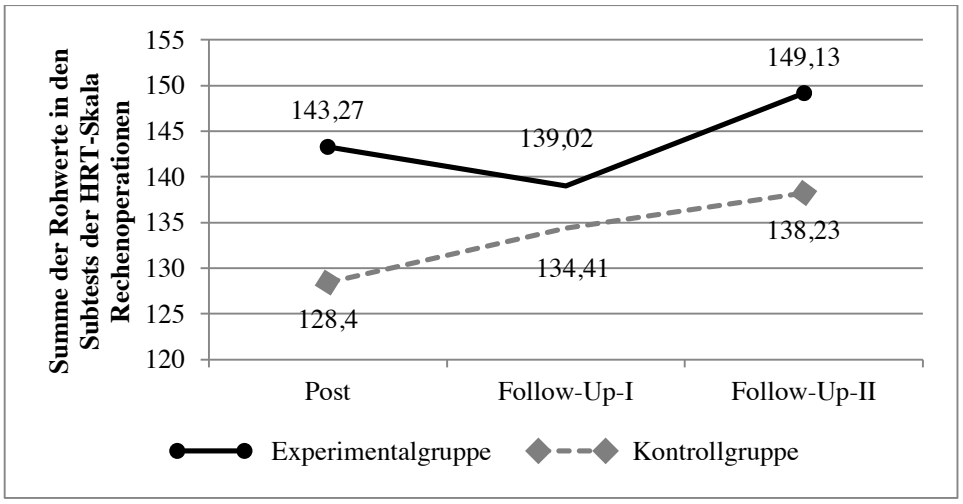


ABBILDUNG 55 POST-FOLLOW-UP-I-FOLLOW-UP-II-ENTWICKLUNG DER SCHÜLERINNEN UND SCHÜLER MIT GERINGEN LESEKOMPETENZEN

Die Voraussetzung der Sphärizität (Homogenität der Varianzen der paarweisen Differenzen zwischen den Messzeitpunkten) ist nicht gegeben ($p_{\text{Mauchly}} = 0.000$), sodass im Folgenden der Huynh-Feldt-Korrekturwert ($p = 0.955$) genutzt wird

Bezüglich der drei zentralen Effekte, die die zweifaktorielle Varianzanalyse ermittelt, zeigt sich Folgendes:

Der Haupteffekt „Messwiederholung“ (Innersubjekteffekt) zeigt an, dass sich die Mittelwerte der Rohwerte beider Gruppen *zusammen* zwischen den drei Messzeitpunkten signifikant unterscheiden ($F(1.875) = 35.199$; $p = 0.000$; $\eta^2 = 0.123$). Die Messzeitpunktbedingung hat einen Einfluss auf die langfristige Entwicklung der mathematischen Basiskompetenzen der Schülerinnen und Schüler.

Auch der zweite Haupteffekt „Gruppenzugehörigkeit“ (Zwischensubjekteffekt) ist signifikant ($F(1) = 8.265$; $p = 0.004$; $\eta^2 = 0.032$). Die Mittelwerte *zwischen* den beiden Gruppen unterscheiden sich zu den drei Messzeitpunkten signifikant voneinander.

Der Interaktionseffekt ist ebenso signifikant ($F(1.875) = 12.776$; $p = 0.000$; $\eta^2 = 0.049$). Die Entwicklung der mathematischen Kompetenzen von der Post-, über die Follow-Up-I- zur Follow-Up-II-Erhebung wird durch die Gruppenzugehörigkeit und durch den Messzeitpunkt wechselseitig beeinflusst.

Die visuelle Analyse der Entwicklung (siehe Abbildung 55) deutet auf einen Rückgang der Rohwerte in der Experimentalgruppe vom Post- zum Follow-Up-I-Messzeitpunkt hin, sodass Detailanalysen vorgenommen werden, um den Interaktionseffekt differenziert aufklären zu können. Mittels eines t-Tests bei unabhängigen Stichproben zeigt sich mit einem kleinen Effekt ein signifikanter Rückgang der Leistungen in der Experimentalgruppe zwischen dem Post- und Follow-Up-I-Messzeitpunkt ($p_t = 0.000$; $d_{korr} = -0.350$). Doch auch die weitere Entwicklung vom Follow-Up-I- zum Follow-Up-II-Messzeitpunkt unterscheidet sich mit einem kleinen Effekt signifikant zwischen den beiden Gruppen ($p_t = 0.003$; $d_{korr} = 0.217$), wobei sich hier die Experimentalgruppe stärker entwickelt. Auch innerhalb des Follow-Up-II unterscheiden sich die beiden Gruppen voneinander ($p_t = 0.004$, $d = 0.354$). Die Experimentalgruppe weist hier deutlich höhere Werte auf als die Kontrollgruppe. Die Ergebnisse der α-Abwärtsadjustierung zeigen, dass die errechneten Signifikanzwerte der Detailanalyse auch nach der α-Abwärtsadjustierung weiterhin als signifikant angesehen werden können (siehe Tabelle 65).

Es liegt eine – nach der hier zugrunde gelegten Definition – langfristige Wirksamkeit der Fördermaßnahme bei Schülerinnen und Schülern mit geringen Lesekompetenzen vor. Die zentralen Werte der Varianzanalyse sind in Tabelle 58 zusammengefasst und durch die Ergebnisse der Detailanalysen ergänzt.

TABELLE 58 ZENTRALE KENNWERTE ZUR HYPOTHESE 4B – LANGFRISTIGE WIRKSAMKEIT BEI GERINGEN LESEKOMPETENZEN

	Experimentalgruppe			**Kontrollgruppe**						
t	N	M	SD	N	M	SD	p	η^2 (d)	d_{korr*}	d_{korr**}
Post	111	143,27	30,85	141	128,40	27,83	0.000	0.049 (0.454)	–	–
FU-I		139,02	30,85		134,41	27,27			–0.350	–
FU-II		149,13	30,81		138,23	28,56			–0.141	0.209

Anmerkungen:
Zweifaktorielle Varianzanalyse mit Messwiederholung auf einem Faktor (Faktor 1: Messwiederholung, Faktor 2: Gruppe); Ergebnisse des Interaktionseffekts (p, η^2, d) und der Detailanalysen
* Effektstärke mit Bezug auf den Post-Messzeitpunkt
** Effektstärke mit Bezug auf den Follow-Up-I-Messzeitpunkt

9.4.4 Wirksamkeit bei Schülerinnen und Schülern mit einem risikobehafteten Sozialverhalten

Im Rahmen dieser Studie werden Schülerinnen und Schüler mit einem risikobehafteten Sozialverhalten über einen durchschnittlichen T-Wert in der Skala „Sozialverhalten“ in der Lehrereinschätzliste für das Sozial- und Lernverhalten (LSL) (Petermann & Petermann, 2013) von kleiner als 40 definiert. Die Überprüfung auf Normalverteilung der Rohwerte der Subskala Rechenoperationen des HRT 1–4 mittels des Kolmogorov-Smirnov-Anpassungstests zeigt, dass in der Experimentalgruppe die Rohwerte in allen Messzeitpunkten normalverteilt sind und in der Kontrollgruppe lediglich der Follow-Up-II-Messzeitpunt nicht ($p_{K\text{-}S_FU\text{-}II} = 0.049$) (für Details siehe Anhang 5).

ERGEBNISSE ZUR HYPOTHESE 5A

Die Hypothese 5a nimmt an, dass die Schülerinnen und Schüler mit einem risikobehafteten Sozialverhalten der Experimentalgruppe im Prä-Post-Vergleich eine bessere Entwicklung in den mathematischen Basiskompetenzen (gemessen durch die Subskala Rechenoperationen des HRT 1–4) zeigen als die Schülerinnen und Schüler mit einem risikobehafteten Sozialverhalten aus der Kontrollgruppe.

Die Abbildung 56 zeigt, dass die Schülerinnen und Schüler der Experimentalgruppe im Post-Messzeitpunkt 10 Aufgaben mehr richtig lösen als im Prä-Messzeitpunkt. In der Kontrollgruppe sind es 3,88 Aufgaben. Die visuelle Analyse weist auf einen stärkeren Leistungszuwachs in der Experimentalgruppe hin.

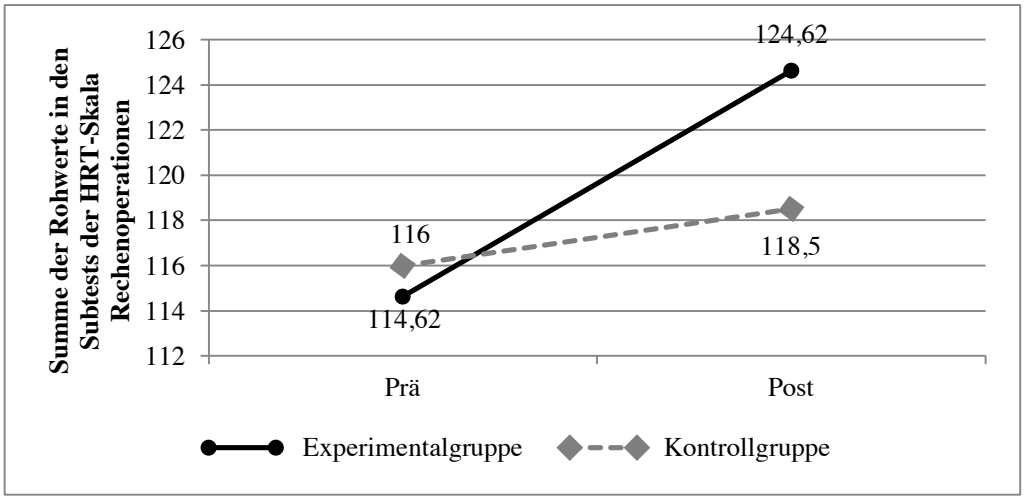


ABBILDUNG 56 PRÄ-POST-ENTWICKLUNG DER SCHÜLERINNEN UND SCHÜLER MIT EINEM RISIKOBEHAFTETEN SOZIALVERHALTEN

Bezüglich der drei zentralen Effekte zeigt sich Folgendes:

Der Haupteffekt „Messwiederholung“ (Innersubjekteffekt) zeigt an, dass sich die Mittelwerte der Rohwerte beider Gruppen *zusammen* zwischen den beiden Messzeitpunkten signifikant unterscheiden ($F(1) = 7.114$; $p = 0.012$; $\eta^2 = 0.173$). Für beide Gruppen *zusammen* zeigen sich also signifikante Unterschiede in den mathematischen Basiskompetenzen zwischen der Prä- und Post-Erhebung.

Der zweite Haupteffekt „Gruppenzugehörigkeit“ (Zwischensubjekteffekt) ist nicht signifikant ($F(1) = 0.067$; $p = 0.797$; $\eta^2 = 0.002$).

Auch der Interaktionseffekt ist nicht signifikant ($F(1) = 2.561$; $p = 0.119$; $\eta^2 = 0.070$). Die Entwicklung der mathematischen Basiskompetenzen von der Prä- zur Posttestung unterscheidet sich *nicht* in wechselseitiger Abhängigkeit von der Gruppenzugehörigkeit und dem Messzeitpunkt. Die Schülerinnen und Schüler mit einem risikobehafteten Sozialverhalten in der Experimentalgruppe entwickeln sich im Prä-Post-Vergleich nicht besser als diejenigen mit einem risikobehafteten Sozialverhalten der Kontrollgruppe.

Es liegt – nach der hier zugrunde gelegten Definition – keine kurzfristige Wirksamkeit der Fördermaßnahme bei Schülerinnen und Schülern mit einem risikobehafteten Sozialverhalten vor. Die zentralen Werte zur Hypothese 5a sind in Tabelle 59 zusammengefasst.

TABELLE 59 ZENTRALE KENNWERTE ZUR HYPOTHESE 5A – KURZFRISTIGE WIRKSAMKEIT BEI EINEM RISIKOBEHAFTETEN SOZIALVERHALTEN

	Experimentalgruppe			**Kontrollgruppe**				
t	N	M	SD	N	M	SD	*p*	η^2 (*d*)
Prä	17	114,62	27,06	19	116,00	27,60	0.119	0.070 (0.549)
Post		124,62	31,30		118,50	26,91		

Anmerkungen:
Zweifaktorielle Varianzanalyse mit Messwiederholung auf einem Faktor (Faktor 1: Messwiederholung, Faktor 2: Gruppe); Ergebnisse des Interaktionseffekts *(p, η^2, d)*

ERGEBNISSE ZUR HYPOTHESE 5B

Die Hypothese 5b nimmt an, dass die Schülerinnen und Schüler mit einem risikobehafteten Sozialverhalten der Experimentalgruppe im Zeitraum zwischen der Post-, Follow-Up-I- und Follow-Up-II Erhebung eine bessere Entwicklung in den mathematischen Basiskompetenzen (gemessen durch die Subskala Rechenoperationen des HRT 1–4) zeigen als die Schülerinnen und Schüler mit einem risikobehafteten Sozialverhalten in der Kontrollgruppe. Da die Varianzanalyse grundsätzlich nur die Fälle einschließt, für die zu jedem Messzeitpunkt Werte vorhanden sind, können sich die Post-Werte hier im Vergleich zur Hypothese 5a unterscheiden.

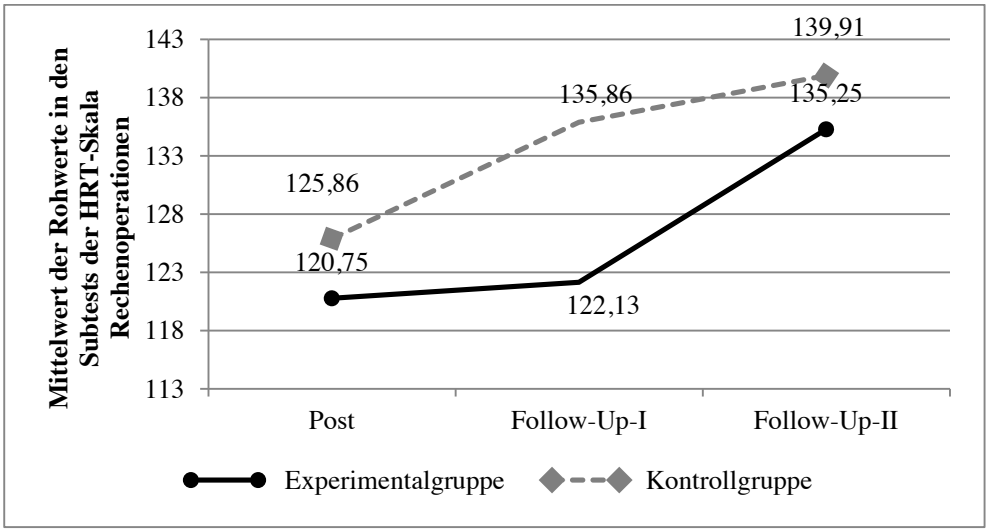


ABBILDUNG 57 POST-FOLLOW-UP-I-FOLLOW-UP-II-ENTWICKLUNG DER SCHÜLERINNEN UND SCHÜLER MIT EINEM RISIKOBEHAFTETEN SOZIALVERHALTEN

Die Voraussetzung der Sphärizität (Homogenität der Varianzen der paarweisen Differenzen zwischen den Messzeitpunkten) ist gegeben ($p_{\text{Mauchly}} = 0.293$).

Bezüglich der drei zentralen Effekte, die die zweifaktorielle Varianzanalyse ergibt, zeigt sich Folgendes:

Der Haupteffekt „Messwiederholung“ (Innersubjekteffekt) zeigt an, dass sich die Werte beider Gruppen *zusammen* zwischen den drei Messzeitpunkten signifikant unterscheiden ($F(2) = 12.989$; $p = 0.000$; $\eta^2 = 0.382$). Für beide Gruppen *zusammen* zeigt sich, dass die Messzeitpunktbedingung einen Einfluss auf die Entwicklung der mathematischen Kompetenzen der Schülerinnen und Schüler nimmt.

Der zweite Haupteffekt „Gruppenzugehörigkeit“ (Zwischensubjekteffekt) ist nicht signifikant ($F(1) = 0.529$; $p = 0.475$; $\eta^2 = 0.025$). Die Mittelwerte *zwischen* den beiden Gruppen unterscheiden sich zu den drei Messzeitpunkten nicht signifikant voneinander.

Der Interaktionseffekt ist hier ebenso nicht signifikant ($F(2) = 1.646$; $p = 0.205$; $\eta^2 = 0.073$). Die Entwicklung der mathematischen Kompetenzen von der Post-, über die Follow-Up-I- zur Follow-Up-II-Erhebung unterscheidet sich nicht in kombinierter Abhängigkeit vom Messzeitpunkt und der Gruppenzugehörigkeit.

Es liegt – nach der hier zugrunde gelegten Definition – keine langfristige Wirksamkeit der Fördermaßnahme bei Schülerinnen und Schülern mit einem risikobehafteten Sozialverhalten vor.

Die zentralen Werte zur Hypothese 5b sind in der Tabelle 60 einzusehen.

TABELLE 60 ZENTRALE KENNWERTE ZUR HYPOTHESE 5B – LANGFRISTIGE WIRKSAMKEIT BEI EINEM RISIKOBEHAFTETEN SOZIALVERHALTEN

	Experimentalgruppe			**Kontrollgruppe**				
t	N	M	SD	N	M	SD	*p*	η^2 (*d*)
Post	12	120,75	33,26	11	125,86	23,94	0.205	0.073 (0.561)
FU-I		122,13	31,25		135,86	17,18		
FU-II		135,25	29,87		139,91	20,93		

Anmerkungen:
Zweifaktorielle Varianzanalyse mit Messwiederholung auf einem Faktor (Faktor 1: Messwiederholung, Faktor 2: Gruppe); Ergebnisse des Interaktionseffekts *(p, η^2, d)*

9.4.5 Wirksamkeit bei Schülerinnen und Schülern mit einem risikobehafteten Lernverhalten

Im Rahmen dieser Studie werden Schülerinnen und Schüler mit einem risikobehafteten Lernverhalten über einen durchschnittlichen T-Wert in der Skala „Lernverhalten“ in der Lehrereinschätzliste für das Sozial- und Lernverhalten (Petermann & Petermann, 2013) von kleiner als 40 definiert. Die Überprüfung auf Normalverteilung der Rohwerte der Subskala Rechenoperationen des HRT 1–4 mittels des Kolmogorov-Smirnov-Anpassungstests zeigt, dass in der Experimental- und Kontrollgruppe die Rohwerte in allen Messzeitpunkten außer im Prä-Messzeitpunt in der Experimentalgruppe ($p_{\text{K-S_Prä}} = 0.019$) normalverteilt sind (für Details siehe Anhang 5).

ERGEBNISSE ZUR HYPOTHESE 6A

Die Hypothese 6a nimmt an, dass die Schülerinnen und Schüler mit einem risikobehafteten Lernverhalten der Experimentalgruppe im Prä-Post-Vergleich eine bessere Entwicklung in den mathematischen Basiskompetenzen (gemessen durch die Subskala Rechenoperationen des HRT 1–4) zeigen als die Schülerinnen und Schüler mit einem risikobehafteten Lernverhalten in der Kontrollgruppe.

Die Abbildung 58 zeigt, dass die Schülerinnen und Schüler mit einem risikobehafteten Lernverhalten im Post-Messzeitpunkt durchschnittlich 12,96 Aufgaben mehr richtig lösen als im Prä-Messzeitpunkt. Bei den Schülerinnen und Schülern der Kontrollgruppe sind es 1,80 Aufgaben mehr. Die visuelle Analyse weist auf einen stärkeren Zuwachs in der Experimentalgruppe hin.

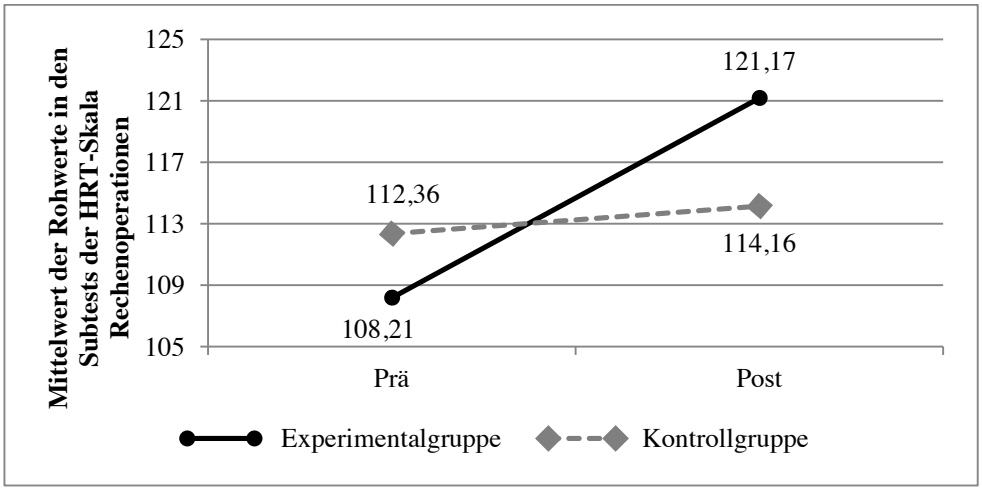


ABBILDUNG 58 PRÄ-POST-ENTWICKLUNG DER SCHÜLERINNEN UND SCHÜLER MIT EINEM RISIKOBEHAFTETEN LERNVERHALTEN

Bezüglich der drei zentralen Effekte zeigt sich Folgendes:

Der Haupteffekt „Messwiederholung" (Innersubjekteffekt) zeigt an, dass sich die Mittelwerte der Rohwerte beider Gruppen *zusammen* zwischen den beiden Messzeitpunkten signifikant unterscheiden ($F(1) = 9.727$; $p = 0.003$; $\eta^2 = 0.192$). Für beide Gruppen *zusammen* zeigt sich also eine signifikante Verbesserung in den mathematischen Kompetenzen zwischen der Prä- und Post-Erhebung.

Der zweite Haupteffekt „Gruppenzugehörigkeit" (Zwischensubjekteffekt) ist hingegen nicht signifikant ($F(1) = 0.032$; $p = 0.859$; $\eta^2 = 0.001$).

Der Interaktionseffekt ist signifikant ($F(1) = 5.567$; $p = 0.023$; $\eta^2 = 0.120$). Die Entwicklung der mathematischen Basiskompetenzen von der Prä- zur Posttestung unterscheidet sich kurzfristig in wechselseitiger Abhängigkeit von der Gruppenzugehörigkeit und dem Messzeitpunkt. Die Schülerinnen und Schüler mit einem risikobehafteten Lernverhalten in der Experimentalgruppe entwickeln sich mit einem mittleren Effekt im Prä-Post-Vergleich besser als diejenigen mit einem risikobehafteten Lernverhalten der Kontrollgruppe.

Es liegt eine – nach der hier zugrunde gelegten Definition – kurzfristige Wirksamkeit der Fördermaßnahme bei Schülerinnen und Schülern mit einem risikobehafteten Lernverhalten vor.

Die zentralen Werte zur Hypothese 6a sind in Tabelle 61 zusammengefasst.

TABELLE 61 ZENTRALE KENNWERTE ZUR HYPOTHESE 6A – KURZFRISTIGE WIRKSAMKEIT BEI EINEM RISIKOBEHAFTETEN LERNVERHALTEN

	Experimentalgruppe			**Kontrollgruppe**				
t	N	M	SD	N	M	SD	*p*	η^2 (*d*)
Prä	21	108,21	25,29	22	112,36	27,90	0.023	0.120 (0.739)
Post		121,17	28,45		114,16	27,21		

Anmerkungen:
Zweifaktorielle Varianzanalyse mit Messwiederholung auf einem Faktor (Faktor 1: Messwiederholung, Faktor 2: Gruppe); Ergebnisse des Interaktionseffekts *(p, η^2, d)*

ERGEBNISSE ZUR HYPOTHESE 6B

Die Hypothese 6b nimmt an, dass die Schülerinnen und Schüler mit einem risikobehafteten Lernverhalten der Experimentalgruppe im Zeitraum zwischen der Post-, Follow-Up-I- und Follow-Up-II Erhebung eine bessere Entwicklung in den mathematischen Basiskompetenzen (gemessen durch die Subskala Rechenoperationen des HRT 1–4) zeigen als die Schülerinnen und Schüler mit einem risikobehafteten Lermverhalten in der Kontrollgruppe. Da die Varianzanalyse grundsätzlich nur die Fälle einschließt, für die zu jedem Messzeitpunkt Werte vorhanden sind, können sich die Post-Werte hier im Vergleich zur Hypothese 6a unterscheiden.

Die Abbildung 59 zeigt, dass die Schülerinnen und Schüler der Experimentalgruppe im Follow-Up-I-Messzeitpunkt durchschnittlich 3,14 Aufgaben mehr richtig lösen als im Post-Messzeitpunkt und im Follow-Up-II-Messzeitpunkt weitere 7,22. Die Kontrollgruppe löst im Follow-Up-I-Messzeitpunkt durchschnittlich 9,03 Aufgaben mehr richtig als im Post-Messzeitpunkt und im Follow-Up-II-Messzeitpunkt weitere 3,94 Aufgaben.

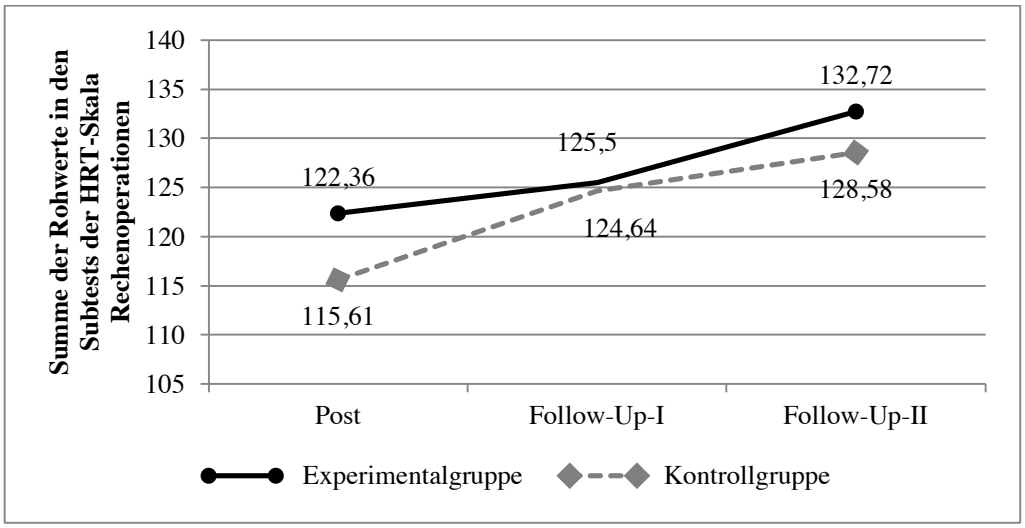


ABBILDUNG 59 POST-FOLLOW-UP-I-FOLLOW-UP-II-ENTWICKLUNG DER SCHÜLERINNEN UND SCHÜLER MIT EINEM RISIKOBEHAFTETEN LERNVERHALTEN

Die Voraussetzung der Sphärizität (Homogenität der Varianzen und Kovarianzen) ist gegeben ($p_{\text{Mauchly}} = 0.627$). Bezüglich der drei zentralen Effekte, die die zweifaktorielle Varianzanalyse ermittelt, zeigt sich Folgendes:

Der Haupteffekt „Messwiederholung“ (Innersubjekteffekt) zeigt an, dass sich die Mittelwerte der Rohwerte beider Gruppen *zusammen* zwischen den drei Messzeitpunkten signifikant unterscheiden ($F(2) = 8.536$; $p = 0.000$; $\eta^2 = 0.201$). Für beide Gruppen *zusammen* zeigen sich also signifikante Unterschiede in den mathematischen Basiskompetenzen über den Zeitraum von der Post-, der Follow-Up-I- zur Follow-Up-II-Erhebung, die Messzeitpunktbedingung hat einen Einfluss auf die Entwicklung der mathematischen Kompetenzen der Schülerinnen und Schüler.

Der zweite Haupteffekt „Gruppenzugehörigkeit“ (Zwischensubjekteffekt) ist nicht signifikant ($F(1) = 0.223$; $p = 0.640$; $\eta^2 = 0.007$). Die Mittelwerte *zwischen* den beiden Gruppen unterscheiden sich zu den drei Messzeitpunkten nicht signifikant voneinander.

Auch der Interaktionseffekt ist nicht signifikant ($F(2) = 0.546$; $p = 0.582$; $\eta^2 = 0.016$). Die Entwicklung der mathematischen Basiskompetenzen von der Post-, über die Follow-Up-I- zur

Follow-Up-II-Erhebung unterscheidet sich nicht in wechselseitiger Abhängigkeit vom Messzeitpunkt und der Gruppenzugehörigkeit.

Es liegt – nach der hier zugrunde gelegten Definition – keine langfristige Wirksamkeit der Fördermaßnahme bei Schülerinnen und Schülern mit einem risikobehafteten Lernverhalten vor.

Die zentralen Werte zur Hypothese 6b sind in der Tabelle 62 einzusehen.

TABELLE 62 ZENTRALE KENNWERTE ZUR HYPOTHESE 6B – LANGFRISTIGE WIRKSAMKEIT BEI EINEM RISIKOBEHAFTETEN LERNVERHALTEN

	Experimentalgruppe			**Kontrollgruppe**				
t	N	M	SD	N	M	SD	*p*	η^2 *(d)*
Post	18	122,36	30,42	18	115,61	25,41	0.582	0.016 (0.255)
FU-I		125,50	27,80		124,64	22,52		
FU-II		132,72	29,18		128,58	24,24		

Anmerkungen:
Zweifaktorielle Varianzanalyse mit Messwiederholung auf einem Faktor (Faktor 1: Messwiederholung, Faktor 2: Gruppe); Ergebnisse des Interaktionseffekts *(p, η^2, d)*

9.4.6 Wirksamkeit bei Schülerinnen und Schülern mit einem sonderpädagogischen Unterstützungsbedarf im Bereich Lernen

Im folgenden Hypothesenkomplex werden ausschließlich die Schülerinnen und Schüler mit einem sonderpädagogischen Unterstützungsbedarf im Bereich Lernen berücksichtigt. Die Auskunft über vorhandene sonderpädagogische Unterstützungsbedarfe erteilten die Lehrkräfte der Schülerinnen und Schüler. Die Überprüfung auf Normalverteilung der Rohwerte der Subskala Rechenoperationen des HRT 1–4 mittels des Kolmogorov-Smirnov-Anpassungstests zeigt, dass in der Experimental und Kontrollgruppe die Rohwerte in allen Messzeitpunkten normalverteilt sind (für Details siehe Anhang 5).

ERGEBNISSE ZUR HYPOTHESE 7A

Die Hypothese 7a nimmt an, dass die Schülerinnen und Schüler mit einem sonderpädagogischen Unterstützungsbedarf im Bereich Lernen der Experimentalgruppe im Prä-Post-Vergleich eine bessere Entwicklung in den mathematischen Kompetenzen (gemessen durch die Subskala Rechenoperationen des HRT 1–4) zeigen als die Schülerinnen und Schüler aus der Kontrollgruppe.

Die Abbildung 60 zeigt, dass die Schülerinnen und Schüler in der Experimentalgruppe im Post-Messzeitpunkt durchschnittlich 7,00 Aufgaben mehr richtig lösen als im Prä-Messzeitpunkt. Die Schülerinnen und Schüler der Kontrollgruppe lösen im Vergleich durchschnittlich 2,53 Aufgaben mehr richtig. Die visuelle Analyse weist auf einen stärkeren Zuwachs der mathematischen Leistungen in der Experimentalgruppe hin.

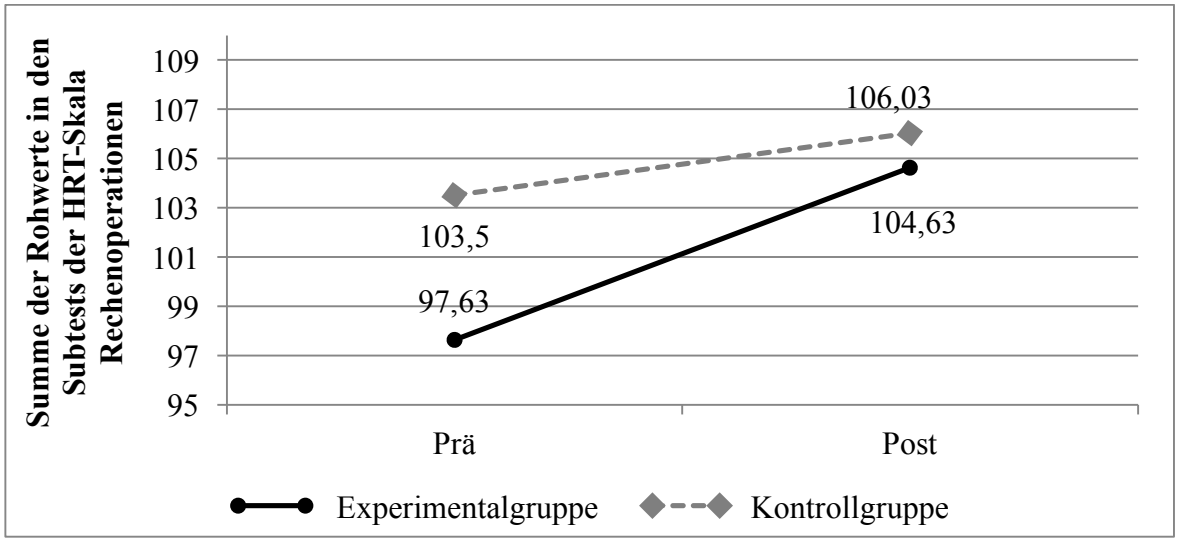


ABBILDUNG 60 PRÄ-POST-ENTWICKLUNG DER SCHÜLERINNEN UND SCHÜLER MIT EINEM SONDERPÄDAGOGISCHEN UNTERSTÜTZUNGSBEDARF IM BEREICH LERNEN

Bezüglich der drei zentralen Effekte zeigt sich Folgendes:

Der Haupteffekt „Messwiederholung“ (Innersubjekteffekt) zeigt an, dass sich die Mittelwerte der Rohwerte beider Gruppen *zusammen* zwischen den beiden Messzeitpunkten nicht signifikant unterscheiden ($F(1) = 2.765$; $p = 0.111$; $\eta^2 = 0.112$). Für beide Gruppen *zusammen* zeigt sich also *keine* Verbesserung in den mathematischen Kompetenzen zwischen der Prä- und Post-Erhebung.

Der zweite Haupteffekt „Gruppenzugehörigkeit“ (Zwischensubjekteffekt) ist ebenso nicht signifikant ($F(1) = 0.065$; $p = 0.802$; $\eta^2 = 0.003$).

Auch der Interaktionseffekt ist nicht signifikant ($F(1) = 0.608$; $p = 0.444$; $\eta^2 = 0.027$). Die Entwicklung der mathematischen Basiskompetenzen von der Prä- zur Posttestung unterscheidet sich kurzfristig nicht in wechselseitiger Abhängigkeit von der Gruppenzugehörigkeit und dem Messzeitpunkt. Die Schülerinnen und Schüler mit einem sonderpädagogischen Unterstützungsbedarf im Bereich Lernen in der Experimentalgruppe entwickeln sich im Prä-Post-Vergleich nicht besser als diejenigen der Kontrollgruppe.

Es liegt – nach der hier zugrunde gelegten Definition – keine kurzfristige Wirksamkeit der Fördermaßnahme bei Schülerinnen und Schülern mit einem sonderpädagogischen Unterstützungsbedarf vor.

Die zentralen Kennwerte der Hypothese 7a sind in Tabelle 63 zusammengefasst.

TABELLE 63 ZENTRALE KENNWERTE ZUR HYPOTHESE 7A – KURZFRISTIGE WIRKSAMKEIT BEI EINEM UNTERSTÜTZUNGSBEDARF IM BEREICH LERNEN

	Experimentalgruppe			**Kontrollgruppe**				
t	N	M	SD	N	M	SD	*p*	η^2 *(d)*
Prä	8	97,63	38,70	16	103,50	26,44	0.444	0.027 (0.333)
Post		104,63	48,69		106,03	28,62		

Anmerkungen:
Zweifaktorielle Varianzanalyse mit Messwiederholung auf einem Faktor (Faktor 1: Messwiederholung, Faktor 2: Gruppe); Ergebnisse des Interaktionseffekts *(p, η^2, d)*

ERGEBNISSE ZUR HYPOTHESE 7B

Die Hypothese 7b nimmt an, dass die Schülerinnen und Schüler mit einem sonderpädagogischen Unterstützungsbedarf im Bereich Lernen der Experimentalgruppe im Zeitraum zwischen der Post-, Follow-Up-I- und Follow-Up-II Erhebung eine bessere Entwicklung in den mathematischen Kompetenzen (gemessen durch die Subskala Rechenoperationen des HRT 1–4) zeigen als die Schülerinnen und Schüler der Kontrollgruppe. Da die Varianzanalyse grundsätzlich nur die Fälle einschließt, für die zu jedem Messzeitpunkt Werte vorhanden sind, können sich die Post-Werte hier im Vergleich zur Hypothese 7a unterscheiden.

Die Abbildung 61 zeigt, dass sich die Schülerinnen und Schüler der Experimentalgruppe in jedem Messzeitpunkt unter dem Niveau der Schülerinnen und Schüler der Kontrollgruppe befinden.

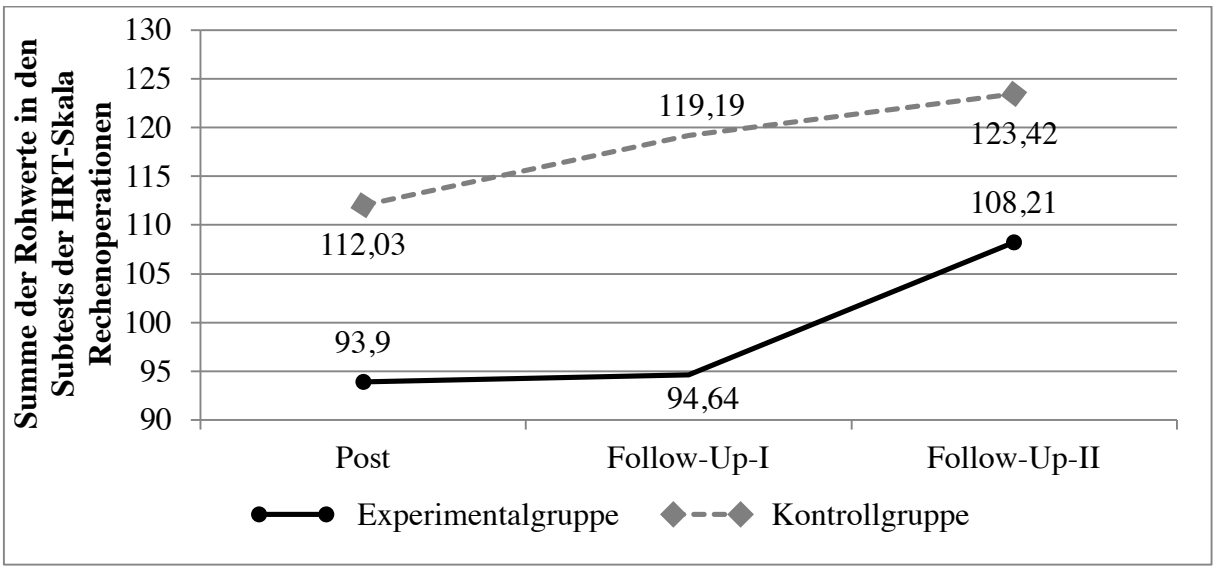


ABBILDUNG 61 POST-FOLLOW-UP-I-FOLLOW-UP-II-ENTWICKLUNG DER SCHÜLERINNEN UND SCHÜLER MIT EINEM SONDERPÄDAGOGISCHEN UNTERSTÜTZUNGSBEDARF IM BEREICH LERNEN

Die Voraussetzung der Sphärizität (Homogenität der Varianzen und Kovarianzen) ist gegeben ($p_{\text{Mauchly}} = 0.184$).

Bezüglich der drei zentralen Effekte, die die zweifaktorielle Varianzanalyse ermittelt, zeigt sich Folgendes:

Der Haupteffekt „Messwiederholung“ (Innersubjekteffekt) zeigt an, dass sich die Mittelwerte der Rohwerte beider Gruppen *zusammen* zwischen den drei Messzeitpunkten signifikant unterscheiden ($F(2) = 8.519$; $p = 0.001$; $\eta^2 = 0.321$). Für beide Gruppen *zusammen* zeigen sich also signifikante Verbesserungen in den mathematischen Kompetenzen über den Zeitraum von der Post-, der Follow-Up-I- zur Follow-Up-II-Erhebung, die Messzeitpunktbedingung hat einen Einfluss auf die Entwicklung der mathematischen Kompetenzen der Schülerinnen und Schüler.

Der zweite Haupteffekt „Gruppenzugehörigkeit“ (Zwischensubjekteffekt) ist nicht signifikant ($F(1) = 2.437$; $p = 0.136$; $\eta^2 = 0.119$). Die Mittelwerte *zwischen* den beiden Gruppen unterscheiden sich zu den drei Messzeitpunkten nicht signifikant voneinander.

Auch der Interaktionseffekt ist nicht signifikant ($F(2) = 1.023$; $p = 0.370$; $\eta^2 = 0.054$). Die Entwicklung der mathematischen Basiskompetenzen von der Post-, über die Follow-Up-I- zur

Follow-Up-II-Erhebung unterscheidet sich nicht in wechselseitiger Abhängigkeit vom Messzeitpunkt und der Gruppenzugehörigkeit.

Es liegt – nach der hier zugrunde gelegten Definition – keine langfristige Wirksamkeit der Fördermaßnahme bei Schülerinnen und Schülern mit einem sonderpädagogischen Unterstützungsbedarf im Bereich Lernen vor. Die zentralen Kennwerte zur Hypothese 7b sind in der Tabelle 64 zusammengefasst.

TABELLE 64 ZENTRALE KENNWERTE ZUR HYPOTHESE 7B – LANGFRISTIGE WIRKSAMKEIT BEI EINEM UNTERSTÜTZUNGSBEDARF IM BEREICH LERNEN

	Experimentalgruppe			**Kontrollgruppe**				
t	N	M	SD	N	M	SD	*p*	η^2 *(d)*
Post	7	93,00	38,78	13	112,04	21,68	0.370	0.054 (0.478)
FU-I		94,64	40,52		119,19	19,59		
FU-II		108,21	37,83		123,42	19,49		

Anmerkungen:
Zweifaktorielle Varianzanalyse mit Messwiederholung auf einem Faktor (Faktor 1: Messwiederholung, Faktor 2: Gruppe); Ergebnisse des Interaktionseffekts (*p*, η^2, *d*)

9.4.7 Ergebnisse der Hypothesenkomplexe I und II im Überblick

Folgende Tabelle 65 gibt einen Überblick über die zentralen Ergebnisse der Hypothesen 1 bis 7 (jeweils a und b) aus den Hypothesenkomplexen I und II. Dargestellt sind – neben den deskriptiven Werten – die Ergebnisse zu den Interaktionseffekten, Innersubjekteffekten und Zwischensubjekteffekten. Rückbeziehend auf das Kapitel 8.7 werden die Definitionen der kurz- und langfristigen Wirksamkeit noch einmal dargelegt:

KURZFRISTIGE WIRKSAMKEIT:

> *Die Schülerinnen und Schüler entwickeln sich im Prä-Post-Vergleich in wechselseitiger Abhängigkeit von der Gruppenzugehörigkeit und der Zeit (Interaktionseffekt). Dabei zeigt die Experimentalgruppe eine bessere Entwicklung als die Kontrollgruppe.*

LANGFRISTIGE WIRKSAMKEIT:

> *a) Die Schülerinnen und Schüler entwickeln sich im Post-Follow-Up-I-Follow-Up-II-Vergleich in wechselseitiger Abhängigkeit von der Gruppenzugehörigkeit und der Zeit (Interaktionseffekt). Dabei zeigt die Experimentalgruppe im Post-Follow-Up-I- und im Follow-Up-I-Follow-Up-II-Vergleich keine schlechtere Entwicklung als die Kontrollgruppe.*
>
> *b) Liegt eine kurzfristige, jedoch keine langfristige Wirksamkeit nach a) vor, spricht ein signifikanter Zwischensubjekteffekt (Haupteffekt) im Post-Follow-Up-I-Follow-Up-II-Vergleich zum Vorteil der Experimentalgruppe für einen persistenten Entwicklungsanschub und somit für die langfristige Wirksamkeit der Förderung.*

Die schraffierten Felder in der Tabelle 65 zeigen an, in welchen Hypothesen die Nullhypothese zu Gunsten der Alternativhypothese verworfen wird – folglich eine Wirksamkeit angenommen werden kann. Die entscheidenden Kennwerte dafür sind entsprechend fett markiert.

TABELLE 65 ERGEBNISSE DER HYPOTHESENKOMPLEXE I UND II IM TABELLARISCHEN ÜBERBLICK

		Experimentalgruppe			**Kontrollgruppe**			**Interaktionseffekt**			**Inns.eff.***	**Zws.eff.***	**Detailanalysen**		
	t	N	M	SD	N	M	SD	*p*	η^2	*d**	*p*	*p*	$p_{t_Post_FU\text{-}I}$ (d_{korr})	$p_{t_FU\text{-}I_FU\text{-}II}$ (d_{korr})	$p_{t_Post_FU\text{-}II}$ (d_{korr})
Gesamtgruppe (H1)	Prä	233	131,92	28,75	344	129,66	28,82	**0.000**	0.027	0.333	0.000	0.058	–	–	–
	Post		143,86	31,37		136,82	30,10								
	Post	190	144,98	31,74	287	138,04	29,65	0.000	0.034	0.375	0.000	0.261	0.000 (–0.281)	0.000 (0.179)	0.048 (–0.103)
	FU-I		141,39	30,79		143,02	29,54								
	FU-II		151,66	30,62		147,83	30,68								
Geringes Vorwissen (H2)	Prä	68	97,85	17,14	110	97,40	12,18	**0.000**	0.067	0.536	0.000	0.070	–	–	–
	Post		113,65	24,94		105,17	16,67								
	Post	50	114,08	26,76	85	106,43	16,12	0.016	0.031	0.357	0.016	0.310	0.002 (–0.440)	0.100 (0.257)	0.281 (–0.158)
	FU-I		112,53	27,18		114,01	19,67								
	FU-II		122,65	26,72		118,28	18,93								

		Experimentalgruppe			Kontrollgruppe			Interaktionseffekt			Inns.eff.*	Zws.eff.*	Detailanalysen		
	t	N	M	SD	N	M	SD	*p*	η^2	*d**	*p*	*p*	$p_{t_Post_FU\text{-}I}(d)$	$p_{t_FU\text{-}I_FU\text{-}II}(d)$	$p_{t_Post_FU\text{-}II}(d)$
Geringe kognitive Fähigkeiten (H3)	Prä	49	120,24	30,85	58	116,30	27,67	**0.007**	0.068	0.540	0.000	0.173	–	–	–
	Post		133,18	29,10		121,60	29,34								
	Post	38	130,30	31,09	43	118,43	25,97	0.134	0.025	0.320	0.000	0.159	–	–	–
	FU-I		128,64	29,16		124,74	23,11								
	FU-II		136,29	26,95		128,86	22,71								
Geringe Lesekompetenz (H4)	Prä	141	128,30	28,56	189	121,02	26,61	**0.000**	0.043	0.424	0.000	0.000	–	–	–
	Post		141,83	30,68		128,42	28,25								
	Post	111	143,27	30,85	141	128,40	27,83	0.000	0.049	0.454	0.000	**0.004**	0.000 (–0.350)	0.003 (0.217)	0.074 (–0.141)
	FU-I		139,02	30,85		134,41	27,27								
	FU-II		149,13	30,81		138,23	28,56								
Risikobehaftetes Sozialverhalten (H5)	Prä	17	114,62	27,06	19	116,00	27,60	0.119	0.070	0.549	0.012	0.797	–	–	–
	Post		124,62	31,30		118,50	26,91								
	Post	12	120,75	33,26	11	125,86	23,94	0.205	0.073	0.561	0.000	0.475	–	–	–
	FU-I		122,13	31,25		135,86	17,18								
	FU-II		135,25	29,87		139,91	20,93								

		Experimentalgruppe			Kontrollgruppe			Interaktionseffekt			Inns.eff.*	Zws.eff.*	Detailanalysen		
	t	N	M	SD	N	M	SD	*p*	η^2	*d**	*p*	*p*	$p_{t_Post_FU\text{-}I}(d)$	$p_{t_FU\text{-}I_FU\text{-}II}(d)$	$p_{t_Post_FU\text{-}II}(d)$
Risikobehaftetes Lernverhalten (H6)	Prä	21	108,21	25,29	22	112,36	27,90	**0.023**	0.120	0.739	0.003	0.859	–	–	–
	Post		121,17	28,45		114,16	27,21								
	Post	18	122,36	30,42	18	115,61	25,41	0.582	0.016	0.255	0.000	0.640	–	–	–
	FU-I		125,50	27,80		124,64	22,52								
	FU-II		132,72	29,18		128,58	24,24								
Sonderpädagogischer Unterstützungsbedarf LE (H7)	Prä	8	97,63	38,70	18	103,50	26,44	0.444	0.027	0.333	0.111	0.802	–	–	–
	Post		104,63	48,69		106,03	28,62								
	Post	7	93,00	38,78	13	112,04	21,68	0.370	0.054	0.478	0.001	0.136	–	–	–
	FU-I		94,64	40,52		119,19	19,59								
	FU-II		108,21	37,83		123,42	19,49								

Anmerkungen: Inns.eff. = Innersubjekteffekt (Haupteffekt), Zws.eff. = Zwischensubjekteffekt (Haupteffekt), *d** = eigene Umrechnung aus dem $\boldsymbol{\eta}^2$

Folgende Tabelle 66 gibt einen Überblick über die Teststärke, für diejenigen Hypothesen, bei denen die Nullhypothese angenommen werden soll (zum Vorgehen siehe Kapitel 0.). Es zeigt sich, dass die Annahme der Nullhypothese in den gefragten Fällen zulässig ist, da die Teststärke jeweils über 90 % liegt. Die Wahrscheinlichkeit, die Nullhypothese anzunehmen, obwohl in Wirklichkeit die Alternativhypothese gilt, ist somit ausreichend gering.

TABELLE 66 ANALYSE DER TESTSTÄRKE IM HYPOTHESENBLOCK I UND II

Hypothese	Teststärke (Power)	Annahme der Nullhypothese
H_2^1	1,000	zulässig
H_2^2	1,000	zulässig
H_2^3	0,999	zulässig
H_1^5	0,999	zulässig
H_2^5	0,999	zulässig
H_2^6	0,917	zulässig
H_1^7	0,995	zulässig
H_2^7	0,999	zulässig

Anmerkung: Nach Rasch und Kollegen (2006a, S. 80) sollte die Power mindestens 90 % betragen, damit die Nullhypothese nicht fälschlicherweise angenommen wird.

EFFEKTSTÄRKENBASIERTE SUKZESSIVE α-ABWÄRTSADJUSTIERUNG

Da im Rahmen der Detailanalysen der langfristigen Wirksamkeit im Hypothesenblock I und II in drei Teilgruppen mehrere t-Tests an derselben Stichprobe durchgeführt werden, ist zur Vermeidung der α-Fehlerkumulierung eine effektstärkenbasierte sukzessive α-Abwärtsadjustierung angezeigt. Diese wird in Tabelle 67 für die Detailanalysen in der langfristigen Wirksamkeit der Gesamtgruppe, bei einem geringen Vorwissen und bei einer geringen Lesekompetenz durchgeführt.

TABELLE 67 EFFEKTSTÄRKENBASIERTE SUKZESSIVE α-ABWÄRTSADJUSTIERUNG IM HYPOTHESENBLOCK I UND II

Platz	α-Fehler-Adjustierung	Merkmal	d_{korr}	p	Nach Adju. signifikant?
Detailanalysen in der langfristigen Wirksamkeit in der Gesamtgruppe					
1	$\frac{\alpha}{1} = .0500$	Mathematische Kompetenzen Post–Follow-Up-I	–0.281	0.000	ja
2	$\frac{\alpha}{2} = 0.250$	Mathematische Kompetenzen Follow-Up-I–Follow-Up-II	0.179	0.000	ja
3	$\frac{\alpha}{3} = 0.166$	Mathematische Kompetenzen Post–Follow-Up-II	–0.103	0.048	ja

Platz	α-Fehler-Adjustierung	Merkmal	d_{korr}	p	Nach Adju. signifikant?
Detailanalysen in der langfristigen Wirksamkeit bei einem geringen Vorwissen					
1	$\frac{\alpha}{1} = .0500$	Mathematische Kompetenzen Post–Follow-Up-I	–0.440	0.002	ja
2	$\frac{\alpha}{2} = 0.250$	Mathematische Kompetenzen Follow-Up-I–Follow-Up-II	0.257	n. s.	–
3	$\frac{\alpha}{3} = 0.166$	Mathematische Kompetenzen Post–Follow-Up-II	–0.158	n. s.	–
Detailanalysen in der langfristigen Wirksamkeit bei einer geringen Lesekompetenz					
1	$\frac{\alpha}{1} = .0500$	Mathematische Kompetenzen Post–Follow-Up-I	–0.350	0.000	ja
2	$\frac{\alpha}{2} = 0.250$	Mathematische Kompetenzen Follow-Up-I–Follow-Up-II	0.217	0.003	ja
3	$\frac{\alpha}{3} = 0.166$	Mathematische Kompetenzen Post–Follow-Up-II	–0.141	n. s.	–

Die Ergebnisse der α-Abwärtsadjustierung zeigen, dass die errechneten Signifikanzwerte, die auch in der zusammenfassenden Tabelle 65 dargestellt sind, auch nach der α-Abwärtsadjustierung weiterhin als signifikant angesehen werden können. Bei den Schülerinnen und Schülern mit einem geringen Vorwissen liegen auf dem Platz 2 und 3 keine signifikanten Ergebnisse vor, sodass keine Adjustierung vorgenommen werden muss. Ähnlich ist es beim Platz 3 bei den Schülerinnen und Schülern mit einer geringen Lesekompetenz.

Ergänzende Ergebnisse

Der Subtest Rechenoperationen des Heidelberger Rechentests 1–4 umfasst die sechs Bereiche Addition, Subtraktion, Multiplikation, Division, Ergänzungsaufgaben und Größer-Kleiner-Vergleiche. Für die bisherigen Ergebnisse und Analysen wurde der durchschnittliche Rohwert aller sechs Subtests hinzugezogen. Folgende Tabelle 68 zeigt ergänzend auf deskriptiver Ebene die Entwicklung in den sechs einzelnen Bereichen vom Prä- bis zum Follow-Up-II-Messzeitpunkt. Es werden die Rohwerte genutzt. Daneben werden die Differenzen zwischen dem Prä- und Post-Messzeitpunkt und dem Post- und Follow-Up-II-Messzeitpunkt angegeben. Jeder der sechs Subtests umfasst insgesamt 40 zu lösende Rechenaufgaben.

Tabelle 68 Ergänzende Ergebnisse: Entwicklung in den sechs Skalen des HRT-Subtests „Rechenoperationen“ (Rohwerte, deskriptiv)

	Prä	Post	FU-I	FU-II	Δ Prä-Post	Δ Post-FU-II
Addition EG	26,98	28,29	27,86	29,39	+ 1,31	+ 1,10
Addition KG	26,87	27,34	27,46	28,79	+ 0,47	+ 1,45
Subtraktion EG	25,58	26,45	25,94	28,23	+ 0,87	+ 1,78
Subtraktion KG	25,14	25,67	26,24	26,96	+ 0,53	+ 1,29
Multiplikation EG	20,94	22,77	20,72	22,93	+ 1,83	+ 0,16
Multiplikation KG	20,70	21,38	21,39	22,02	+ 0,68	+ 0,64

	Prä	Post	FU-I	FU-II	Δ Prä-Post	Δ Post-FU-II
Division EG	19,61	21,61	20,08	22,24	+ 2,00	– 0,37
Division KG	18,66	19,95	20,87	21,58	+ 1,29	+ 1,63
Ergänzungs-aufgaben EG	12,06	15,10	14,42	16,85	+ 3,04	+ 1,75
Ergänzung-saufgaben KG	13,02	15,09	16,36	17,26	+ 2,07	+ 2,17
Größer-Kleiner-Vergleich EG	26,55	28,74	29,88	31,91	+ 2,19	+ 3,17
Größer-Kleiner-Vergleich KG	25,34	27,39	28,83	30,17	+ 2,05	+ 2,78

Im Prä-Post-Vergleich ist zu erkennen, dass die Experimentalgruppe in allen sechs Bereichen einen stärkeren Zuwachs verzeichnet als die Kontrollgruppe. Die größten Vorsprünge zeigen sich in den Bereichen Addition, Multiplikation, Division und in den Ergänzungsaufgaben. Eher geringer sind die Unterschiede zur Kontrollgruppe in den Bereichen Subtraktion und Größer-Kleiner-Vergleich.

In der langfristigen Entwicklung (Post-Follow-Up-II) ist zu erkennen, dass beide Untersuchungsgruppen in allen sechs Subtests – bis auf die Experimentalgruppe im Bereich Division – einen Punktezuwachs verzeichnen. In den Bereichen Subtraktion und Ergänzungsaufgaben weist die Experimentalgruppe insgesamt einen stärkeren Zuwachs auf. In den anderen vier Bereichen ist die Kontrollgruppe stärker. Die Kontrollgruppe weist über die vier Messzeitpunkte in jedem Bereich eine kontinuierliche Entwicklung auf. Hingegen ist in der Experimentalgruppe in den Bereichen Addition, Subtraktion, Multiplikation und in den Ergänzungsaufgaben ein leichter Rückgang zwischen zwischen der Post- und Follow-Up-I-Erhebung zu verzeichnen. Die Ergebnisse sind hier deskriptiv dargestellt.

9.5 Hypothesenblock III

Die Hypothese 8 nimmt an, dass die Lernfreude und die Leistungsheterogenität der Klasse den Punktezuwachs im HRT von der Prä- zur Posterhebung als Prädikatoren vorhersagen.

Die Prüfung auf Normalverteilung zeigt, dass die Werte der Skala Lernfreude normalverteilt sind, bei der Leistungsheterogenität der Klasse ist diese Voraussetzung jedoch verletzt. Detaillierte Informationen zu den Kennwerten finden sich im Anhang 6. Zu Grunde liegen die Angaben von 89,62 % der Schülerinnen und Schüler der Experimentalgruppe.

Der Durbin-Watson-Test ergibt einen Wert von 1,623, sodass der Zusammenhang der Variablen als linear angenommen und eine Autokorrelation der Residualwerte ausgeschlossen werden kann. Die visuelle Analyse des Histogramms indiziert die Homoskedastizität, sodass auch diese Voraussetzung als gegeben gilt (vgl. Bühner & Ziegler, 2009, S. 718).

Die Regressionsanalyse ergibt ein R^2 von 0.036 (R^2_{korr} = 0.027). Es wird also 3,6 % der Varianz in den Daten durch die beiden abhängigen Variablen erklärt (siehe Modellübersicht I in Tabelle 69). Der Signifikanzwert des Modells bleibt unter der α-Fehler-Wahrscheinlichkeit ($F(2) = 4.002$; $p = 0.020$).

TABELLE 69 MODELLÜBERSICHT I

Prädiktoren	R	R^2	R^2_{korr}	β	T	*p*	Semipartielle Korrelation^2
Lernfreude	0.189	0.036	0.027	0.188	2.813	0.005	0.0354
Leistungsheterogenität der Klasse				0.029	0.441	0.659	0.0009

Die Analyse der Koeffizienten ergibt für den Prädiktor „Lernfreude“ einen signifikanten Wert ($p = 0.005, \beta = 0.188$), für den Prädiktor „Leistungsheterogenität der Klasse“ hingegen einen nicht signifikanten Wert ($p = 0.659, \beta = 0.029$). Es leistet ausschließlich der Prädiktor Lernfreude einen signifikanten Beitrag zur Vorhersage des Leistungszuwachses im HRT von der Prä- zur Posterhebung. Die quadrierten Semipartialkorrelationen zeigen, dass sich die gemeinsame Varianz von dem Prädiktor Leistungsheterogenität der Klasse und dem Leistungszuwachs im HRT von der Prä- zur Posterhebung fast vollständig durch die Unterschiede in der Lernfreude erklären lässt. Die Werte der Kollinearitätsdiagnose zeigen, dass von einer mäßigen Kollinearität auszugehen ist.

Wird der Prädiktor „Leistungsheterogenität der Klasse“ herausrausgenommen, ergibt sich ausschließlich mit dem Prädiktor Lernfreude folgendes zweites Modell (siehe Tabelle 70):

TABELLE 70 MODELLÜBERSICHT II

Prädiktor	R	R^2	R^2_{korr}	β	T	*p*
Lernfreude	0.186	0.035	0.030	0.186	2.80	0.006

Es wird in diesem zweiten Modell 3,5 % der Varianz in den Daten durch die abhängige Variable „Lernfreude“ erklärt. Der Signifikanzwert des Modells bleibt auch hier unter der α-Fehler-Wahrscheinlichkeit ($F(2) = 4.002; p = 0.006$).

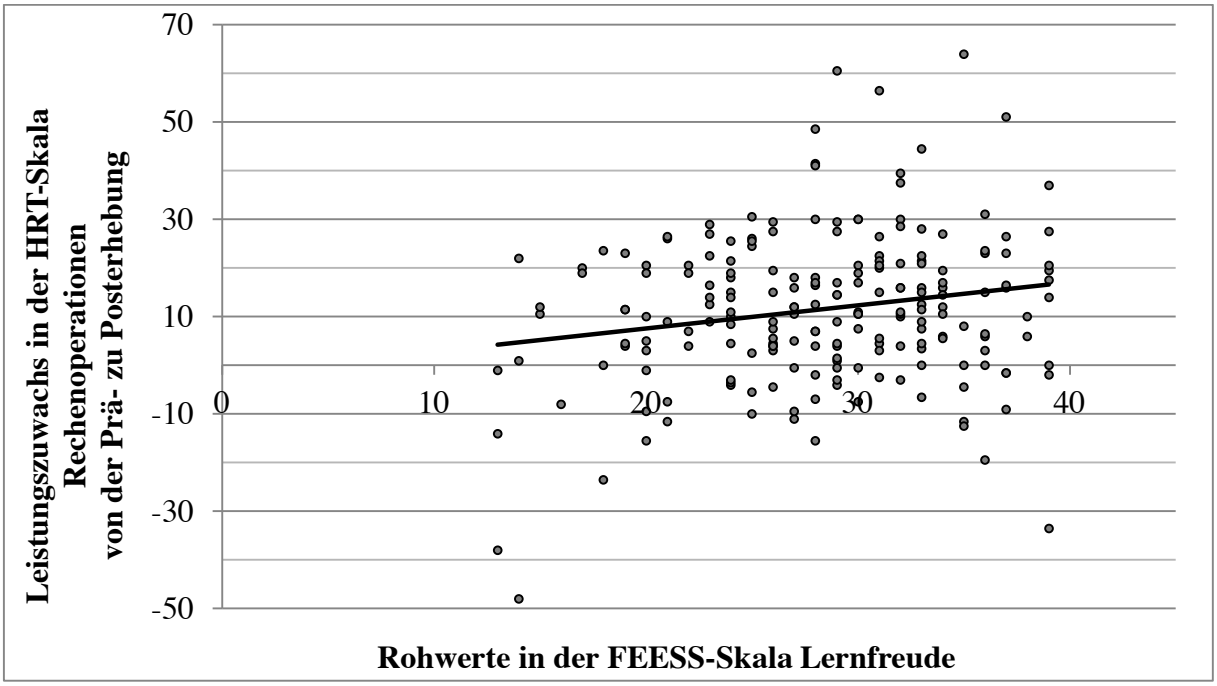


ABBILDUNG 62 KORRELATION ZWISCHEN DEM PRÄDIKTOR LERNFREUDE UND DEM LEISTUNGSZUWACHS IM HRT VON DER PRÄ- ZUR POSTERHEBUNG INKLUSIVE DER REGRESSIONSGERADEN

Die Regressionsgleichung für den Zusammenhang zwischen der Lernfreude und dem Leistungszuwachs im HRT von der Prä- zur Posterhebung (siehe Abbildung 62) lautet:

$$y_{Leistungszuwachs} = 0{,}41 * x_{Lernfreude} - 0{,}31$$

Der Korrelationskoeffizient beträgt $r = 0.186$ ($p = 0.003$).

9.6 Hypothesenblock IV

Der vierte Hypothesenblock beinhaltet die Hypothesen zur Implementationsqualität der Fördermaßnahme aus Sicht der Schülerinnen und Schüler (Hypothesen 9 und 10) und aus Sicht der Interventionsleiterinnen und -leiter (Hypothesen 11–14).

9.6.1 Implementationsqualität aus Sicht der Schülerinnen und Schüler

Insgesamt füllen 250 Schülerinnen und Schüler der Experimentalgruppe (96,15 %) nach Abschluss der Fördermaßnahme den Implementationsfragebogen mit den beiden Skalen „Akzeptanz“ und „Angemessenheit“ aus. Der Implementationsfragebogen ist im Anhang 3 einzusehen. Für die umfassenden Ergebnisse des Implementationsfragebogens auf Itemebene wird hier auf den Anhang 4 verwiesen.

Die Abbildung 63 zeigt deskriptiv die durchschnittliche Zustimmung zu den beiden Implementationsskalen „Akzeptanz“ und „Angemessenheit“.

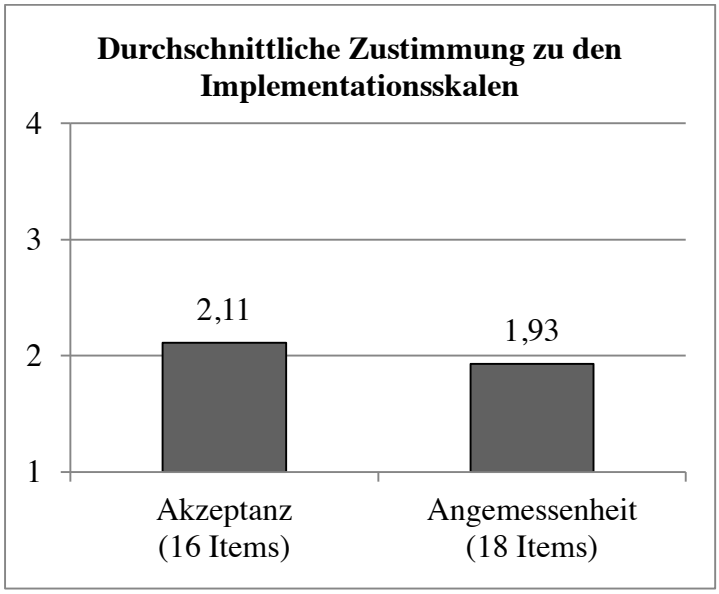


ABBILDUNG 63 DIE DURCHSCHNITTLICHE ZUSTIMMUNG ZU DEN IMPLEMENTATIONSSKALEN AUS SICHT DER SCHÜLERINNEN UND SCHÜLER

Bei den 16 bzw. 18 zugrundeliegenden Items handelt es sich um positiv gerichtete Aussagen, zu denen eine Zustimmung durch die Schülerinnen und Schüler in vier Abstufungen erfolgt. Die Antwortskala umfasst 1 = stimmt genau, 2 = stimmt ziemlich, 3 = stimmt kaum und 4 = stimmt gar nicht. Die Ergebnisse zu den beiden Skalen werden im Folgenden inferenzstatistisch analysiert.

ERGEBNISSE ZUR HYPOTHESE 9

Die Hypothese 9 nimmt an, dass die Fördermaßnahme für die Schülerinnen und Schüler der Experimentalgruppe zufriedenstellend ist und sie diese „akzeptieren“ (Kriterium „acceptability“ nach Petermann, 2015b, S. 123). Angenommen wird, dass sie den 16 Aussagen zur „Akzeptanz“ (AK) durchschnittlich „ziemlich“ zustimmen (entspricht dem Skalenwert 2 von 4). Ausgeschlossen ist dabei die Frage zur Notenvergabe durch die Schülerinnen und Schüler (AK_{Note}), da sie ein abweichendes Skalenniveau aufweist (Note von 1–6).

Der Mittelwert der Skala Akzeptanz über alle 16 positiv gerichteten Aussagen hinweg liegt bei 2,11 (SD: 0,71) (siehe Abbildung 63). Die Prüfung auf Normalverteilung zeigt, dass für die Werte der Skala Akzeptanz keine Normalverteilung vorliegt ($p_{K\text{-}S} = 0.000$). Detailangaben dazu finden sich im Anhang 7. Da somit die Voraussetzung auf Normalverteilung für den t-test verletzt ist, wird auf den nonparametrischen Wilcoxon-Test zurückgegriffen. Der Wilcoxon-Test ergibt eine Signifikanz von $p = 0.106$, sodass keine Abweichung von dem zu prüfenden Mittelwert 2 vorliegt. Die Nullhypothese wird zu Gunsten der Alternativhypothese verworfen.

ERGÄNZENDE ERGEBNISSE ZUR HYPOTHESE 9

Folgende Tabelle 71 gibt ergänzend einen Überblick über die Ergebnisse der Subskalen der Skala „Akzeptanz“. Die höchste Zustimmung erhält die Zufriedenheit mit den Methoden, gefolgt von der Zufriedenheit insgesamt und der Zustimmung zum eigenen positiven Lernerfolg. Die Schülerinnen und Schüler geben der Fördermaßnahme durchschnittlich die Schulnote 2,72. Die umfangreichen Ergebnisse auf Itemebene finden sich im Anhang 4.

TABELLE 71 ERGEBNISSE DER SKALA „AKZEPTANZ“ DES IMPLEMENTATIONSFRAGEBOGENS DER SCHÜLERINNEN UND SCHÜLER AUF SUBSKALENEBENE (DESKRIPTIV)

Variable	Operationalisierung	Anzahl Items	Kürzel	M	SD
Akzeptanz	• Einschätzung des eigenen Lernerfolgs	5	AK_{LE}	2,17	0,88
	• Zufriedenheit mit den Methoden	9	AK_{ZF}	2,07	0,71
	• Zufriedenheit insgesamt	2	AK_{ZI}	2,13	0,96
	• Geben einer Schulnote	1	AK_{Note}	2,72	1,57

ERGEBNISSE ZUR HYPOTHESE 10

Die Hypothese 10 nimmt an, dass die Schülerinnen und Schüler der Experimentalgruppe verschiedene Aspekte der Fördermaßnahme als angemessen empfinden (Kriterium „appropriateness“ nach Petermann, 2015b, S. 123). Sie stimmen den 18 positiv gerichteten Aussagen zur Angemessenheit (AN) der Maßnahme durchschnittlich „ziemlich“ zu (entspricht dem Skalenwert 2 von 4).

Der Mittelwert der Skala Angemessenheit (AN) über die 18 Aussagen hinweg liegt bei 1,93 (SD: 0.47) (siehe Abbildung 63). Die Prüfung auf Normalverteilung zeigt, dass für die Skala Angemessenheit keine Normalverteilung vorliegt ($p_{K\text{-}S} = 0.007$). Detailangaben dazu finden sich im Anhang 7.

Da somit die Voraussetzung auf Normalverteilung für den t-test bei einer Stichprobe verletzt ist, wird auf den nonparametrischen Wilcoxon-Test zurückgegriffen. Der Wilcoxon-Test ergibt eine Signifikanz von $p = 0.004$, sodass eine positive Abweichung von dem erwarteten Mittelwert 2 in Richtung 1 (stimmt genau) vorliegt. Die Nullhypothese wird – aufgrund der positiven Abweichung vom erwarteten Wert – zu Gunsten der Alternativhypothese verworfen.

ERGÄNZENDE ERGEBNISSE ZUR HYPOTHESE 10

Folgende Tabelle 72 gibt einen ergänzenden Überblick über die sechs Subskalen der Skala „Angemessenheit“. Die höchste Zustimmung erfährt hier der Einsatz des Verhaltensfeedbacks, gefolgt von dem formativen Leistungsfeedback (Lernverlaufsdiagnostik), dem Advance-Organizer und dem Einsatz der Arbeitsblätter. Zustimmungswerte größer als 2 erhalten die Direkte Instruktion und die peer-mediierten Lernformen (peergestütztes und kooperatives Lernen). Die umfangreichen Ergebnisse auf Itemebene finden sich im Anhang 4.

TABELLE 72 ERGEBNISSE DER SKALA „ANGEMESSENHEIT“ DES IMPLEMENTATIONSFRAGEBOGENS DER SCHÜLERINNEN UND SCHÜLER AUF SUBSKALENEBENE (DESKRIPTIV)

Variable	Operationalisierung	Anzahl Items	Kürzel	M	SD
Angemessenheit	• Lernförderlicher Einsatz der Arbeitsblätter	6	AN_{AB}	1,94	0,55
	• Lernförderlicher Einsatz des Advance Organizers	2	AN_{AO}	1,83	0,84
	• Lernförderlicher Einsatz der Direkten Instruktion	2	AN_{DI}	2,15	0,86
	• Lernförderlicher Einsatz des formativen Leistungsfeedbacks	2	AN_{CBM}	1,76	0,85
	• Lernförderlicher Einsatz der peer-medi-ierten Lernformen	2	AN_{PM}	2,37	0,92
	• Lernförderlicher Einsatz des Verhal-tensfeedbacks	4	AN_{VF}	1,68	0,48

9.6.2 Implementationsqualität aus Sicht der Interventionsleiterinnen und-leiter

Der Implementationsfragebogen wird im Anschluss an die Durchführung der Maßnahme von allen 13 Interventionsleiterinnen und -leitern (100 %) ausgefüllt. Der Implementationsfragebogen findet sich im Anhang 1. Deskriptive Ergebnisse zu allen Items des Implementationsfragebogens finden sich im Detail im Anhang 2.

Die Abbildung 64 zeigt deskriptiv die durchschnittliche Zustimmung zu den fünf Implementationsskalen Akzeptanz, Übernahme, Angemessenheit, Machbarkeit und Wiedergabetreue aus Sicht der Interventionsleiterinnen und -leiter.

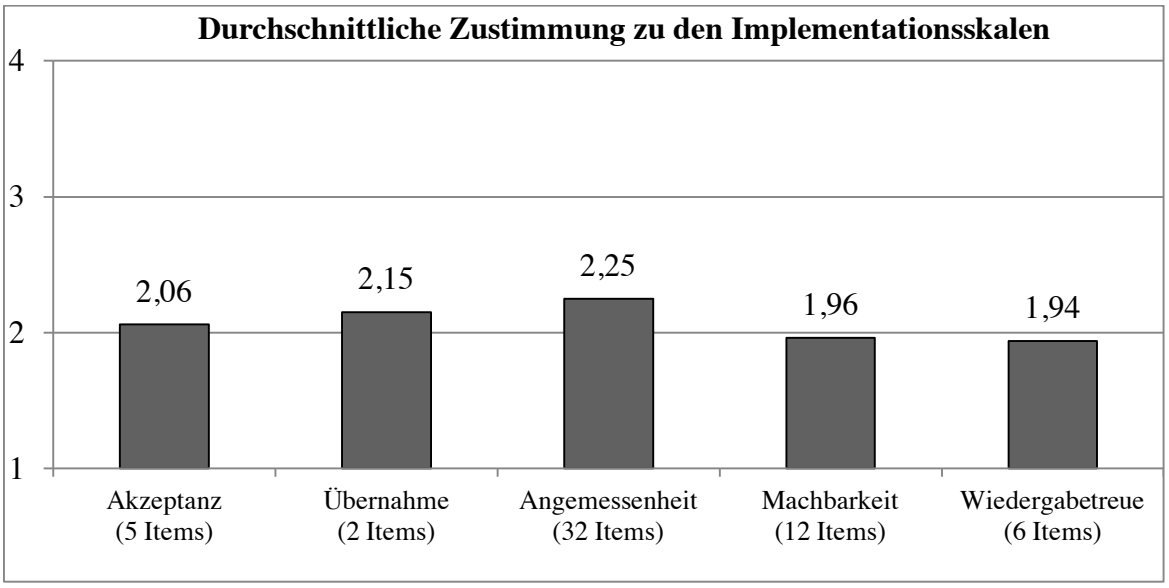


ABBILDUNG 64 DIE DURCHSCHNITTLICHE ZUSTIMMUNG ZU DEN IMPLEMENTATIONSSKALEN AUS SICHT DER INTERVENTIONSLEITERINNEN UND -LEITER

Bei den zugrundeliegenden Items der Subskalen handelt es sich um positiv gerichtete Aussagen. Die Antwortskala umfasst 1 = stimmt genau, 2 = stimmt ziemlich, 3 = stimmt kaum und 4 = stimmt gar nicht. Die Ergebnisse werden im Folgenden inferenzstatistisch analysiert.

Ergebnisse zur Hypothese 11

Die Hypothese 11 nimmt an, dass die Interventionsleiterinnen und -leiter mit der Fördermaßnahme zufrieden sind und diese akzeptieren (Kriterium „acceptability" nach Petermann, 2015b, S. 123). Sie stimmen den fünf Aussagen zur „Akzeptanz" insgesamt (AK) „ziemlich" zu (entspricht dem Skalenwert 2 von 4). Ausgeschlossen ist hier die Frage zur Notenvergabe (AK_{Note}), da diese Frage ein abweichendes Skalenniveau aufweist (Schulnote von 1–6).

Die Prüfung auf Normalverteilung zeigt, dass die Skala „Akzeptanz" normalverteilt ist ($p_{K\text{-}S}$ = 0.058). Detailangaben dazu finden sich im Anhang 7.

Der Mittelwert der Skala Akzeptanz über alle fünf Aussagen hinweg liegt bei 2,06 (SD: 0,386) (siehe Abbildung 64). Der t-Test bei einer Stichprobe ergibt einen Signifikanzwert von $p = 0.576$, sodass der Mittelwert nicht von dem erwarteten Wert 2 abweicht. Die Nullhypothese wird zu Gunsten der Alternativhypothese verworfen.

Ergänzende Ergebnisse zur Hypothese 11

Folgende Tabelle 73 gibt einen ergänzenden Überblick über die Werte der Subskala „Akzeptanz". Neben der Zufriedenheit mit einem Wert von 2,06 ist die Schulnote angegeben, die die Interventionsleiterinnen und -leiter der Fördermaßnahme geben. Die Note befindet sich mit einem Wert von 2,69 im guten bis befriedigenden Bereich. Deskriptive Ergebnisse zu allen Items der Skala finden sich im Detail im Anhang 2.

Tabelle 73 Ergebnisse der Skala „Akzeptanz" des Implementationsfragebogens der Interventionsleiterinnen und -leiter (deskriptiv)

Variable	Operationalisierung	Anzahl Items	Kürzel	M	SD
Akzeptanz	• Zufriedenheit insgesamt	5	AK_{ZI}	2,06	0,39
	• Geben einer Schulnote	1	AK_{Note}	2,69	0,95

Ergebnisse zur Hypothese 12

Die Hypothese 12 nimmt an, dass die Interventionsleiterinnen und -leiter die Maßnahme „übernehmen" würden (Kriterium „adoption" nach Petermann, 2015b, S. 123), das heißt, einer Implementation in die schulische Praxis positiv gegenüberstehen. Sie stimmen den zwei Aussagen zur Übernahme (Ü) „ziemlich" zu (entspricht dem Skalenwert 2 von 4). Die zugehörigen beiden Aussagen zu der Skala sind in Anhang 1 einzusehen. Die Skala Übernahme ist nicht in weitere Subskalen aufgeteilt.

Die Prüfung auf Normalverteilung zeigt, dass die Skala „Übernahme" nicht normalverteilt ist ($p_{K\text{-}S}$ = 0.006). Detailangaben dazu finden sich im Anhang 7. Da somit die Voraussetzung auf Normalverteilung für den t-Test verletzt ist, wird auf den non-parametrischen Wilcoxon-Test zurückgegriffen. Der Mittelwert der Skala Übernahme (Ü) über beide Aussagen hinweg liegt

bei 2,15 (SD: 0,689). Der non-parametrische Wilcoxon-Test ergibt einen Signifikanzwert von $p = 0.518$, sodass der Mittelwert nicht von dem erwarteten Wert 2 abweicht. Die Nullhypothese wird zu Gunsten der Alternativhypothese verworfen.

Ergebnisse zur Hypothese 13

Die Hypothese 13 nimmt an, dass die Interventionsleiterinnen und -leiter die Maßnahme hinsichtlich verschiedener Aspekte als passend empfinden (Kriterium „appropriateness" nach Petermann, 2015b, S. 123). Sie stimmen den 32 Aussagen zur Angemessenheit (AN) durchschnittlich „ziemlich" zu (entspricht dem Skalenwert 2 von 4). Die Prüfung auf Normalverteilung zeigt, dass die Skala „Angemessenheit" normalverteilt ist ($p_{K\text{-}S} = 0.200$). Detailangaben dazu finden sich im Anhang 7.

Der Mittelwert der Skala Angemessenheit (AN) über alle 32 Aussagen hinweg liegt bei 2,25 (SD: 0,285). Der t-Test ergibt einen Signifikanzwert von $p = 0.008$, sodass der Mittelwert signifikant von dem erwarteten Wert 2 abweicht. Die Nullhypothese wird angenommen.

Ergänzende Ergebnisse zur Hypothese 13

Folgende Tabelle 74 gibt einen ergänzenden Überblick über die Werte der Subskala „Angemessenheit".

Bezüglich der vier Bereiche der „Passung" erhält die Passung der Inhalte die höchste Zustimmung gefolgt von der Passung der Methoden allgemein, der Passung für die Zielgruppe und der Passung der Zeitplanung. Hinsichtlich der Lernförderlichkeit der Methoden erhält die höchste Zustimmung der Advance Organizer, gefolgt von den Spielen und dem formativen Leistungsfeedback (Lernverlaufsdiagnostik), dem Verhaltensfeedback, den Arbeitsblättern und der Direkten Instruktion. Die geringste Zustimmung erhält der lernförderliche Einsatz der peer-mediierten Methoden (peergestütztes und kooperative Lernformen).

Deskriptive Ergebnisse zu allen Items der Skala finden sich im Detail im Anhang 2.

Tabelle 74 Ergebnisse der Skala „Angemessenheit" des Implementationsfragebogens der Interventionsleiterinnen und -leiter (deskriptiv)

Variable	Operationalisierung	Anzahl Items	Kürzel	M	SD
Angemessenheit	• Passung für die Zielgruppe	9	AN_{ZG}	2,47	0,30
	• Passung der Inhalte	1	AN_{PI}	1,69	0,48
	• Passung der Methoden allgemein	1	AN_{PME}	2,10	0,64
	• Passung der Zeitplanung	4	AN_{ZP}	2,84	0,33
	• Lernförderlicher Einsatz des Advance Organizers	2	AN_{AO}	1,46	0,52
	• Lernförderlicher Einsatz der Arbeitsblätter	1	AN_{AB}	2,08	0,67
	• Lernförderlicher Einsatz der Direkten Instruktion	2	AN_{DI}	2,27	0,60
	• Lernförderlicher Einsatz der Spiele	1	AN_{SP}	1,69	0,63

Variable	Operationalisierung	Anzahl Items	Kürzel	M	SD
	• Lernförderlicher Einsatz der Peer-Mediierten Methoden	3	AN_{PM}	2,82	0,60
	• Lernförderlicher Einsatz des formativen Leistungsfeedbacks	2	AN_{CBM}	1,69	0,52
	• Lernförderlicher Einsatz des Verhaltensfeedbacks	6	AN_{VF}	1,98	0,55

ERGEBNISSE ZUR HYPOTHESE 14

Die Hypothese 14 nimmt an, dass die Interventionsleiterinnen und -leiter die Maßnahme als im schulischen Setting umsetzbar und machbar einschätzen (Kriterium „feasibility" nach Petermann, 2015b, S. 123). Sie stimmen den 12 Aussagen zu verschiedenen Aspekten der „Machbarkeit" (M) durchschnittlich „ziemlich" zu (entspricht dem Skalenwert 2 von 4). Die Prüfung auf Normalverteilung zeigt, dass die Werte der Skala „Machbarkeit" nicht normalverteilt sind ($p_{K\text{-}S} = 0.005$). Detailangaben dazu finden sich im Anhang 7.

Da somit die Voraussetzung auf Normalverteilung für den t-test verletzt ist, wird auf den non-parametrischen Wilcoxon-Test zurückgegriffen. Der Mittelwert der Skala Machbarkeit (M) über alle 12 Aussagen hinweg liegt bei 1,96 (SD: 0,335) (siehe Abbildung 64). Der non-parametrische Wilcoxon-Test ergibt einen Signifikanzwert von $p = 0.813$, sodass der Mittelwert nicht von dem erwarteten Wert 2 abweicht. Die Nullhypothese wird zu Gunsten der Alternativhypothese verworfen.

ERGÄNZENDE ERGEBNISSE ZUR HYPOTHESE 14

Folgende Tabelle 74 gibt einen ergänzenden Überblick über die Werte der Subskala „Machbarkeit". Eine hohe Zustimmung erhält darin die Eignung des Manuals gefolgt von der Umsetzbarkeit mit den persönlichen Ressorucen. Die geringste Zustimmung erhält die Umsetzbarkeit mit den Ressorucen im Feld Schule. Deskriptive Ergebnisse zu allen Items der Skala finden sich im Detail im Anhang 2.

TABELLE 75 ERGEBNISSE DER SKALA „MACHBARKEIT" DES IMPLEMENTATIONSFRAGEBOGENS DER INTERVENTIONSLEITERINNEN UND -LEITER (DESKRIPTIV)

Variable	Operationalisierung	Anzahl Items	Variable	M	SD
Machbarkeit	• Umsetzbarkeit mit den Ressourcen im Feld Schule	6	M_{SR}	2,27	0,43
	• Eignung des Manuals	5	M_{M}	1,62	0,38
	• Umsetzbarkeit mit den persönlichen Ressourcen	1	M_{PR}	1,85	0,55

ERGEBNISSE ZUR HYPOTHESE 15

Die Hypothese 15 nimmt an, dass die Interventionsleiterinnen und -leiter die Maßnahme entsprechend der Vorgaben umsetzen (Kriterium „fidelity" nach Petermann, 2015b, S. 123). Sie

stimmen den fünf Aussagen zur Wiedergabetreue (W) durchschnittlich „ziemlich“ zu (entspricht dem Skalenwert 2 von 4). Die Skala Wiedergabetreue ist nicht in weitere Subskalen aufgeteilt.

Die Prüfung auf Normalverteilung zeigt, dass die Werte der Skala „Wiedergabetreue“ normalverteilt sind ($p_{K\text{-}S} = 0.200$). Detailangaben dazu finden sich im Anhang 7. Der Mittelwert der Skala Wiedergabetreue (W) über alle fünf Aussagen hinweg liegt bei 1,94 (SD: 0,435). Der t-test ergibt einen Signifikanzwert von $p = 0.619$, sodass der Mittelwert nicht signifikant von dem erwarteten Wert 2 abweicht. Die Nullhypothese wird zu Gunsten der Alternativhypothese verworfen. Deskriptive Ergebnisse zu allen Items der Skala finden sich im Detail im Anhang 2.

9.6.3 Ergebnisse des Hypothesenkomplexes IV im Überblick

Die folgende Tabelle 76 zeigt die zentralen Ergebnisse zur Implementationsqualität aus Sicht der Schülerinnen und Schüler und der Interventionsleiterinnen und -leiter. Liegt die durchschnittliche Zustimmung zu einer Skala nach inferenzstatistischer Prüfung bei 2 oder darunter (< 2), wird die Nullhypothese zu Gunsten der Alternativhypothese abgelehnt. Liegt der Wert der durchschnittlichen Zustimmung jedoch über 2, wird die Nullhypothese angenommen. Die entscheidenden Kennwerte dafür sind entsprechend fett markiert.

TABELLE 76 ERGEBNISSE DES HYPOTHESENKOMPLEXES IV IM TABELLARISCHEN ÜBERBLICK

	n	Hypo-these	Skala	An-zahl Items	M	SD	p_t	p_w	H0 verworfen/beibehalten
Schülerinnen und Schüler	250	H_9	Akzeptanz	16	2,11	0,71	–	**0.106**	verworfen
		H_{10}	Angemessenheit	16	**1,93**	0,47	–	**0.004**	verworfen
Interventionsleiterinnen und -leiter	13	H_{11}	Akzeptanz	5	2,06	0,39	**0.576**	–	verworfen
		H_{12}	Übernahme	2	2,15	0,69	–	**0.518**	verworfen
		H_{13}	Angemessenheit	32	2,25	0,29	**0.008**	–	beibehalten
		H_{14}	Machbarkeit	12	1,96	0,34	–	**0.813**	verworfen
		H_{15}	Wiedergabe-treue	5	1,94	0,44	**0.619**	–	verworfen

Anmerkungen: p_t = t-test bei einer Stichprobe (Normalverteilung liegt vor), p_w = Wilcoxon-Test (Normalverteilung liegt nicht vor); p_t und p_w beziehen sich auf die Abweichung vom Erwartungswert 2 (entspricht: „stimmt ziemlich“).

10 Interpretation und Diskussion

Die Ergebnisse der vorliegenden Untersuchung sind im Kapitel 9 hypothesengeleitet dargelegt. Es folgt nun die Interpretation sowie Diskussion der Ergebnisse (Kapitel 10.1) und die kritische Analyse der Forschungsmethodik (Kapitel 10.2).

10.1 Interpretation und Diskussion der Ergebnisse

10.1.1 Hypothesenblock I und II

HYPOTHESE 1A UND 1B: WIRKSAMKEIT BEI DER GESAMTGRUPPE

Die Hypothese 1a beinhaltet die Annahme, dass die Schülerinnen und Schüler der Experimentalgruppe im Prä-Post-Vergleich eine bessere mathematische Entwicklung zeigen als diejenigen der Kontrollgruppe. Der Interaktionseffekt in der zweifaktoriellen Varianzanalyse mit Messwiederholung nimmt einen signifikanten Wert an ($F(1) = 15.796$; $p = 0.000$; $\eta^2 = 0.027$). Die Umrechnung des η^2 des Interaktionseffekts in Cohen's d führt zu einer kleinen Effektstärke von $d = 0.333$. Die Schülerinnen und Schüler der Experimentalgruppe entwickeln sich im Prä-Post-Vergleich signifikant besser als diejenigen der Kontrollgruppe. *Für die Hypothese 1a wird die Nullhypothese zu Gunsten der Alternativhypothese verworfen.*

Die Hypothese 1b betrachtet die langfristige Entwicklung und nimmt an, dass die Schülerinnen und Schüler der Experimentalgruppe im Zeitraum zwischen der Post-, Follow-Up-I- und Follow-Up-II-Erhebung eine bessere Entwicklung in den mathematischen Basisompetenzen zeigen als die Schülerinnen und Schüler der Kontrollgruppe. Auch in der Hypothese 1b ist der Interaktionseffekt signifikant ($F(2) = 16.508$; $p = 0.000$; $\eta^2 = 0.034$), sodass sich die Entwicklung der mathematischen Basiskompetenzen über die drei Messzeitpunkte in Abhängigkeit von der Gruppenzugehörigkeit voneinander unterscheidet. Detailanalysen zeigen jedoch, dass die Schülerinnen und Schüler der Experimentalgruppe zwischen der Post- und Follow-Up-I-Erhebung im Vergleich zu denjenigen der Kontrollgruppe einen signifikanten Rückgang in ihren mathematischen Leistungen zeigen ($p_t = 0.000$; $d_{korr} = -0.281$). Auch die Entwicklung vom Follow-Up-I- zum Follow-Up-II-Messzeitpunkt unterscheidet sich signifikant, wobei hier wiederum die Experimentalgruppe eine bessere Entwicklung aufweist ($p_t = 0.000$; $d_{korr} = 0.179$). Innerhalb des Follow-Up-II-Messzeitpunkts zeigen sich keine signifikanten Unterschiede zwischen den beiden Gruppen ($p_t = 0.076$, $d = 0.125$). Auch der direkte Vergleich der Entwicklung vom Post- zum Follow-Up-II-Messzeitpunkt ergibt keinen Vorsprung der Experimentalgruppe ($p_t = 0.048$; $d = -0.103$). Es können für die Gesamtgruppe keine langfristigen Fördererfolge festgestellt werden. *Es wird für die Hypothese 1b die Nullhypothese angenommen.*

Nach Hattie und Kollegen (2013, S. 21) sind ab einer Effektstärke von $d = 0.40$ Veränderungen durch eine Förderung im schulischen Alltag sichtbar. Da sich die Effektstärke der vorliegenden Maßnahme unter dieser Grenze verortet, ist die praktische Relevanz der vorliegenden Fördermaßnahme bezogen auf die Gesamtgruppe kurzfristig als eher gering einzustufen.

Da in Deutschland bisher keine universelle Maßnahme zur Förderung mathematischer Kompetenzen zu Beginn der Sekundarstufe I vorliegt, ist die Einordnung bzw. der Vergleich der Stärke des Fördereffekts schwierig. Die Programme „Rechenspiele mit Elfe und Mathis II“ (Lenhard et al., 2010) und „Calcularis“ (Von Aster et al., 2016) beinhalten auf indizierter Ebene eine Förderung der – wie hier definierten – mathematischen Basiskompetenzen in der Sekundarstufe I, sodass ein grundsätzlicher Vergleich mit diesen Programmen legitim ist. Ein Vergleich mit den Programmen „Kalkulie“ (Gerlach et al., 2007) und „Mathe sicher können“ (Prediger et al., 2014; Selter et al., 2014) hingegen ist aufgrund ausstehender Evaluationsergebnisse nicht möglich.

Evaluationsergebnisse der „Rechenspiele mit Elfe und Mathis II“ im Prä-Post-Design ergeben nach einer acht- bis zehnwöchigen Zeitspanne, in der einmal in der Woche eine Einzelförderung über den Zeitraum einer Schulstunde stattfindet, kleine Effektstärken von $d = .33$ (Jahrgangsstufe drei) und $d = .19$ (Jahrgangsstufe vier). Das Förderprogramm „Calcularis“ führt bei einer Einzelförderung fünfmal in der Woche über 20 Minuten zwischen der zweiten und fünften Klasse über zwölf Wochen zu einer sehr kleinen Effektstärke von $d = .16$ (Addition im HRT), einer kleinen Effektstärke von $d = .31$ (Addition im Computertest) bzw. einer mittleren Effektstärke von $d = .52$ (Subtraktion im HRT) und $d = .39$ (Subtraktion im Computertest). Die vorliegende Untersuchung erzielt im Prä-Post-Vergleich für die Gesamtgruppe eine Effektstärke von $d = 0.333$ und reiht sich damit tendenziell in die Wirksamkeit anderer deutschsprachiger Programme zur Förderung mathematischer Kompetenzen in der Sekundarstufe I ein. Dabei muss beachtet werden, dass es sich bei der vorliegenden untersuchten Maßnahme um eine *universelle* Förderung handelt und diese nach Ise und Kollegen (2012, S. 189) im Vergleich zur Einzel- oder Kleingruppenförderung sogar eine geringere Wirksamkeit erwarten lässt.

Wird die Förderdauer in den o. a. Evaluationen der Programme betrachtet („Calcularis“: 1.200 Minuten, „Rechenspiele mit Elfe und Mathis II“: 450 Minuten, vorliegende Untersuchung: 720 Minuten), zeigt sich, dass eine längere Förderdauer offenbar *nicht* zwingend mit höheren Effekten zusammenhängt und die hier gewählte Förderdauer von 720 Minuten einen durchschnittlichen Umfang darstellt. Die hier und bei „Calcularis“ vorliegenden kleinen Effekte entsprechen dabei nicht den Befunden von Ise und Kollegen (2012, S. 189), die ab einem Förderumfang von 600 Minuten eigentlich einen großen Effekt erwarten.

Bezüglich der langfristigen Wirksamkeit zeigen die Ergebnisse zur Hypothese 1b, dass diese nicht eingetreten ist, da die Experimentalgruppe nach Abschluss der Maßnahme einen signifikanten Leistungsrückgang mit einer kleinen Effektstärke von $d_{korr} = -0.281$ verzeichnet. Die Power von 1,00 legitimiert die Annahme der Nullhypothese. Der Zuwachs um 11,94 Rohwertpunkte zwischen der Prä- und Post-Erhebung ist gefolgt von einem Rückgang um 3,59 Rohwertpunkte zwischen der Post- und Follow-Up-I-Erhebung. Dass die am Ende einer Intervention erzielten Zuwächse langfristig nicht aufrechterhalten oder gesteigert werden können, ist im Rahmen von Interventionsstudien als ein möglicher Ausgang zu betrachten und zu diskutieren (Hagenauer, 2010, S. 247). Auch von Aster und Kollegen (2016, S. 242) stellen in der Evaluation des Programms „Calcularis“ einen Leistungsrückgang nach der Interventionsphase fest. Sie begründen dies mit einem zu kurzen Trainingszeitraum (6 Wochen). Für die vorliegende Untersuchung sind weitere Gründe für den zwischenzeitigen Leistungsrückgang in Betracht zu ziehen: Die in der Maßnahme geförderten mathematischen Basiskompetenzen fungieren als „notwendige Voraussetzung“ (Ehlert et al., 2013, S. 239) für den weiteren mathematischen

Kompetenzerwerb. Das Verständnis der systematischen Rechenoperationen und das Regelwissen bilden die Grundlage für komplexe Rechenoperationen im weiteren Mathematikunterricht (Gaidoschik, 2008; Grube, 2005; Moser Opitz, 2007). Nach dem Abschluss der Maßnahme stehen die erlernten Inhalte im Mathematikunterricht jedoch nicht mehr primär im Vordergrund und werden von weiterführenden curricularen Inhalten (zum Beispiel Prozent- oder Bruchrechnen) abgelöst. Die neu erlernten Inhalte müssen nun mit dem aktuellen Unterrichtsstoff verknüpft werden. Dem Ende des intensiven Trainings schließt sich somit eine Phase an, in der die gelernten Rechenstrategien nun in vergleichbaren Aufgabenformaten Anwendung finden können, und eine weitere Phase, in der bisherige Vorstellungen und Zugänge ggf. auch verändert werden müssen (Berk, 2011, S. 201). Klauer (2011b, S. 195) beschreibt diesen Effekt als „Umstellung der Lernstrategie“ und formuliert, dass „negativer Lerntransfer wahrscheinlicher wird, wenn eine bereits hoch automatisierte Lernprozedur durch eine neue ersetzt werden soll, selbst wenn die neue an sich leistungsfähiger ist“. Die „Umstellung der Lernstrategien“ könnte dazu führen, dass der schnelle Faktenabruf, wie er im Heidelberger Rechentest 1–4 gefordert ist, vorübergehend nicht auf dem durch die Intervention erzielten Level erfolgen kann und der Outcome im Follow-Up-I-Messzeitpunkt deswegen kurzfristig zurückgeht.

Erschwert wird die Übergangsphase in den regulären Mathematikunterricht nach der Förderung zudem durch den Wechsel der unterrichtenden Lehrkraft, da nun die Mathematiklehrkraft den Unterricht leitet und nicht mehr die Interventionsleiterinnen und -leiter. Dieser Angewöhnungsprozess ist begleitet von – denen der Primarstufe abweichenden – Unterrichtsprinzipien wie zum Beispiel dem Fachlehrerprinzip. Dies wird u. a. im Hagener Übergangsprojekt als die Transition erschwerend identifiziert (Mitzlaff & Wiederhold, 1989). Dass sich die Experimentalgruppe in der Entwicklung vom Follow-Up-I- zum Follow-Up-II-Messzeitpunkt wiederum signifikant stärker als die Kontrollgruppe entwickelt ($p_{\text{t}} = 0.000$; $d_{\text{korr}} = 0.179$), stützt die hier getätigten Überlegungen zum zwischenzeitigen Leistungsrückgang. Es ist anzunehmen, dass zum Schulhalbjahr die Gewöhnung an und das Vertrauen in die Mathematiklehrkraft und an die divergierenden Unterrichtsprinzipien der Sekundarstufe erfolgt sind und die in der Fördermaßnahme erlernten Rechenstrategien erfolgreich in mathematischen Schemata integriert sind. Diese Annahmen werden durch die wieder ansteigenden Werte zum Follow-Up-II gestützt.

Um eine frühzeitige gegenseitige Angewöhnung zwischen der regulären Mathematiklehrkraft und der Lerngruppe direkt im Anschluss an die Transition zu ermöglichen, wäre grundsätzlich zu erwägen, die Mathematiklehrkraft selbst zu schulen und sie die Fördermaßnahme durchführen zu lassen. Ise und Kollegen (2012, S. 189) ermitteln eine höhere Effektstärke, wenn Lehrkräfte eine Intervention durchführen ($d = 0.57$), als wenn dies Studierende bzw. Projektmitarbeiterinnen und -mitarbeiter tun ($d = 0.18$). In der vorliegenden Evaluation erfolgt die Durchführung jedoch durch Studentinnen, um potenzielle Variablen, die Einfluss auf die Wirksamkeit nehmen, gering halten und kontrollieren zu können. Dazu gehört zum Beispiel eine fundierte Einarbeitung in die Fördermaßnahme und eine gute Auseinandersetzung mit den theoretischen Grundlagen sowie eine Schulung der Kompetenzen zum Thema Classroom-Management. Für eine möglichst vergleichbare Umsetzung in den Lerngruppen können auch der Ausbildungsstand und die Unterrichtserfahrung der Interventionsleiterinnen und -leiter auf diesem Wege kontrolliert werden. Die Durchführung der Förderung durch die Mathematiklehrkräfte selbst wäre in einer umrissenen Replikationsstudie zu erwägen.

Um einen zwischenzeitigen Leistungsrückgang künftig zu vermeiden, könnten darüber hinaus Auffrischungssitzungen entwickelt werden, die in regelmäßigen Abständen nach dem Abschluss der Intervention Anwendung finden. In diesen Auffrischungssitzungen könnten die Inhalte und Themen der Förderung mit konkreten Unterrichtsthemen und Aufgabenstellungen verknüpft werden. Dies könnte die langfristige Integration und Verknüpfung der Inhalte erleichtern. Ebenso wäre denkbar, die in der Intervention genutzen Materialien und Methoden zum Beispiel mit Handlungsanweisungen aufzubereiten, so dass die Mathematiklehrkraft diese auch nach Abschluss der Förderung problemlos in ihren alltäglichen Unterricht einbauen kann. So kann die Direkte Instruktion und peergestütztes bzw. kooperatives Lernen auch im Themenbereich Bruch- oder Prozentrechnen umgesetzt werden, sofern die Mathematiklehrkraft über Hinweise für eine gelingende Umsetzung verfügt. Zudem ist zu diskutieren, den Interventionszeitraum zu verlängern, wie die Autoren von „Calcularis“ vorschlagen (Von Aster et al., 2016, S. 242). Auch die Interventionsleiterinnen und -leiter empfinden den zeitlichen Umfang für die geplanten Inhalte in der vorliegenden Untersuchung als zu knapp (siehe Diskussion der Hypothese 13). Jedoch sollten in dieser Frage die konzeptionellen Grundlagen der Fördermaßnahme nicht außer Acht gelassen werden. Die Fördermaßnahme soll den Übergang und den Einstieg in den Mathematikunterricht der Sekundarstufe I erleichtern und den gesamten Grundschulunterricht nicht wiederholen bzw. ersetzen. Eine Ausweitung der universellen Förderung auf fünf oder sechs Wochen wird deswegen eher kritisch gesehen, zumal auch die sich anschließenden curicularen Inhalte der Jahrgangsstufe fünf einen ausreichenden zeitlichen Rahmen benötigen. Der Kritik der Interventionsleiterinnen und -leiter kann jedoch durch innere strukturelle Modifikationen der Maßnahme und einer weiteren Differenzierung begegnet werden, die im Verlauf der Diskussion weiter erörtert werden.

Ein überraschendes Ergebnis stellt – trotz eines nicht signifikanten Zwischensubjekteffekts – die durchweg positive Entwicklung der Kontrollgruppe über die drei Messzeitpunkte dar ($F(1) = 1.266$; $p = 0.261$; $\eta^2 = 0.003$). Forschungsbefunde zeigen, dass nach der Transition in die Sekundarstufe I die Leistungen oftmals nicht gehalten werden können (Böck & Migutsch, 1974; Eder, 1995; Isakson & Jarvis, 1999; McGee et al., 2003; Sirsch, 2000). Diese Befunde können hier – bezogen auf die Entwicklung der gesamten Kontrollgruppe – nicht bestätigt werden.

Hypothese 2a und 2b: Wirksamkeit bei einem geringen Vorwissen

Die Hypothese 2a nimmt an, dass die Schülerinnen und Schüler mit einem geringen Vorwissen in der Experimentalgruppe im Prä-Post-Vergleich eine bessere mathematische Entwicklung zeigen als diejenigen in der Kontrollgruppe. Der Interaktionseffekt in der zweifaktoriellen Varianzanalyse mit Messwiederholung nimmt einen signifikanten Wert an ($F(1) = 12.608$; $p = 0.000$; $\eta^2 = 0.067$). Die Umrechnung des η^2 des Interaktionseffekts in Cohen's d führt zu einem mittleren Effekt von $d = 0.536$. Die Schülerinnen und Schüler mit einem geringen Vorwissen in der Experimentalgruppe entwickeln sich im Prä-Post-Vergleich signifikant besser als diejenigen der Kontrollgruppe. *Für die Hypothese 2a wird die Nullhypothese zu Gunsten der Alternativhypothese verworfen.*

Die Hypothese 2b nimmt an, dass die Schülerinnen und Schüler mit einem geringen Vorwissen in der Experimentalgruppe im Zeitraum zwischen der Post-, Follow-Up-I- und Follow-Up-II-Erhebung eine bessere Entwicklung in den mathematischen Basiskompetenzen zeigen als die

Schülerinnen und Schüler der Kontrollgruppe. Auch in der Hypothese 2b ist der Interaktionseffekt mit einer kleinen Effektstärke ($d = 0.357$) signifikant ($F(1.905) = 4{,}317$; $p = 0.016$; $\eta^2 = 0.031$), sodass die Entwicklung der mathematischen Basiskompetenzen grundsätzlich wechselseitig von der Gruppenzugehörigkeit und dem Messzeitpunkt abhängt. Detailanalysen zeigen allerdings, dass die mathematischen Leistungen der Schülerinnen und Schüler der Experimentalgruppe zwischen der Post- und Follow-Up-I-Erhebung im Vergleich zu denjenigen der Kontrollgruppe mit einem kleinen Effekt signifikant zurückgehen ($p_t = 0.002$; $d_{korr} = -0.440$) und sich der darauffolgende Anstieg zum Follow-Up-II nicht signifikant unterscheidet ($p_t = 0.100$; $d_{korr} = 0.257$). Da auch die Entwicklung vom Post- zum Follow-Up-II-Messzeitpunkt zwischen den Gruppen vergleichbar ist ($p_t = 0.281$; $d = -0.158$), kann nicht von langfristigen Fördererfolgen ausgegangen werden. *Es wird für die Hypothese 2b die Nullhypothese angenommen.*

Die Schülerinnen und Schüler mit einem geringen Vorwissen profitieren von der Fördermaßnahme kurzfristig mit einer mittleren Effektstärke von $d = 0.536$. Da die Effektstärke größer ist als $d = 0.40$, handelt es sich nach Hattie und Kollegen (2013, S. 21) um einen Effekt, der sich im pädagogischen Alltag bemerkbar macht.

Konzeptionell sind für Schülerinnen und Schüler mit einem geringen Vorwissen vor allem die umgesetzten Methoden des Advance Organizers und der Direkten Instruktion lernförderlich, da sie erleichtern, neu erlerntes Wissen mit bereits bekanntem zu verknüpfen und zu strukturieren. Für eine alleinige Förderung durch die Direkte Instruktion attestieren Hattie und Kollegen (2013) allgemein eine Effektstärke von ($d \approx .50$). Speziell für den Bereich Mathematik sind die Befunde sehr unterschiedlich. Sie reichen von mittleren (Baker et al., 2002) bis zu sehr hohen Effektstärken (z. B. Gersten et al., 2009). Der hier ermittelte Effekt von $d = 0.536$ verortet sich auch im Bereich der mittleren Effektstärke. Im Rahmen der vorliegenden Untersuchung ist es jedoch nicht möglich, den Effekt auf diese spezifischen Methoden zurückzuführen. Der Einsatz der Direkten Instruktion unterstützt – theoretisch – Schülerinnen und Schüler mit einem geringen Vorwissen, doch kann der erzielte mittlere Effekt auch zum Beispiel durch das intensive Üben oder das kooperative Lernen befördert werden. Festgehalten werden kann lediglich, dass die vorliegende Fördermaßnahme bei Schülerinnen und Schülern mit geringem Vorwissen kurzfristig zu einem mittleren Effekt führt. Eine hohe Effektstärke kann nicht gefunden werden.

Eine langfristige Wirksamkeit kann bei den Schülerinnen und Schülern mit einem geringen Vorwissen nicht bestätigt werden. Interessant ist hier ein Blick ins Detail: Der Zuwachs in der Experimentalgruppe von 15,8 Rohwertpunkten zwischen dem Prä- und Post-Messzeitpunkt ist gefolgt von einem nur leichten Rückgang um nur 1,55 Rohwertpunkten zwischen dem Post- und Follow-Up-I-Messzeitpunkt. Da sich die Kontrollgruppe zwischen dem Post- und Follow-Up-I-Messzeitpunkt jedoch um 7,58 Rohwertpunkte verbessert und damit im Follow-Up-I-Messzeitpunkt sogar 1,48 Rohwertpunkte über dem Outcome der Experimentalgruppe liegt, kann die Annahme über eine langfristige Wirksamkeit der Fördermaßnahme nicht bestätigt werden. Die Schülerinnen und Schüler mit einem geringen Vorwissen, die an der Maßnahme teilnehmen, zeigen langfristig keinen Vorsprung im Vergleich zu denjenigen der Kontrollgruppe. Die Annahme der Nullhypothese ist mit einer Power von 1,0 zulässig.

Wie Tabelle 65 verdeutlicht, liegen die Schülerinnen und Schüler mit einem geringen Vorwissen in beiden Untersuchungsgruppen in nahezu allen Messzeitpunkten weit unter dem Niveau

derjenigen mit geringen kognitiven Fähigkeiten, einer geringen Lesekompetenz oder einem risikobehafteten Lern- oder Sozialverhalten. Lediglich die Schülerinnen und Schüler mit einem Unterstützungsbedarf im Bereich Lernen zeigen noch schwächere Leistungen. In Anbetracht der belastbaren Stichprobengröße scheint es sich bei den Schülerinnen und Schülern mit einem geringen Vorwissen um eine Gruppe mit einer überaus ungünstigen Lernausgangslage für mathematischen Kompetenzerwerb zu handeln. In diesem Kontext stellt die mittlere Effektstärke der kurzfristigen Wirksamkeit ein erfreuliches Ergebnis dar (siehe Abbildung 65).

Auch im Hypothesenkomplex 2 ist die positive Entwicklung der Schülerinnen und Schüler mit einem geringen Vorwissen in der Kontrollgruppe über die vier Messzeitpunkte hinweg ein Ergebnis, das aufgrund der bestehenden Befunde über einen möglichen Leistungsrückgang nach der Transition überraschend ist (Böck & Migutsch, 1974; Eder, 1995; Isakson & Jarvis, 1999; McGee et al., 2003; Sirsch, 2000). Zwar gelingt den Schülerinnen und Schülern der Experimentalgruppe kurzfristig ein Entwicklungsvorsprung, doch holen die Schülerinnen und Schüler der Kontrollgruppe diesen langfristig wieder auf, sodass sich – neben einem nicht signifikanten Zwischensubjekteffekt ($F(1) = 1.040$; $p = 0.310$; $\eta^2 = 0.008$) – auch die Leistungen am Schuljahresende nicht unterscheiden ($p_t = 0.226$, $d = 0.189$). Um eine langfristige Stabilität oder gar einen Ausbau der Fördereffekte zu ermöglichen, sind auch hier Auffrischungssitzungen denkbar, in denen die erlernten Rechenstrategien mit aktuellen Unterrichtsinhalten verknüpft werden.

Folgendes überraschendes Ergebnis soll zudem beschrieben werden: Zur Definition eines geringen Vorwissens wird in der vorliegenden Untersuchung ein T-Wert in der Prä-Erhebung des HRT 1–4 von kleiner als 40 festgesetzt. Eben diesen Cut-Off-Wert wählen die Autoren des HRT 1–4 – unter Berücksichtigung einer durchschnittlichen Intelligenz – zur Dyskalkulie-Diagnostik. Wird von der Prävalenz einer Rechenstörung von ca. 2–6 % ausgegangen (siehe Tabelle 10), sind die hier auf die Prä-Erhebung bezogenen Quoten (ohne Berücksichtigung des Diskrepanzkriteriums) von 28,85 % (EG) bzw. 30,85 % (KG) als überraschend und besorgniserregend anzusehen, zumal sogar ein Vergleich mit der Normstichprobe am Ende der Klasse 4 zugrundeliegt.

Hypothese 3a und 3b: Wirksamkeit bei geringen kognitiven Fähigkeiten

Die Hypothese 3a beinhaltet die Annahme, dass die Schülerinnen und Schüler mit geringen kognitiven Fähigkeiten in der Experimentalgruppe im Prä-Post-Vergleich eine bessere mathematische Entwicklung zeigen als diejenigen in der Kontrollgruppe. Der Interaktionseffekt in der zweifaktoriellen Varianzanalyse mit Messwiederholung nimmt einen signifikanten Wert an ($F(1) = 7.686$; $p = 0.007$; $\eta^2 = 0.068$). Die Umrechnung des η^2 des Interaktionseffekts in Cohen's d führt zu einem mittleren Effekt von $d = 0.540$. Die Schülerinnen und Schüler mit geringen kognitiven Fähigkeiten in der Experimentalgruppe entwickeln sich im Prä-Post-Vergleich signifikant besser als diejenigen der Kontrollgruppe. *Die Nullhypothese wird zugunsten der Alternativhypothese verworfen.*

Die Hypothese 3b nimmt an, dass die Schülerinnen und Schüler mit geringen kognitiven Fähigkeiten in der Experimentalgruppe im Zeitraum zwischen der Post-, Follow-Up-I- und Follow-Up-II Erhebung eine bessere Entwicklung in den mathematischen Basiskompetenzen zeigen als die Schülerinnen und Schüler der Kontrollgruppe. Für die Hypothese 3b liegt kein

signifikanter Interaktionseffekt vor ($F(2) = 2.037$; $p = 0.134$; $\eta^2 = 0.025$). Auch der Zwischensubjekteffekt ist nicht signifikant ($F(1) = 2.020$; $p = 0.159$; $\eta^2 = 0.025$), sodass nicht von einer Stabilisierung des erzielten Vorsprungs (Prä-Post) ausgegangen werden kann. Die mathematischen Basiskompetenzen der Experimentalgruppe entwickeln sich langfristig nicht unterschiedlich zur Kontrollgruppe, sodass keine langfristigen Fördererfolge zu erkennen sind. *Es wird die Nullhypothese angenommen.*

Trotz ausbleibender Signifikanz liegt die Effektstärke der langfristigen Wirksamkeit ($d = 0.540$) in einem Wertebereich, der eine Sichtbarkeit auch im pädagogischen Alltag erwarten lässt ($d \geq 0.4$) (Hattie et al., 2013, S. 21). Unterstützt werden sollen die Schülerinnen und Schüler mit geringen kognitiven Fähigkeiten in der Fördermaßnahme gerade durch die Entlastung des Arbeitsgedächtnisses in Form der Förderung des Strategiewissens mittels klarer und eindeutiger Instruktionen. Diese erfolgen vor allem durch die Methode der Direkten Instruktion, aber auch durch die Darbietung von Beispielen und Strategie-Checklisten auf den Arbeitsblättern. In der vorliegenden Untersuchung kann die zielgruppenspezifische Wirksamkeit dieser einzelnen Elemente jedoch nicht ermittelt und bestätigt werden.

Bezüglich der langfristigen Wirksamkeit zeigt sich eine vergleichbare Entwicklung wie bei den Schülerinnen und Schülern mit einem geringen Vorwissen (siehe Hypothesenblock 2). Die Leistungen in der Experimentalgruppe gehen auch hier zum Follow-Up-I-Messzeitpunkt um 1,66 Rohwertpunkte zurück; in der Experimentalgruppe steigt in dieser Zeitspanne der Outcome um 6,31 Rohwertpunkte. Die Leistungen der Experimentalgruppe bleiben hier deskriptiv zwar in jedem Messzeitpunkt über denen der Kontrollgruppe, was grundsätzlich für einen persistenten Entwicklungsvorsprung spricht – jedoch ist der Zwischensubjekteffekt nicht signifikant, sodass sich die Leistungsentwicklung der beiden Gruppen langfristig nicht unterscheidet. In der Hypothese 3b muss zudem berücksichtigt werden, dass der Stichprobenumfang klein ist. Die vorliegende Stichprobengröße beträgt lediglich $n_{EG} = 38$ und $n_{KG} = 43$. Die ermittelte Teststärke von 0,999 lässt die Annahme der Nullhypothese jedoch zu. Bezüglich der Schülerinnen und Schüler mit geringen kognitiven Kompetenzen ist eine Replik mit einer größeren Stichprobe angezeigt.

Auch hier weisen die Schülerinnen und Schüler mit geringen kognitiven Fähigkeiten aus der Kontrollgruppe eine durchweg positive Entwicklung auf, was als überraschendes Ergebnis interpretiert werden kann. Es gelingt ihnen, den Entwicklungsvorsprung der Schülerinnen und Schüler der Experimentalgruppe im Follow-Up-I zu verringern, jedoch nicht aufzuholen. Am Ende des Schuljahres liegen sie im Follow-Up-II deskriptiv immer noch unter dem erreichten Niveau der Experimentalgruppe zum Zeitpunkt des *Post*-Messzeitpunkts. Diese Befunde müssen jedoch als zufällig tituliert werden, da sie inferenzstatistisch nicht abgesichert werden können.

Einschränkend muss angeführt werden, dass die Schülerinnen und Schüler mit geringen kognitiven Fähigkeiten hier ein unterdurchschnittlich ausgeprägtes *logisches Denken* aufweisen, da die kognitiven Fähigkeiten in der vorliegenden Studie mit dem CFT 20-R (längere Testzeit) erhoben werden, der die sog. „fluide Intelligenz" misst. Es liegen hier *keine* Informationen zur Intelligenz im mehrdimensionalen Sinne vor, die zum Beispiel auch das Sprachverständnis, die Arbeitsgedächtniskapazitäten, die Verarbeitungsgeschwindigkeit oder das Allgemeinwissen umfasst und zum Beispiel im WISC-IV berücksichtigt wird. Somit kann *keine* Aussage bzgl. der Wirksamkeit in Abhängigkeit dieses mehrdimensionalenn Intelligenzkonstrukts getätigt

werden, sondern nur in Bezug auf die fluide Intelligenz. In einer Replikation wäre der Einsatz eines Instruments, das die sog. mehrdimensionale Intelligenz erfasst, zu erwägen. Gerade in Bezug auf die anvisierte Entlastung des Arbeitsgedächtnisses durch die Förderung sind hier differenzierte Ergebnisse aufschlussreich.

Hypothese 4a und 4b: Wirksamkeit bei einer geringen Lesekompetenz

Die Hypothese 4a nimmt an, dass die Schülerinnen und Schüler mit einer geringen Lesekompetenz in der Experimentalgruppe im Prä-Post-Vergleich eine bessere mathematische Entwicklung zeigen als diejenigen in der Kontrollgruppe. Der Interaktionseffekt in der zweifaktoriellen Varianzanalyse mit Messwiederholung nimmt einen signifikanten Wert an ($F(1) = 714.675$; $p = 0.000$; $\eta^2 = 0.043$). Die Umrechnung des η^2 des Interaktionseffekts in Cohen's d führt zu einem kleinen Effekt von $d = 0.424$. Die Schülerinnen und Schüler mit einer geringen Lesekompetenz in der Experimentalgruppe entwickeln sich im Prä-Post-Vergleich signifikant besser als diejenigen der Kontrollgruppe. *Die Nullhypothese wird zu Gunsten der Alternativhypothese verworfen.*

Die Hypothese 4b nimmt an, dass die Schülerinnen und Schüler mit einer geringen Lesekompetenz in der Experimentalgruppe im Zeitraum zwischen der Post-, Follow-Up-I- und Follow-Up-II-Erhebung eine bessere Entwicklung in den mathematischen Basiskompetenzen zeigen als die Schülerinnen und Schüler der Kontrollgruppe. Auch in der Hypothese 4b ist der Interaktionseffekt signifikant ($F(1.875) = 12.776$; $p = 0.000$; $\eta^2 = 0.049$), sodass die langfristige Entwicklung der mathematischen Basiskompetenzen grundsätzlich wechselseitig von der Gruppenzugehörigkeit und dem Messzeitpunkt abhängt. Detailanalysen zeigen auch hier, dass die mathematischen Leistungen der Schülerinnen und Schüler der Experimentalgruppe zwischen der Post- und Follow-Up-I-Erhebung im Vergleich zu denjenigen der Kontrollgruppe mit einem kleinen Effekt signifikant zurückgehen ($p_t = 0.000$; $d_{korr} = -0.350$), danach zum Follow-Up-II jedoch auch mit einem kleinen Effekt signifikant wieder ansteigen und dabei die Experimentalgruppe eine stärkere Entwicklung zeigt ($p_t = 0.003$; $d_{korr} = 0.217$). Auch innerhalb des Follow-Up-II-Messzeitpunkts zeigt sich ein signifikanter Unterschied zwischen den beiden Gruppen ($p_t = 0.004$, $d = 0.354$). Da der Haupteffekt „Gruppenzugehörigkeit" (Zwischensubjekteffekt) ebenso signifikant ist und die Experimentalgruppe in den drei Messzeitpunkten ein signifikant höheres Kompetenzniveau verzeichnet als die Kontrollgruppe, weisen die Ergebnisse darauf hin, dass trotz eines zwischenzeitigen Rückgangs zwischen Post und Follow-Up-I ein langfristiger Entwicklungsvorsprung der Experimentalgruppe besteht. *Es wird die Nullhypothese zu Gunsten der Alternativhypothese verworfen.*

Eine oftmals geäußerte Kritik an Diagnostik und Förderung bei Schülerinnen und Schülern mit geringen Lesekompetenzen ist die Hürde des Leseverständnisses für zum Beispiel mathematische Inhalte (Landerl & Fussenegger, 2006). Im Rahmen der hier vorliegenden Fördermaßnahme wird der schriftsprachliche Anteil insgesamt sehr gering gehalten. Die Rechenstrategien werden zu Beginn der Unterrichtsstunde in Form Direkter Instruktion an der Tafel präsentiert und mündlich erarbeitet. Auch bei der Hinführung mit dem Advance Organizer wird nahezu auf die Verwendung von Schriftsprache verzichtet. Die darauffolgenden Arbeitsphasen dienen vornehmlich der Förderung der Rechenstrategien, was zwar nicht gänzlich ohne die Verwendung von Schriftsprache auskommt (zum Beispiel Instruktionen auf den Arbeitsblättern oder beim peergestützten Lernen) (siehe Abbildung 40), doch ist der Anteil insgesamt gering. Dies

kann eine weniger diskriminierende, dafür aber intensive Förderung mathematischer Basiskompetenzen der Schülerinnen und Schüler mit unzureichenden Lesekompetenzen ermöglichen, sodass sie sich in ihren mathematischen Basiskompetenzen signifikant verbessern können. Abschließend bestätigt werden kann diese Überlegung jedoch nicht.

Auch nach dem Abschluss der Fördermaßnahme zeichnet sich ein persistenter Entwicklungsvorsprung ab. Auch wenn in der Experimentalgruppe ein zwischenzeitiger signifikanter Leistungsrückgang um 4,25 Rohwertpunkte erfolgt, steigen die Leistungen zum Follow-Up-II wieder um 10,11 Rohwertpunkte an. Der signifikante Zwischensubjekteffekt zeigt dabei an, dass sich die Experimentalgruppe über die drei Messzeitpunkte hinweg auf einem stärkeren Kompetenzniveau befindet als die Kontrollgruppe. Der durch die Fördermaßnahme erzielte Vorsprung kann langfristig aufrecht erhalten bleiben, sodass die Nullhypothese verworfen wird.

Im Rahmen der Förderung wird zur grundsätzlichen Entlastung des Arbeitsgedächtnisses die Cognitive-Load-Theory hinzugezogen, die dem Arbeitsgedächtnis eine zentrale Rolle mit begrenzter Kapazität für die Prozesse der Informationsverarbeitung zuschreibt (Ayres & Paas, 2012) und Hinweise auf dessen Entlastung gibt. Bei Schülerinnen und Schüler mit geringen Lesekompetenzen zeigen Befunde, dass es deutliche Zusammenhänge zwischen den Lesekompetenzen und der Fähigkeit zum phonologischen Rekodieren im Arbeitsgedächtnis gibt (Klicpera, Schabmann & Gasteiger-Klicpera, 2013). Die Befundlage bzgl. vergleichbarer Ursachen von Lese- und Rechenschwierigkeiten ist jedoch derzeit noch wenig einheitlich. Doch könnten die vorliegenden positiven Befunde dafür sprechen, dass die Entlastung des Arbeitsgedächtnisses auch förderlich für die mathematische Kompetenzentwicklung von Schülerinnen und Schülern mit geringer Lesekompetenz ist. Da jedoch keine Arbeitsgedächtniskapazitäten erhoben werden (siehe Diskussion der Hypothese 1a), kann diese Überlegung abschließend nicht bestätigt werden.

In diesem Hypothesenkomplex vier sind die hohen Prävalenzen der Schülerinnen und Schüler mit Leseschwierigkeiten auffallend. In der Experimentalgruppe weisen in der Prä-Erhebung 60,00 % und in der Kontrollgruppe 52,66 % der Schülerinnen und Schüler einen (unterdurchschnittlichen) Lesequotienten von < 90 auf, zumal die Referenzgruppe Schülerinnen und Schüler am Ende der Klasse 4 darstellt. Zudem weisen die Schülerinnen und Schüler mit einer geringen Lesekompetenz – wie in Tabelle 65 erkenntlich – im Vergleich zu den anderen Gruppen mit einer ungünstigen Lernausgangslage in der Prä-Testung die höchsten Werte im HRT 1–4 auf. Die Ausgangswerte sind dabei auch nur etwas geringer als die der Gesamtgruppe. Es wird festgestellt, dass es sich folglich nicht um *die* Gruppe mit der ungünstigsten Lernausgangslage handelt. Dies widerspricht den Befunden, die von massiven Rechenschwierigkeiten bei geringen Lesekompetenzen berichten (siehe Landerl & Fussenegger, 2006).

Die verhältnismäßig gute Lernausgangslage kann durch die Förderung noch weiter verstärkt werden. Die Schülerinnen und Schüler der Experimentalgruppe verbessern sich kurzfristig um 13,53 Rohwertpunkte, diejenigen der Kontrollgruppe um 7,40 Rohwertpunkte. Die Effektstärken des Interaktionseffekts liegen kurz- und langfristig über $d = .40$, was auf einen praktisch relevanten Fördereffekt hinweist (Hattie et al., 2013, S. 21).

Im Rahmen weiterer Analysen sollte vertiefend die Wirksamkeit bei Schülerinnen und Schüler mit *komorbiden* Lese- und Rechenschwierigkeiten fokussiert werden.

Hypothese 5a und 5b: Wirksamkeit bei einem risikobehafteten Sozialverhalten

Die Hypothese 5a nimmt an, dass die Schülerinnen und Schüler mit einem risikobehafteten Sozialverhalten in der Experimentalgruppe im Prä-Post-Vergleich eine bessere mathematische Entwicklung zeigen als diejenigen in der Kontrollgruppe. Der Interaktionseffekt in der zweifaktoriellen Varianzanalyse mit Messwiederholung nimmt keinen signifikanten Wert an ($F(1) = 2.561$; $p = 0.119$; $\eta^2 = 0.070$). Die Schülerinnen und Schüler mit einem risikobehafteten Sozialverhalten in der Experimentalgruppe entwickeln sich im Prä-Post-Vergleich nicht besser als diejenigen der Kontrollgruppe. *Die Nullhypothese wird angenommen.*

Die Hypothese 5b nimmt an, dass die Schülerinnen und Schüler mit einem risikobehafteten Sozialverhalten in der Experimentalgruppe im Zeitraum zwischen der Post-, Follow-Up-I- und Follow-Up-II-Erhebung eine bessere Entwicklung in den mathematischen Basiskompetenzen zeigen als die Schülerinnen und Schüler der Kontrollgruppe. Für die Hypothese 5b liegt kein signifikanter Interaktionseffekt vor ($F(2) = 1.646$; $p = 0.205$; $\eta^2 = 0.073$). Die mathematischen Basiskompetenzen der Experimentalgruppe entwickeln sich in wechselseitiger Abhängigkeit von der Gruppenzugehörigkeit und den Messzeitpunkten langfristig *nicht* besser als die der Kontrollgruppe, sodass keine langfristigen Fördererfolge zu erkennen sind. *Die Nullhypothese wird angenommen.*

In absoluten Werten profitieren die Schülerinnen und Schüler der Experimentalgruppe kurzfristig von der Maßnahme, da sie in der Post-Messung durchschnittlich zehn Aufgaben mehr richtig lösen. Dieser Zuwachs kann inferenzstatistisch jedoch nicht abgesichert werden, entspricht aber einer mittleren Effektstärke von $d = 0.549$. Die Effektstärke der kurzfristigen Wirksamkeit weist auf einen praktische relevanten Effekt hin (Hattie et al., 2013, S. 21).

Beachtet werden muss in der Interpretation der Ergebnisse die geringe Stichprobengröße. Im Prä-Post-Vergleich sind nur 17 (EG) bzw. 19 (KG) Schülerinnen und Schüler in die Analysen einbezogen. Zugleich sind die Standardabweichungen in den gemessenen Werten sehr hoch, sodass eine Überschneidung der Konfidenzintervalle wahrscheinlicher wird. Die Analyse der Stichprobenumfangsplanung a priori mit dem Programm G*Power ergibt, dass die vorhandene Stichprobe grundsätzlich groß genug ist, einen mittleren Effekt bei einer Teststärke von 80 % aufzudecken.

In beiden Untersuchungsgruppen zeigen die Schülerinnen und Schüler mit einem risikobehafteten Sozialverhalten in der Prä-Testung des HRT 1–4 im Vergleich zu den anderen Gruppen mit ungünstigen Lernvoraussetzungen sehr niedrigere mathematische Leistungen (siehe Tabelle 65). Die Befunde können hier folglich einen Zusammenhang zwischen einem risikobehafteten Sozialverhalten und Schwierigkeiten im Bereich Mathematik bestätigen (Linderkamp & Grünke, 2007; Ricken, 2014; Rolus-Borgward, 1997). Auch wenn kein signifikanter Interaktions- und Zwischensubjekteffekt festzustellen ist, zeigt der Innersubjekteffekt, dass sich die Schülerinnen und Schüler der beiden Gruppen *gemeinsam* zwischen der Prä- und Post-Erhebung signifikant entwickeln, was grundsätzlich als positives Ergebnis gewertet werden kann.

Der Übergang in die Sekundarstufe I ist grundsätzlich dadurch gekennzeichnet, dass sich die Schülerinnen und Schüler in ihren neuen Klassenkonstellationen kennenlernen müssen (Griebel & Niesel, 2011) (siehe Kapitel 2.4.1). Nicht alle Schülerinnen und Schüler sind sich aus der Primarstufe bekannt. Die Phasen der Gruppenbildung innerhalb der Lerngruppen sind dabei

sehr intensiv. Gerade die Schülerinnen und Schüler mit einem schwierigen Sozialverhalten haben Probleme damit, sich in neuen Gruppenkonstellationen zu orientieren und ihre Rolle zu finden (Linderkamp & Grünke, 2007), wie auch das Hagener Übergangsprojekt im Kontext der Transition zeigt (Mitzlaff & Wiederhold, 1989). Diese Gruppenprozesse tangieren in der Folge die erfolgreiche Umsetzung von Unterricht und Förderung; mathematisches Lernen steht eventuell erst an nachgeordneter Stelle. Dies kann sich in der nicht vorhandenen kurz- und langfristigen Wirksamkeit in dem vorliegenden Hypothesenkomplex widerspiegeln.

Im Rahmen der Fördermaßnahme sollen gerade die Schülerinnen und Schüler mit einem risikobehafteten Sozialverhalten durch den Einsatz kooperativer und peergestützter Lernformen in ihrer mathematischen Kompetenzentwicklung unterstützt werden. Jedoch dürfen sich die Schülerinnen und Schüler in der vorliegenden Untersuchung ihren Partner bzw. ihre Partnerin und die Mitglieder der Kleingruppe konzeptionell *nicht* selbst aussuchen. Dies kann gerade für die Schülerinnen und Schülern mit einem risikobehafteten Sozialverhalten sehr herausfordernd sein. Gerade in Anbetracht der Schülerinnen und Schüler mit einem schwierigen Sozialverhalten ist in einer Replikation zu erwägen, ob sich die Schülerinnen und Schüler ihren Lernpartner bzw. ihre Lernpartnerin und die Kleingruppen selbst aussuchen können. Auch sollte dann stärker auf die Häufigkeit der Durchführung der peer-mediierten Lernformen geachtet werden, da ein zu häufiger Einsatz die Attraktivität und die Effektivität beeinträchtigen kann (Slavin et al., 2009; Smith et al., 2014). Zudem kann nicht davon ausgegangen werden, dass alle Schülerinnen und Schüler überhaupt mit peergestütztem oder kooperativem Lernen vertraut sind. Zu erwägen ist, die grundsätzlichen Gelingensbedingungen (Haag & Streber, 2011) dieses Förderansatzes zu Beginn vertieft mit den Schülerinnen und Schülern zu thematisieren und zu üben, bevor sie diese eigenständig durchführen. Dies entspricht den Befunden, die zeigen, dass die Qualität peergestützten und kooperativen Lernens entscheidend für den Fördererfolg ist und nicht die Implementation an sich (Kunter & Trautwein, 2013). Auch die Ergebnisse zur Implementation zeigen sowohl auf Seiten der Interventionsleiterinnen und -leiter (siehe Diskussion zur Hypothese 13) als auch auf Seiten der Schülerinnen und Schüler (siehe Diskussion zur Hypothese 9), dass die kooperativen und peergestützten Methoden nicht wie erwartet akzeptiert und angenommen werden.

Einzubeziehen sind weiterhin Befunde, die zeigen, dass die Effektivität peergestützter Förderung im Laufe der Schuljahre sinkt (Kunsch et al., 2007; Rohrbeck et al., 2003; Xin, 1996b), was die eher geringe Attraktivität der Methoden für die Schülerinnen und Schüler in der vorliegenden Untersuchung erklären kann.

Im Post-Follow-Up-I-Follow-Up-II-Vergleich reduziert sich die Stichprobengröße weiter auf 12 (EG) bzw. 11 (KG) Schülerinnen und Schüler. Die Analyse der Stichprobenumfangsplanung mit dem Programm G*Power zeigt auch hier, dass ein mittlerer Effekt bei einer Teststärke von 80 % mit der vorhandenen Stichprobe grundsätzlich herausgefunden werden kann. Auch hier ist auf die großen Standardabweichungen in den Messzeitpunkten zu verweisen, die eine große Heterogenität der kleinen Stichprobe impliziert.

Positiv hervorzuheben ist, dass der signifikante Innersubjekteffekt langfristig auf eine grundsätzlich positive Entwicklung der Schülerinnen und Schüler beider Untersuchungsgruppen hinweist. Zudem zeigen die Schülerinnen und Schüler der Experimentalgruppe durchweg eine po-

sitive Entwicklung ohne einen zwischenzeitigen Leistungsrückgang, wie es bei den Schülerinnen und Schülern mit einem geringen Vorwissen, geringen kognitiven Fähigkeiten oder einer geringen Lesekompetenz der Fall ist.

Ein letzter Aspekt der Diskussion in diesem Hypothesenkomplex betrifft den Einsatz der Lehrereinschätzliste. Die Einschätzung des Sozialverhaltens mittels der Lehrereinschätzliste für das Lern- und Sozialverhalten (2013) erfolgt über die Mathematiklehrkräfte, die auf Grundlage der Beobachtung in den vier Förderwochen bei den Schülerinnen und Schülern die Kriterien Kooperation, Selbstwahrnehmung, Selbstkontrolle, Einfühlungsvermögen, Selbstbehauptung und Sozialkontakt beurteilen sollen. Dabei melden sechs Lehrkräfte einer Schule zurück, dass es ihnen auf Grund der kurzen Beobachtungszeitspanne nicht möglich erscheint, die Einschätzungen vorzunehmen, sodass es zu einem Dropout von 52,13 % in der Kontrollgruppe kommt. Die Bedenken werden von den anderen Lehrkräften der Lerngruppen jedoch nicht geäußert. Berichtet wird aber von einer grundsätzlich geringen Motivation und einer hohen zeitlichen Belastung beim Ausfüllen der Listen. Diese Aspekte sind in Fragen der Validität und Reliabilität der Einschätzungen – und damit auch der hier vorgenommenen Einstufungen der Schülerinnen und Schüler – zu berücksichtigen. Zur Absicherung der Einschätzung des Lern- und Sozialverhaltens wäre künftig der zusätzliche Einsatz einer Schülerselbsteinschätzung, zum Beispiel durch die Schülereinschätzliste (Petermann & Petermann, 2014), zu erwägen.

Hypothese 6a und 6b: Wirksamkeit bei einem risikobehafteten Lernverhalten

Die Hypothese 6a nimmt an, dass die Schülerinnen und Schüler mit einem risikobehafteten Lernverhalten in der Experimentalgruppe im Prä-Post-Vergleich eine bessere mathematische Entwicklung zeigen als diejenigen in der Kontrollgruppe. Der Interaktionseffekt in der zweifaktoriellen Varianzanalyse mit Messwiederholung nimmt einen signifikanten Wert mit einer mittleren Effektstärke von $d = 0.739$ an ($F(1) = 5.567$; $p = 0.023$; $\eta^2 = 0.120$). Die Schülerinnen und Schüler mit einem risikobehafteten Lernverhalten in der Experimentalgruppe entwickeln sich im Prä-Post-Vergleich besser als diejenigen der Kontrollgruppe. *Die Nullhypothese wird zu Gunsten der Alternativhypothese verworfen.*

Die Hypothese 6b beinhaltet die Annahme zur langfristigen Wirksamkeit der Fördermaßnahme. Sie sagt aus, dass die Schülerinnen und Schüler mit einem risikobehafteten Lernverhalten in der Experimentalgruppe im Zeitraum zwischen der Post-, Follow-Up-I- und Follow-Up-II-Erhebung eine bessere Entwicklung in den mathematischen Basiskompetenzen zeigen als die Schülerinnen und Schüler der Kontrollgruppe. Für die Hypothese 6b liegt kein signifikanter Interaktionseffekt vor ($F(2) = 0.546$; $p = 0.582$; $\eta^2 = 0.016$). Auch der Haupteffekt „Gruppenzugehörigkeit“ (Zwischensubjekteffekt) nimmt keinen signifikanten Wert an ($F(1) = 0.223$; $p = 0.640$; $\eta^2 = 0.007$), sodass sich die beiden Gruppen langfristig auf einem vergleichbaren Entwicklungsniveau befinden. Die mathematischen Kompetenzen der Experimentalgruppe entwickeln nicht besser als die der Kontrollgruppe, sodass keine langfristigen Fördererfolge zu erkennen sind. *Die Nullhypothese wird angenommen.*

Die Schülerinnen und Schüler mit einem risikobehafteten Lernverhalten erzielen in der Prä-Testung sehr geringe Werte im HRT 1–4, die nur bei den Schülerinnen und Schülern mit einem sonderpädagogischen Unterstützungsbedarf im Bereich Lernen und bei denjenigen mit einem geringen Vorwissen noch niedriger sind. Die kurzfristige Wirksamkeit ist dabei mit einer Ef-

fektstärke von $d = 0.739$ versehen. Somit handelt es sich um eine Gruppe mit durchaus ungünstigen Voraussetzungen für mathematisches Lernen, die zugleich kurzfristig am stärksten von der Förderung profitiert. Der Wert steht gar an der Schwelle zu einer großen Effektstärke.

Die zugrundeliegende Definition des „Lernverhaltens" durch die Lehrereinschätzliste für Sozial- und Lernverhalten (Petermann & Petermann, 2013) ist dabei eher weit gefächert. Sie umfasst sowohl motivationale als auch metakognitive Aspekte erfolgreichen Lernens. Für die Schülerinnen und Schüler in der Experimentalgruppe, die dahingehend eine risikobehaftete Ausgangssituation aufweisen, scheint die Fördermaßnahme eine sehr gute Unterstützung für mathematisches Lernen darzustellen. Dies mag auf einer Makroebene begründet sein, durch den sehr strukturierten und manualisierten Unterricht, auf der Mesoebene durch die interne Orientierungsmöglichkeit (Advance Organizer), den klaren Strategieinstruktionen (Direkte Instruktion) oder auch den eindeutigen Aufgaben auf den Arbeitsblättern, die durch Beispielaufgaben, Strategie-Checkliste und Möglichkeiten der Selbstkontrolle versehen sind. Motivationsfördernd auf einer Mikroebene könnte zudem das Verhaltensfeedback und die Lernverlaufsdiagnostik gewesen sein. Auch diese Überlegungen können aufgrund fehlender Befunde abschließend nicht bestätigt werden.

Die Einschätzung des Lernverhaltens erfolgt hier ebenso wie die des Sozialverhaltens durch die hospitierenden Lehrkräfte. Sie beurteilen bei den Schülerinnen und Schülern die Bereiche Anstrengungsbereitschaft, Konzentration, Selbstständigkeit beim Lernen und Sorgfalt beim Lernen. Ebenso wie für den Bereich Sozialverhalten ist in der Kontrollgruppe ein Dropout von 52,13 % zu verzeichnen, sodass die Analysen auf einer deutlich reduzierten Stichprobengröße erfolgen. In die Analyse der kurzfristigen Wirksamkeit gehen lediglich 21 (EG) bzw. 22 (KG) Schülerinnen und Schüler ein, wobei die Stichprobengröße in den Untersuchungsgruppen mit der reduzierten Kontrollgruppe nun vergleichbar groß ist. Die Analyse der Stichprobenumfangsplanung mit dem Programm G*Power bei einem erwarteten mittleren Effekt und einer Teststärke von 80 % ergibt, dass die Stichprobe grundsätzlich ausreichend ist, um einen Effekt abbilden zu können. In der Analyse der langfristigen Wirksamkeit befinden sich sowohl in der Experimental- als auch in der Kontrollgruppe 18 Schülerinnen und Schüler. Auch diese Stichprobe ist laut der Umfangsplanung durch das Programm G*Power groß genug, um einen mittleren Effekt bei einer Teststärke von 80 % aufzudecken.

Die Schülerinnen und Schüler der Kontrollgruppe holen den Vorsprung der Experimentalgruppe zum Follow-Up-I-Messzeitpunkt wieder auf, übertreffen diese jedoch nicht. Der nicht signifikante Zwischensubjekteffekt zeigt an, dass sich die beiden Gruppen – bei grundsätzlich positiver Entwicklung (Innersubjekteffekt) – auf einem vergleichbaren Entwicklungsniveau bewegen. Auch hier zeichnet sich die Relevanz von regelmäßigen Auffrischungssitzungen ab, die die Kompetenzentwicklung der Schülerinnen und Schüler der Experimentalgruppe langfristig stärker unterstützen könnten. Am Ende des Schuljahres (Follow-Up-II) befinden sich die Schülerinnen und Schüler beider Untersuchungsgruppen im Vergleich jedoch weiterhin auf einem sehr niedrigen Entwicklungsniveau, unter dem lediglich die Schülerinnen und Schüler mit einem sonderpädagogischen Unterstützungsbedarf oder mit einem geringen Vorwissen liegen. Trotz kurzfristigen Fördererfolgs und einer tendenziell positiven langfristigen Entwicklung darf das weiterhin niedrige Entwicklungsniveau nicht missachtet oder unterschätzt werden.

Überraschend ist auch hier die durchweg positive Entwicklung der Kontrollgruppe über die vier Messzeitpunkte, da gerade für die Gruppe der Schülerinnen und Schüler mit einem risikobehafteten Lernverhalten der Übergang in die Sekundarstufe I eine große Herausforderung darstellen kann (Griebel & Niesel, 2011).

Hypothese 7a und 7b: Wirksamkeit bei einem sonderpädagogischen Unterstützungsbedarf im Bereich Lernen

Die Hypothese 7a besagt, dass die Schülerinnen und Schüler mit einem sonderpädagogischen Unterstützungsbedarf im Bereich Lernen in der Experimentalgruppe im Prä-Post-Vergleich eine bessere mathematische Entwicklung zeigen als diejenigen in der Kontrollgruppe. Der Interaktionseffekt in der zweifaktoriellen Varianzanalyse mit Messwiederholung ist nicht signifikant ($F(1) = 0.608$; $p = 0.444$; $\eta^2 = 0.027$). Die Schülerinnen und Schüler mit einem sonderpädagogischen Unterstützungsbedarf im Bereich Lernen entwickeln sich im Prä-Post-Vergleich nicht besser als diejenigen der Kontrollgruppe. *Die Nullhypothese wird angenommen.*

Die Hypothese 7b beinhaltet die Annahme der langfristigen Wirksamkeit der Fördermaßnahme. Sie besagt, dass die Schülerinnen und Schüler mit einem sonderpädagogischen Unterstützungsbedarf im Bereich Lernen der Experimentalgruppe im Zeitraum zwischen der Post-, Follow-Up-I- und Follow-Up-II-Erhebung eine bessere Entwicklung in den mathematischen Kompetenzen zeigen als die Schülerinnen und Schüler der Kontrollgruppe. Für die Hypothese 7b liegt kein signifikanter Interaktionseffekt vor ($F(2) = 1.023$; $p = 0.370$; $\eta^2 = 0.054$). Die mathematischen Basiskompetenzen der Experimentalgruppe entwickeln sich in wechselseitiger Abhängigkeit von der Gruppenzugehörigkeit und den Messzeitpunkten langfristig *nicht* besser als die der Kontrollgruppe, sodass keine langfristigen Fördererfolge zu erkennen sind. *Es wird die Nullhypothese angenommen.*

In die Analyse der kurzfristigen Wirksamkeit sind lediglich acht (EG) bzw. 16 (KG) Schülerinnen und Schüler einbezogen. Durch diese Stichprobengröße ist die Aufdeckung eines mittleren Effekts bei einer Teststärke von 80 % noch knapp möglich. Die Schülerinnen und Schüler der Experimentalgruppe verzeichnen im Prä-Post-Vergleich zwar einen stärkeren absoluten Zuwachs, doch unterscheidet sich dieser nicht signifikant von dem der Kontrollgruppe. Da weder der Innersubjekt-, Zwischensubjekt- noch Interaktionseffekt signifikante Werte annimmt, muss die gesamte positive Entwicklung der beiden Gruppen als zufällig erachtet werden. Zudem befindet sich auch die Effektstärke in einem kleinen Wertebereich ($d = 0.333$), der keine praktische Relevanz impliziert. Die Fördermaßnahme scheint bei Schülerinnen und Schülern mit einem sonderpädagogischen Unterstützungsbedarf im Bereich Lernen keine lernförderliche Wirkung zu erzielen.

Die Fördermaßnahme basiert auf dem Responsiven Handlungsmodells und verortet sich darin auf der universellen Ebene. Schülerinnen und Schüler, die auf dieser Ebene nicht „responsiv" sind, bei denen der Unterricht folglich keinen Lernzuwachs ermöglicht, bedürfen einer weiteren Förderung auf der zweiten selektiven Ebene. Es zeichnet sich in den Hypothesen 1–6 ab, dass die hier vorgestellte Fördermaßnahme für einen Großteil der Schülerinnen und Schüler unterstützend wirkt, diejenigen mit einem Unterstützungsbedarf im Bereich Lernen jedoch eine intensivere Unterstützung benötigen. Für diese Schülerinnen und Schüler reicht die vorgenommene Differenzierung „nach unten" nicht aus. Hier ist die Entwicklung weiterer Differenzie-

rung innerhalb der Fördermaßnahme auf selektiver und indizierter Ebene angezeigt (für konkrete Vorschläge siehe Kapitel 10.2). Dass es sich bei der Gruppe der Schülerinnen und Schüler mit einem sonderpädagogischen Unterstützungsbedarf im Bereich Lernen um eine sehr leistungsheterogene Gruppe handelt, zeigen die hohen Standardabweichungen, die hier erstmalig in einem Wertebereich von über 40 liegen. Von dieser weiten Streuung berichten auch Gebhardt und Kollegen (2013) in ihrer Querschnittstudie. Wird die Lernausgangslage in der Prä-Erhebung betrachtet, wird zudem deutlich, dass die Schülerinnen und Schüler der Experimentalgruppe – im Vergleich zu den anderen betrachteten Sub-Gruppen – absolut die niedrigsten Werten erreichen. In der Kontrollgruppe erreichen nur diejenigen mit einem geringen Vorwissen noch niedrigere Werte.

In der Analyse der langfristigen Wirksamkeit verringert sich die Experimentalgruppe um eine Person auf sieben Schülerinnen und Schüler. Diese Verringerung hat zugleich auch eine *Verringerung* des Mittelwerts in der Post-Erhebung um 11,63 Rohwertpunkte zur Folge. Die Experimentalgruppe verringert sich in dem Zuge um fünf Personen auf 13 Schülerinnen und Schüler, was hier jedoch eine *Erhöhung* des Mittelwerts in der Post-Erhebung um 6,01 Rohwertepunkte mit sich bringt. Dies zeigt die großen interindividuellen Unterschiede im mathematischen Leistungsspektrum dieser Schülerinnen und Schüler und zugleich die schwierige Messbarkeit dieser Variable. Auch wenn die Analyse des Stichprobenumfangs mit G*Power ergibt, dass ein mittlerer Effekt mit der vorhandenen Stichprobe – gerade noch – aufgedeckt werden kann, fordern die Ergebnisse eine Replikation mit einer größeren Stichprobe, um valide Aussagen über die Wirksamkeit bei den Schülerinnen und Schülern mit einem sonderpädagogischen Unterstützungsbedarf im Bereich Lernen zu ermöglichen. Auch die Erhebung der Kompetenzen in Form einer Gruppentestung ist hier zu überdenken, da oftmals komorbide Schwierigkeiten im emotionalen und sozialen Bereich vorliegen, die im Rahmen einer Gruppentestung zu Tage treten können und auch in der vorliegenden Untersuchung beobachtet werden. In der NELS:88-Studie werden zum Beispiel 5,4 % der Schülerinnen und Schüler aus der Untersuchung ausgeschlossen, da die Erhebungsinstrumente nicht zu ihrem Leistungsniveau passen (Kaufman & Bradbury, 1992). Es sind bei diesen Schülerinnen und Schülern folglich sehr sensible Messinstrumente einzusetzen. Somit ist in der vorliegenden Untersuchung die Trennschärfe des HRT 1–4 im untersten Leistungsbereich zu hinterfragen.

Effektstärken der Hypothesen 1–7 im Vergleich

Folgende Abbildung 65 gibt einen Überblick über die Effektstärken der Interaktionseffekte in der Analyse der kurz- und langfristigen Wirksamkeit in den Hypothesen 1–7. Die Effektstärken werden hier unabhängig der Signifikanzen dargestellt. Neben der Klassifikation nach Cohen (1988) ist zudem die 0.4-Grenze angegeben, ab der Hattie und Kollegen (2013) einen praktisch relevanten Fördereffekt verorten.

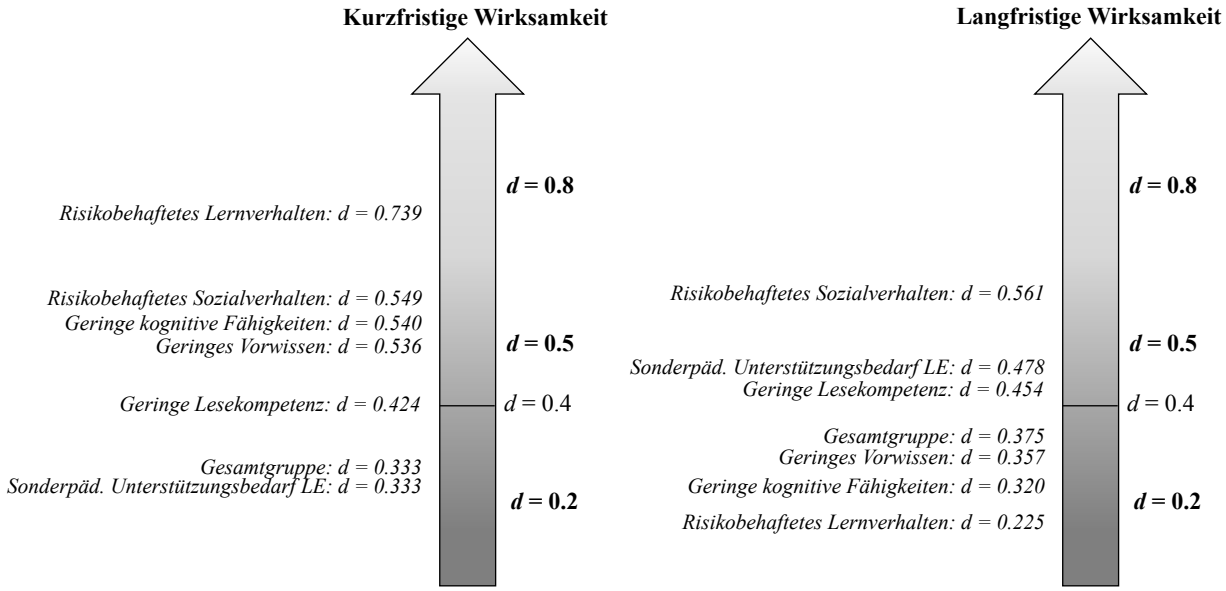


ABBILDUNG 65 VERGLEICH DER EFFEKTSTÄRKEN DES INTERAKTIONSEFFEKTS IN DEN HYPOTHESENKOMPLEXEN 1–7

Im Rahmen der kurzfristigen Wirksamkeit liegen die Schülerinnen und Schüler, bis auf diejenigen mit einem Unterstützungsbedarf im Förderschwerpunkt Lernen, im Effektstärkenbereich > 0.4. Bezogen auf die langfristige Wirksamkeit befinden sich über dieser Schwelle nur diejenigen mit einem risikobehafteten Sozialverhalten, einer geringen Lesekompetenz und einem Unterstützungsbedarf im Bereich Lernen.

Auffällig ist der Rückgang der Effektstärke in der langfristigen Wirksamkeit bei den Schülerinnen und Schülern mit einem risikobehafteten Lernverhalten, die kurzfristig die höchste Effektstärke verzeichnen. Auch diejenigen mit geringen kognitiven Fähigkeiten und einem geringen Vorwissen fallen in der langfristigen Wirksamkeit um einige Plätze zurück. Die Effektstärken bei den Schülerinnen und Schülern mit einem risikobehafteten Sozialverhalten und einer geringen Lesekompetenz verändern sich kaum und bleiben auf einem ähnlichen Niveau. Die Schülerinnen und Schüler mit einem Unterstützungsbedarf im Bereich Lernen zeigen in der langfristige Perspektive größere Werte.

10.1.2 Hypothesenblock III

Die multiple lineare Regression zeigt, dass der Leistungszuwachs innerhalb der Experimentalgruppe in der HRT-Skala Rechenoperationen zwischen der Prä- und Posterhebung zu einem kleinen Anteil durch den Prädiktor Lernfreude aufgeklärt werden kann. Der Prädiktor nimmt dabei einen signifikanten Wert an, verzeichnet jedoch lediglich ein R^2 von 3,5 %, sodass nur dieser geringe Varianzanteil durch die abhängige Variable erklärt werden kann. Die Lernfreude hat also grundsätzlich einen Einfluss auf den Leistungszuwachs im HRT von der Prä- zur Posterhebung, doch sollte dieser in seiner Bedeutsamkeit in der vorliegenden Untersuchung nicht zu viel Gewicht beigemessen werden. *Für die Hypothese H9a wird die Nullhypothese zu Gunsten der Alternativhypothese verworfen.*

Der Prädiktor der Leistungsheterogenität der Klasse nimmt keinen signifikanten Wert an. *Für die Hypothese H9b wird die Nullhypothese angenommen.*

Das Ergebnis bestätigt, dass die Lernfreude bedeutsam ist für schulisches Lernen (Klauer, 2011b, S. 206). Dies führt auch Matthes (2006) in seinem Handlungsmodell unter dem zentralen Aspekt der „Motivation“ auf (siehe Kapitel 3.1.1). Auch wenn der Fokus in der Sekundarstufe I stärker auf der Vermittlung von curricularen Inhalten liegt (Racherbäumer & Kohnen, 2014), zeigen die Ergebnisse eindeutig, dass auch die Lernfreude der Schülerinnen und Schüler gefördert werden sollte, da sie einen Einfluss auf den Lernerfolg jedes Einzelnen hat.

Zu bemängeln ist hier, dass die Lernfreude in der vorliegenden Untersuchung nicht in der Kontrollgruppe erhoben wird. Dies würde differenzierte Erkenntnisse bzgl. der grundsätzlichen Entwicklung der mathematischen Kompetenzen zu Beginn der Sekundarstufe I in Abhängigkeit der Lernfreude ermöglichen. Positiv hervorgehoben wird, dass die Lernfreude jedoch überhaupt Berücksichtigung im Rahmen der vorliegenden Untersuchung gefunden hat, da dieser Aspekt oftmals vernachlässigt wird (Fritz & Ricken, 2008, S. 14).

Die Annahme, dass jeder einzelne Schüler bzw. jede Schülerin von einer leistungsstarken Klasse profitiert (Kompositionseffekt, siehe Opdenakker & Van Damme, 2007), kann hier nicht bestätigt werden. Die Ergebnisse in der vorliegenden Untersuchung zeigen, dass die Wirksamkeit der Maßnahme nicht von der Leistungsheterogenität der Klasse abhängt. Bezüglich des Faktors der Leistungsheterogenität der Klasse ist zu diskutieren, inwiefern die durchschnittliche Standardabweichung der gesamten Klasse in der Prätestung ein zuverlässiger Indikator für die Leistungsheterogenität ist. Alternativ könnte in Folgeuntersuchungen das „korrigierte“ Leistungsniveau der *restlichen* Klasse (die Versuchsperson selbst ausgeschlossen) genutzt werden.

10.1.3 Hypothesenblock IV

IMPLEMENTATIONSQUALITÄT – BEURTEILUNG DER SCHÜLERINNEN UND SCHÜLER

HYPOTHESE 9

Die Hypothese 9 nimmt an, dass die Fördermaßnahme für die Schülerinnen und Schüler der Experimentalgruppe zufriedenstellend ist und sie diese akzeptieren (Implementationskriterium „acceptability“). Sie stimmen den Items zur *Akzeptanz* (AK) der Maßnahme durchschnittlich „ziemlich“ zu (entspricht dem Skalenwert 2 von 4). Die Analyse ergibt, dass die Schülerinnen und Schüler den positiv gerichteten Aussagen zur Akzeptanz der Maßnahme durchschnittlich mit einem Wert von 2,11 zustimmen, der sich nicht signifikant vom Erwartungswert 2 unterscheidet ($p = 0.106$). *Die Nullhypothese wird zu Gunsten der Alternativhypothese verworfen.*

Die ergänzende deskriptive Analyse zu den drei Subskalen (Eigener Lernerfolg, Zufriedenheit mit den Methoden, Zufriedenheit insgesamt) zeigt, dass die durchschnittlichen Zustimmungen nicht massiv voneinander abweichen. Die Schülerinnen und Schüler sehen einen eigenen Lernerfolg und empfinden die eingesetzten Methoden als lernförderlich. Die Zufriedenheit und die Akzeptanz der Fördermaßnahme werden insgesamt durchschnittlich mit 2,13 bewertet, was als zufriedenstellendes Ergebnis angenommen wird.

Die durchschnittlich gegebene Schulnote liegt bei 2,72 mit einer Standardabweichung von 1,57. Hier zeigt sich eine breite Streuung, sodass einige Schülerinnen und Schüler die Fördermaßnahme sehr gut, andere hingegen eher unzureichend benoten. Hier wäre in Folgearbeiten zu

prüfen, welche weiteren Faktoren die Zufriedenheit beeinflussen. Denkbär wäre, die Unterrichtsqualität der Interventionsleiterin und -leiter oder das Lernklima innerhalb der Klasse zu berücksichtigen (Hattie et al., 2013).

Aufschlussreich sind die Ergebnisse auf Itemebene: Erwartungsgemäß fällt hier die Einschätzung des Lernerfolgs im Bereich der Addition höher aus als im Bereich des Geteiltrechnens. Bzgl. der Freude an den Methoden erhalten die mathematisch ausgerichteten Spiele zum Stundenabschluss die höchste Zustimmung (1,76). Die niedrigste Zustimmung (entspricht einem hohen Wert) hingegen das Üben im Trainingslager (2,44). Dies widerspricht den Befunden von Calhoon und Fuchs (2003), in denen Schülerinnen und Schülern aus den Klassen 9–12 peer-mediierten Lernformen grundsätzlich positiv gegenüberstehen. Jedoch zeigen Smith und Kollegen (Smith et al., 2014), dass die Freude an der Methode bei einem zu frequentierten Einsatz zurückgehen kann. In der vorliegenden Studie fehlen qualitative Aussagen der Schülerinnen und Schüler, um die eher geringe Zustimmung näher interpretieren zu können.

Hypothese 10

Die Hypothese 10 nimmt an, dass die Schülerinnen und Schüler der Experimentalgruppe die Maßnahme als für sie passend empfinden (Implementationskriterium „appropriateness“) und den Items zur *Angemessenheit* (AN) der Maßnahme somit durchschnittlich „ziemlich“ zustimmen (entspricht dem Skalenwert 2 von 4). Die Ergebnisse zeigen, dass die Schülerinnen und Schüler die positiv gerichteten Aussagen durchschnittlich mit einem Wert von 1,93 beurteilen, der sich in eine positive Richtung signifikant vom Erwartungswert 2 unterscheidet ($p = 0.004$). Die Schülerinnen und Schüler beurteilen die Angemessenheit der Maßnahme gar besser als angenommen. *Die Nullhypothese wird zu Gunsten der Alternativhypothese verworfen.*

Im Vergleich zur Akzeptanz der Fördermaßnahme (siehe Hypothese 8) divergieren die Werte in den sechs Subskalen der Angemessenheit stärker untereinander. Es wird deutlich, dass die Schülerinnen und Schüler den positiv gerichteten Aussagen zur lernförderlichen Unterstützung der Arbeitsblätter (1,94), des Advance Organizers (1,83), der Lernverlaufsdiagnostik (1,76) und des Verhaltensfeedbacks (1,68) mit durchschnittlichen Bewertungen kleiner als zwei eher uneingeschränkt zustimmen. Etwas weniger, jedoch immer noch „ziemliche“ Zustimmung mit Werten kleiner als 2,5 finden die Aussagen zur lernförderlichen Unterstützung durch die Direkte Instruktion (2,15) und die peer-mediierten Lernformen (2,37), die sowohl das kooperative als auch das peergestützte Lernen umfassen.

Die Beurteilung der Fördermaßnahme durch die Schülerinnen und Schüler ergibt insgesamt zufriedenstellende Ergebnisse. Die Methoden des kooperativen und peergestützten Lernens (peer-mediierte Lernformen) bilden das Schlusslicht und werden insgesamt als am wenigsten lernförderlich beurteilt und bereiten auch weniger Freude (siehe Hypothese 9) als zum Beispiel die Strategieinstruktion oder das Lern- und Verhaltensfeedback. Gerade schwache Schülerinnen und Schüler profitieren jedoch von peer-mediierten Methoden (Spörer, 2009). Hier schließen sich die bereits getätigten Überlegungen an (siehe Hypothese 5), den Schülerinnen und Schülern das Ziel und die explizite Durchführung dieser Fördermethoden stärker nahezubringen, um die – für einen sichtbaren Fördererfolg notwendige – Qualität der Umsetzung der Methode zu steigern (Haag & Streber, 2011). Denn letztlich ist die Qualität der Implementation der Methoden entscheidend und nicht die Quantität (Engelmann, 2007).

Die Ergebnisse des Implementationsfragebogens ermöglichen insgesamt keine *qualitativen* Hinweise auf die Akzeptanz und die Zufriedenheit mit spezifischen Aspekten der Fördermaßnahe. Zudem muss bei der Interpretation der Ergebnisse beachtet werden, dass es sich nicht um einen standardisierten Fragebogen handelt, sodass die Ergebnisse eher als tendenziell angesehen werden müssen.

Implementationsqualität – Beurteilung der Interventionsleiterinnen und -leiter

Bei der Beurteilung durch die Interventionsleiterinnen und -leiter muss grundsätzlich beachtet werden, dass die Werte auf Angaben von lediglich 13 Personen beruhen und deswegen ebenso eher als tendenziell anzusehen sind.

Hypothese 11

Die Hypothese 11 nimmt an, dass die Interventionsleiterinnen und -leiter mit der Maßnahme zufrieden sind und sie diese akzeptieren (Implementationskriterium „acceptability"). Sie stimmen den Items zur Akzeptanz (AK) der Maßnahme durchschnittlich „ziemlich" zu (entspricht dem Skalenwert 2 von 4). Die Ergebnisse zeigen, dass die Interventionsleiterinnen und -leiter die positiv gerichteten Aussagen mit einem durchschnittlichen Wert von 2,06 beurteilen, der sich nicht signifikant vom angenommenen Erwartungswert 2 unterscheidet ($p = 0.576$). *Die Nullhypothese wird zu Gunsten der Alternativhypothese verworfen.*

Die Interventionsleiterinnen und -leiter geben der Fördermaßnahme – ähnlich wie die Schülerinnen und Schüler – insgesamt die Schulnote 2,69 mit einer Standardabweichung von 0,947, die eine hohe Variabilität der Beurteilung widerspiegelt. Ein Blick in die Items zeigt, dass die Interventionsleiterinnen und -leiter die Fördermaßnahme zu Beginn sehr gerne durchgeführt haben (1,85), die Zufriedenheit nach der Hälfte der Fördermaßnahme sinkt (2,23) und zum Ende wieder steigt (1,92). Die Aussage, dass die Fördermaßnahme die Erwartungen insgesamt erfüllt habe, verortet sich mit einer Beurteilung von 2,5 exakt zwischen „stimmt ziemlich" und „stimmt kaum". Die Ergebnisse zur Akzeptanz der Maßnahme aus Sicht der Interventionsleiterinnen und -leiter werden insgesamt als zufriedenstellend angesehen.

Hypothese 12

Die Hypothese 12 nimmt an, dass die Interventionsleiterinnen und -leiter einer Übernahme der Maßnahme in den schulischen Alltag positiv gegenüberstehen. Der Mittelwert der Skala Übernahme liegt bei 2,15 und weicht mit einer Signifikanz von $p = 0.518$ nicht vom erwarteten Mittelwert 2 ab. *Die Nullhypothese kann verworfen werden, sodass die Alternativhypothese angenommen wird.*

Da die Skala lediglich aus zwei Items besteht, sind die Gütekriterien der Zuverlässigkeit und der inhaltlichen Passgenauigkeit nicht im vollen Umfang gegeben. Die Bereitschaft, die Fördermaßnahme ein weiteres Mal durchzuführen (2,08) ist etwas höher, als diese weiterzuempfehlen (2,23). Hier zeigt sich, dass – obwohl die Interventionsleiterinnen und -leiter die Fördermaßnahme nach der Hälfte der Zeit weniger gerne durchgeführt haben (siehe Hypothese 11) – sie diese mit einer recht hohen Bereitschaft noch einmal durchführen würden. Die positive Akzeptanz und Übernahmebereitschaft kann als Zeichen einer gelungenen Implementation im Rahmen der vorliegenden Untersuchung verstanden werden. Auch hier würden sich qualitative

Interviews anbieten, um herauszufinden, weshalb die Bereitschaft, die Fördermaßnahme weiterzuempfehlen, entsprechend geringer ausfällt.

HYPOTHESE 13

Die Hypothese 13 nimmt an, dass die Interventionsleiterinnen und -leiter die Maßnahme insgesamt als angemessen empfinden (Implementationskriterium „appropriateness"). Sie stimmen den Items zur *Angemessenheit* (AN) der Maßnahme durchschnittlich „ziemlich" zu. Die inferenzstatistische Analyse ergibt einen Mittelwert von 2,25, der signifikant vom erwarteten Wert 2 abweicht. *Die Nullhypothese wird angenommen.*

Die Skala Angemessenheit umfasst im Vergleich zu den anderen Skalen die größte Itemanzahl (32). Die Werte der elf Subskalen zeigen, dass einer Passung des Advance Organizers (1,46), der Inhalte (1,69), der Spiele (1,69), der Lernverlaufsdiagnostik (1,69) und des Verhaltensfeedbacks (1,98) mit einer durchschnittlichen Beurteilung von kleiner als 2 sogar besser als angenommen bewertet werden. In einem Wertebereich zwischen 2 und 2,5 befinden sich der lernförderliche Einsatz Arbeitsblätter (2,08), die grundsätzliche Passung der Methoden (2,10) und der Direkten Instruktion (2,27). Auch die Passung für die Zielgruppe verortet sich in diesem Bereich (2,47), wobei dabei der Aussage der lernförderlichen Unterstützung der Schülerinnen und Schüler mit einem sonderpädagogischen Unterstützungsbedarf kaum zugestimmt wird (3,58). Diese Auffassung bestätigt sich in der bereits diskutierten nicht vorhandenen kurz- und langfristigen Wirksamkeit bei gerade dieser Zielgruppe (siehe Diskussion der Hypothese 7). Hier ist bereits erörtert, dass eine weitere Differenzierung notwendig ist, die über eine Förderung auf der universellen Ebene hinausgeht. Niedrige Zustimmung erhält die Subskala zu der Angemessenheit der peer-mediierten Methoden, worunter sowohl die Tandem- als auch die Kleingruppenarbeit verortet ist. Gerade das Trainingslager wird darin als wenig lernförderlich angesehen (3,25). Die Passung der Zeitplanung erhält mit einem Wert von 2,84 die am wenigsten positive Beurteilung. Neun der Interventionsleiterinnen und -leiter stimmen der Aussage gänzlich zu, dass zu wenig Zeit für die Umsetzung der Inhalte der Fördermaßnahme zur Verfügung steht. Hier lässt sich eindeutig erkennen, dass die Inhalte und die Zeitplanung nicht ausreichend abgestimmt sind und maßnahmenbezogene strukturelle Modifikationen erfolgen sollten.

HYPOTHESE 14

Die Hypothese 14 nimmt an, dass die Interventionsleiterinnen und -leiter die Maßnahme als im schulischen Setting machbar und umsetzbar empfinden (feasibility). Sie stimmen den Aussagen zur *Machbarkeit* (M) der Maßnahme durchschnittlich „ziemlich" zu (entspricht dem Skalenwert 2 von 4). Der Mittelwert der Skala liegt bei 1,96 und weicht nicht signifikant vom erwarteten Wert 2 ab ($p = 0.813$). Die Subskalen zur Eignung des Manuals (1,62) und der Umsetzbarkeit im Kontext der persönlichen Ressourcen (1,85) werden mit einem durchschnittlichen Wert von kleiner als 2 beurteilt. In der Frage nach der Umsetzbarkeit mit den Ressourcen im Feld Schule ergibt die Zustimmung einen etwas geringeren Wert von 2,27. Hier fallen auf Itemebene vor allem der Materialaufwand (2,46) und Zweifel an der Durchführung durch eine einzige Lehrkraft (2,23) ins Gewicht. Am wenigsten Zustimmung erhält die Aussage, dass in der vorliegenden Untersuchung die betreuende Lehrkraft der Fördermaßnahme positiv gegenübersteht (2,77). Für eine langfristige Implementation der Fördermaßnahme in den schulischen

Alltag ist eine überzeugte Lehrkraft, die in ein unterstützendes Kollegium eingebunden ist, eine der drei zentralen Gelingenskriterien (Oswald et al., 2013). Bei einer Replikation der vorliegenden Studie, in der die Lehrkräfte ggf. selbst die Fördermaßnahme durchführen, sollte ihre Beurteilung mit eingeholt werden. Dies könnte Aufschlüsse darüber ermöglichen, weshalb die Unterstützung im Rahmen der vorliegenden Studie von den Interventionsleiterinnen und -leiter als gering eingeschätzt wird.

HYPOTHESE 15

Die Hypothese 15 nimmt an, dass die Interventionsleiterinnen und -leiter die Maßnahme entsprechend des Manuals umsetzen (fidelity). Sie stimmen den Aussagen zur *Wiedergabetreue* (W) der Maßnahme durchschnittlich „ziemlich" zu (entspricht dem Skalenwert 2 von 4). Die Skala nimmt einen Wert von 1,94 an, der sich nicht signifikant vom erwarteten Mittelwert 2 unterscheidet ($p = 0.619$). *Die Nullhypothese wird zu Gunsten der Alternativhypothese verworfen.*

Hier muss berücksichtigt werden, dass die Skala insgesamt nur aus fünf Items besteht und die Ergebnisse deswegen als tendenziell angesehen werden müssen. Die Beurteilungen auf Itemebene zeigen, dass die Materialien grundsätzlich verwendet und nach Vorgabe umgesetzt werden. Dies fügt sich ein in die gute Beurteilung der Handhabung und Praktikabilität des Manuals (siehe Hypothese 14). Entsprechend der wenig zufriedenstellenden Zeitplanung im Kontext der Stundeninhalte (siehe Hypothese 14) zeigt sich hier entsprechend, dass die Zeitstruktur eher weniger nach Vorgaben eingehalten wird (2,62). Die Zeitstruktur der Unterrichtseinheit (eine bzw. drei Stunde/n) wird hingegen eng nach Vorgabe umgesetzt (2,15).

Insgesamt zeigen sich die Interventionsleiterinnen und -leiter in einem Großteil der erfragten Skalen zufrieden mit der Fördermaßnahme. Sie stimmen den positiv gerichteten Aussagen im Regelfall zu. Die Evaluation der Implementation gibt jedoch auch Hinweise auf erforderliche Modifikationen. Dies wird vor allem in der Skala der Angemessenheit deutlich.

10.2 Methodenkritische Reflexion

Im folgenden Kapitel erfolgt zu Beginn die Reflexion der Entwicklung und Durchführung der Fördermaßnahme (Kapitel 10.2.1) und im Anschluss die Reflexion des Forschungsdesigns, der Erhebungsinstrumente, der Stichprobenrekrutierung und der statistischen Analysen (Kapitel 10.2.2).

10.2.1 Entwicklung und Durchführung der Fördermaßnahme

„Es ist ein tolles Gefühl, wenn das klappt,
wenn aufwendige Arbeit zum Erfolg führt."
(Feedback einer Interventionsleiterin)

Der zeitliche Ablauf der Entwicklung der Fördermaßnahme ist in Kapitel 6.2 bereits dargestellt und erläutert (siehe Tabelle 28).

Im Juni 2014 sind die Förderstunden, Instruktionen und Materialen der Maßnahme abschließend ausgearbeitet, sodass ein Manual erstellt wird, mit dem die Umsetzung der Hauptstudie

im September und Oktober 2014 erfolgen kann. Das Manual gibt den Interventionsleiterinnen und -leitern konkrete Anweisungen und Hinweise zur Durchführung der Förderung und erhöht damit die Vergleichbarkeit der Umsetzung in den einzelnen Lerngruppen (Petermann, 2015a).

Um auch von der fachwissenschaftlichen Gemeinschaft ein Feedback zu erhalten, wird das Projekt im Juni 2014 nach Abschluss der Entwicklungsphase im Rahmen des Fachgesprächs Lernen an der Universität Oldenburg vorgestellt. Zusätzlich unterstützend wäre eine Präsentation des Forschungsprojekts vor einem Fachpublikum schon *während* der Entwicklungsphase gewesen.

Die Durchführung der Pilotstudie im Juni 2014 erweist sich insgesamt als sinnvoll und zielführend. Zu bemängeln ist dabei, dass sich die 30 Schülerinnen und Schüler zu dem Zeitpunkt der Pilotstudie bereits am Ende der fünften Klasse befinden. Dieser Umstand ist aufgrund der Zeitplanung innerhalb des Projekts jedoch nicht zu vermeiden. Dennoch sammeln die beiden Interventionsleiterinnen wertvolle Erfahrungen und erhalten konstruktives Feedback von den Schülerinnen und Schülern und von der betreuenden Mathematiklehrkraft. Das Feedback und die formative Evaluation durch die Interventionsleiterinnen münden im Anschluss in einer Überarbeitung der Fördermaßnahme.

Damit die Erwartungen auf Seiten der Interventionsleiterinnen und -leiter und der Projektleitung definiert und abgesprochen sind, wird vor der Umsetzung der Hauptstudie ein Kontrakt unterschrieben, in dem beide Seiten eine zuverlässige und unterstützende Mitarbeit versprechen. Dieser Kontrakt führt während der Projektphase zu einer zuverlässigen Arbeitskultur und wird als positiv angesehen.

Während der Umsetzung der Hauptstudie im September und Oktober 2014 wird eine summative Evaluation der Fördermaßnahme durch die Implementationsfragebögen verfolgt. In Fragen der Implementations*qualität* zeigt sich im Nachhinein, dass eine *formative* Evaluation differenziertere Aussagen zu spezifischen Arbeitsphasen oder -methoden in den einzelnen Förderstunden ermöglichen kann. Im Rahmen der vorliegenden Arbeit können zwar Tendenzen bzgl. der Zustimmung zu einzelnen Implementationsvariablen herausgefunden werden, doch bleiben qualitative, differenzierte Antworten sowohl seitens der Interventionsleiterinnen und -leiter als auch seitens der Schülerinnen und Schüler aus. Hilfreich wäre hier ein kurzer Fragebogen, den die Interventionsleiterinnen und -leiter auch in der Hauptstudie am Ende jeder Förderstunde zu positiven und negativen Aspekten der durchgeführten Stunde ausfüllen und der Hinweise und Fallstricke bzgl. der Umsetzung der Inhalte und Arbeitsphasen gibt. Ebenso wären ergänzende Interviews mit den Interventionsleiterinnen und -leitern und Schülerinnen und Schülern hilfreich, um detailliertes Feedback zu erhalten.

In der Evaluation mit der Projektgruppe im Juli 2015 wird die Arbeitsatmosphäre in der gesamten Gruppe von den Interventionsleiterinnen und -leitern als positiv herausgestellt. Dies ist keineswegs ein triviales Feedback, da eine gute Arbeitsatmosphäre Voraussetzung für ein erfolgreiches Projekt ist. Befunde zeigen, dass die gegenseitige kollegiale Unterstützung maßgeblich für eine gelungene Implementation von Förderkonzepten ist (Oswald et al., 2013). Somit kommt der Zufriedenheit und dem Gefühl des Unterstützt-Seins der Interventionsleiterinnen und -leiter eine hohe Bedeutsamkeit zu. Die Interventionsleiterinnen und -leiter empfinden zudem das Material- und Diagnostikmanual als zugänglich und handlich. Dies passt zu den posi-

tiven Ergebnissen in der Implementationsskala Wiedergabetreue (siehe Kapitel 9.6.1). Sie berichten später, dass das Plakat des Advance-Organizers auch noch in den Follow-Up-Erhebungen in den Klassen hängt und zum Teil im regulären Mathematikunterricht Anwendung findet, was als positiv erachtet wird. Die Interventionsleiterinnen und -leiter regen darüber hinaus an, zukünftig weniger Arbeitsblätter pro Stunde einzuplanen, da die Bearbeitung und Besprechung mehrerer Arbeitsblätter innerhalb des vorgegebenen Rahmens unter Berücksichtigung der hohen Methodenvielfalt zeitlich nicht umsetzbar ist. Als weniger positiv werden in der Evaluationssitzung u. a. der Zeitdruck am Ende der Förderstunden oder der z. T. schwierige Kontakt bzgl. wichtiger Absprachen mit den Lehrkräften genannt. Auch die „Merksätze" in der Direkten Intstruktion sind zum Teil zu kompliziert und sollen einer Komplexitätsreduktion unterzogen werden. Zudem wird ebenso weniger positiv gesehen, dass das geplante Tandemkonzept aufgrund der begrenzten Anzahl an Interventionsleiterinnen nicht in jeder Klasse umgesetzt werden kann. Die Interventionsleiterinnen und -leiter äußern zudem, dass es bei einer weiteren Programmevaluation sinnvoll wäre, die Diagnostik von Studierenden durchführen zu lassen, die *nicht* zugleich die Fördermaßnahme in der Klasse umsetzen. Es besteht die Vermutung, dass die Schülerinnen und Schüler, die weniger Freude an der Diagnostik haben, dies mit den Interventionsleiterinnen und -leitern assoziieren. Dies kann letztlich eine negative Folge für die Mitarbeit und das Engagement im Rahmen der Fördermaßnahme haben.

Weiterhin ist das Engagement und die Bereitschaft der Lehrkräfte zur Mitarbeit am Projekt zu diskutieren. Die Teilnahme am Projekt geschieht auf Seiten der Lehrkräfte auf freiwilliger Basis. Sie werden während des Projekts in regelmäßigen Abständen über die Projektschritte informiert, sodass transparent ist, wie lange die Durchführung (noch) andauert und zu welchem Zeitpunkt die diagnostischen Erhebungen der Kompetenzen geplant sind. Obwohl die Lehrkräfte auch über den Sinn und Zweck der Lehrereinschätzliste informiert werden, gibt es zum Teil massive Widerstände, diese auszufüllen, da dies eine zu hohe zeitliche Belastung sei und das Verhalten der Schüler nach sechs Wochen noch nicht valide eingeschätzt werden könne. Die sechs Lehrkräfte, die die Bearbeitung verweigern, stammen von derselben Schule, sodass hier vermutlich zuvor eine interne Kommunikation stattgefunden hat. Dies führt letztlich – neben einer weiteren Lehrkraft, die während der Projektphase die Klasse wechseln muss – zu dem hohen Dropout zum Follow-Up-I-Messzeitpunkt in der Kontrollgruppe. Positiv ist jedoch insgesamt festzuhalten, dass die Drop-Out-Quoten in den Schülererhebungenüber das ganze Schuljahr hinweg recht gering sind (siehe Tabelle 50). Dies kann als Zeichen einer guten Projektorganisation gewertet werden.

Bei Fragen oder Unstimmigkeiten bzgl. der Förderung zeigen sich in der vorliegenden Untersuchung Telefongespräche oder auch persönliche Treffen zwischen den Lehrkräften und der Projektleitung als problemlösend. Um Schwierigkeiten mit Lehrkräften künftig frühzeitig entgegenzuwirken, ist ein *früher persönlicher* Kontakt zu erwägen, zum Beispiel in Form eines Projekttreffens zu Beginn der Studie in den Räumlichkeiten der Schulen oder der Universität.

Die teilnehmenden Lehrkräfte werden am Ende der Studie zu einem Projektabschluss in die Räumlichkeiten der Universität eingeladen, bei dem zentrale Ergebnisse präsentiert werden und Feedback eingeholt wird. Leider ist dieser früh angekündigte Termin sehr schlecht besucht, was die Herausforderung der Verzahnung von Wissenschaft und schulischer Praxis verdeutlicht (Križan & Vossen, 2016). Hinweise auf die Zufriedenheit der Lehrkräfte aus der Experimentalgruppe geben vier kurze strukturierte Interviews zur Implementation der Fördermaßnahme,

die einzeln mit der Projektleitung und den Lehrkräften geführt werden. Darin äußern sich die Lehrkräfte grundsätzlich positiv zu der Fördermaßnahme, bemängeln im Gros jedoch auch eine unzureichende Differenzierung und einen Zeitdruck innerhalb der Förderstunden. Dies spiegelt das Feedback der Interventionsleiterinnen und -leiter wider (siehe Hypothesen 11–14).

Letztlich stellt sich die Frage, inwiefern die hier evaluierte Fördermaßnahme in ihrer jetzigen Form implementiert werden kann bzw. an welchen Stellen es Änderungsbedarfe gibt. Die Evaluation der Implementationsqualität zeigt, dass der zeitliche Rahmen zu knapp bemessen ist. Da die Lehrkräfte den Umfang der Förderung (vier Wochen) grundsätzlich als passend beschreiben, wird eine *Ausweitung* weniger erwogen. Alternativ ist eine inhaltliche Flexibilisierung möglich, sodass zum Beispiel die mehr herausfordernden Themen Subtraktion und Multiplikation einen größeren Förderumfang erhalten. Hier wird zum Beispiel auch der Lernerfolg durch die Schülerinnen und Schüler als geringer eingeschätzt (siehe Anhang 4), was sich auch in den ergänzenden Ergebnissen auf Subtestebene des HRT 1–4 zeigt (siehe Tabelle 68). Die Entscheidung der Schwerpunktsetzung ist mit einer gezielten Eingangsdiagnostik steuerbar. Für die sehr schwachen Schülerinnen und Schüler, zum Beispiel mit einem Unterstützungsbedarf im Bereich Lernen, bedarf es einer stärkeren Differenzierung. Hier sollte – nach dem Konzept des Responsiven Handlungsmodells – auf der selektiven Ebene (Stufe 2) das Arbeiten in binnendifferenzierten Sozialformen wie Kleingruppen oder auch Einzelarbeit erwogen werden. Diese Phase sollte dann nach den überaus individuellen Entwicklungsständen der Schülerinnen und Schüler gestaltet werden, wo in Teilen evtl. gar noch die Förderung der Vorläuferkompetenzen angezeigt ist. Grundsätzlich sollte jedoch die Partizipation an der Fördermaßnahme ermöglicht werden, zum Beispiel als Rahmung in der Hinführungs- und Abschlussphase. Denkbar wäre ebenso die Gestaltung von Zusatzstunden für die schwachen Schülerinnen und Schüler, wie es zum Beispiel das Lubo-Programm – für den Bereich emotional-soziale Entwicklung – in seiner dritten Auflage vorsieht (Hillenbrand, Hennemann, Hens & Hövel, 2015). Die Identifikation der Non-Responder kann neben der Eingangsdiagnostik auch über die vorhandene Lernverlaufsdiagnostik erfolgen.

Da in Deutschland bisher kein Förderprogramm auf universeller Ebene für die Sekundarstufe I vorliegt, kann kein Vergleich bzgl. der Differenzierung mit einem mathematischen Förderkonzept getätigt werden. Die Evaluation der Implementation der vorliegenden Untersuchung zeigt, dass die Partizipation der schwachen Schülerinnen und Schülern an den Förderstunden aus Sicht der Interventionsleiterinnen und -leiter *an sich* als schwierig erachtet wird (siehe Hypothese 12), sodass ausschließliche Zusatzstunden dieser Kritik nicht begegnen würden. *Primär* sollte die innere Differenzierung weiter ausgebaut werden, was jedoch langfristig die zusätzliche Entwicklung von Vertiefungsstunden nicht ausschließt. Hier würde es sich anbieten, auf selektiver oder indizierter Ebene bereits evaluierte Förderkonzepte hinzuzuziehen und zu implementieren (zum Beispiel „Kalkulie“) (siehe Tabelle 20). Dies wäre eine optimale Nutzung verschiedener evidenzbasierter Programme auf den drei Ebenen des RTI-Modells. Ergänzend ist bereits die Implementation von Auffrischungssitzungen diskutiert, die an die unmittelbaren Unterrichtsinhalte anknüpfen, um die Wirksamkeit der Förderung langfristig stabilisieren zu können (siehe Diskussion zur Hypothese 1).

Zudem weisen die Ergebnisse der Implementationsfragebögen darauf hin, dass die peer-mediierten Methoden sorgfältiger eingeführt und mit den Schülerinnen und Schülern erarbeitet werden sollten, um die Freude daran und das Potenzial der Lernförderlichkeit zu steigern. Auch

sollte die Frequenz des Einsatzes kritisch überdacht werden. Ob die Schülerinnen und Schüler ihre Lernpartner selbst aussuchen dürfen, sollte je nach sozialem Klassenklima beurteilt werden.

10.2.2 Forschungsmethodik

Abschließend erfolgt eine methodenkritische Diskussion des Forschungsdesigns, der Erhebungsinstrumente, der Stichprobenrekrutierung und der statistischen Analysen.

FORSCHUNGSDESIGN

Bei dem vorliegenden Forschungsdesign handelt es sich um ein starkes Design in Form eines randomisierten Zwei-Gruppen-Plans mit Vortest, Behandlung und drei Nachtests (Rost, 2013, S. 141). Das Design erweist sich grundsätzlich als geeignet, um die hier vorliegenden Forschungsfragen zu beantworten und die Hypothesen zu prüfen. Durch das Design können die kurz- und langfristige Wirksamkeit der Fördermaßnahme evaluiert, mögliche Prädiktoren identifiziert und Aussagen zur Implementationsqualität getroffen werden. Der Zeitpunkt und der Umfang der Durchführung der Maßnahme werden von den Lehrkräften als passend empfunden.

Grenzen zeigt das gewählte Design bei der Evaluation der Effektivität der *einzelnen* implementierten Fördermethoden (zum Beispiel Direkte Instruktion, Advance Organizer etc.). Dieses Forschungsanliegen sollte in sich anschließenden Studien – unter Berücksichtigung eines geeigneten forschungsmethodischen Ansatzes (zum Beispiel Multiple-Baseline-Design) – verfolgt werden. Auch mögliche Zuwendungseffekte werden hier nicht kontrolliert. Hier wäre die gleichzeitige Behandlung der Kontrollgruppe mit einem anderen Förderkonzept (zum Beispiel psychomotorische Förderung) angezeigt (Klauer, 2011b).

Die eingesetzte Lernverlaufsdiagnostik (siehe Abbildung 43) ist im Rahmen der vorliegenden Untersuchung als *Methode* und nicht als Erhebungsinstrument im Rahmen des Forschungsdesigns implementiert. Um den Schülerinnen und Schüler zuverlässiges Feedback zu geben, werden die Testbögen im Vorfeld umfassend auf ihre Güte überprüft (Käter, 2014; Rensing, 2014; Rensing, Käter, Käter & Hillenbrand, 2017; Wunram, 2014). Im Rahmen der Evaluation der Wirksamkeit findet die Lernverlaufsdiagnostik jedoch keine weitere Berücksichtigung. Eine Evaluation der Fördermaßnahme könnte zukünftig formativ durch CBM erfolgen, in der auch die Kontrollgruppe die CBM-Bögen bearbeitet. Für die Evaluation der langfristigen Wirksamkeit können die CBM auch in der Zeit nach der Förderung in regelmäßigen Abständen umgesetzt werden. Dies wäre ein Ausblick für Folgestudien.

Auswahl und Einsatz der Erhebungsinstrumente

Das Forschungsdesign wird – bis auf die beiden Implementationsfragebögen – mit standardisierten Erhebungsinstrumenten versehen, die den hohen wissenschaftlichen Ansprüchen an die Güte von Testverfahren genügen (Döring & Bortz, 2016d) (siehe Kapitel 8.2). Durch den Einsatz standardisierter Instrumente kann gewährleistet werden, dass die zugrundegelegten Dimensionen zuverlässig, objektiv und reliabel erfasst werden. Wie die Tabelle 13 zeigt, ist der Heidelberger Rechentest 1–4 nahezu das einzige standardisierte Instrument, das zu Beginn der fünften Klasse im Rahmen einer Gruppentestung eingesetzt werden kann, das die – wie hier definierten – mathematischen Basiskompetenzen testet und einen ökonomischen zeitlichen Umfang aufweist. Der HRT 1–4 kommt zudem auch in den jüngsten Evaluationen der Rechenprogramme „Calcularis“ (Käser et al., 2013) und „Meister CODY“ (Kuhn et al., 2016) zum Einsatz, was zeigt, dass es sich um ein in der Evaluationsforschung durchaus angewandtes Instrument handelt. Die Autoren des HRT 1–4 nehmen dabei die Rechengeschwindigkeit als validen Indikator für mathematische Kompetenz an. Dieser Ansatz wird auch in weiteren Studien verfolgt: In der LOGIK-Studie werden Mathematikaufgaben mit Zeitbegrenzung gelöst, was Stern (2013, S. 147) als „validen Indikator“ für mathematische Kompetenz sieht. Ebenso werden in der LAU-Studie (Lehmann, Husfeldt, et al., 2001) die Mathematikaufgaben in Form von Speed-Tests gestellt, was grundsätzlich für die Eignung dieses diagnostischen Zugangs im HRT 1–4 spricht. Da der Heidelberger Rechentest 1–4 mit den 240 zu lösenden Items in der Subskala Rechenoperationen einen großen Entwicklungsspielraum bietet, kann er in der vorliegenden Untersuchung über die vier Messzeitpunkte in der fünften Klasse eingesetzt werden. Im Rahmen der vorliegenden Untersuchung bearbeitet im Follow-Up-II-Messzeitpunkt kein Schüler bzw. keine Schülerin alle 240 Items. Um zu überprüfen, inwiefern die Fördermaßnahme auch die mathematische Problemlösekompetenz fördert, wäre der Einsatz eines weiteren standardisierten Instruments, das eben diesen Bereich zuverlässig misst, denkbar. Hier würde sich zum Beispiel der Schweizer Rechentest 4–6 (Subtest Textaufgaben) eignen. Im Rahmen des Gesamtprojekts, in das die vorliegenden Untersuchung eingebettet ist, werden in den Follow-Up-I- und Follow-Up-II-Erhebungen drei Textaufgaben aus dem Basis-Math 4–8 (Moser Opitz et al., 2010) gestellt, die sich aber im Nachhinein als wenig nützlich und aussagekräftig erweisen, zumal sie auch nicht für den Einsatz in Gruppentestungen vorgesehen sind. Die Ergebnisse werden deswegen nicht weiter berücksichtigt.

Bezüglich der kognitiven Kompetenzen kommt in der vorliegenden Untersuchung der erste Teil des CFT 20-R mit verlängerter Testzeit zum Einsatz. Das Verfahren misst die fluide Intelligenz, also das logische Denken. Denkbar wäre zukünftig der Einsatz eines Instruments, das die sog. mehrdimensionale Intelligenz erfasst (zum Beispiel WISC-IV), um die Wirksamkeit der Förderung in Abhängigkeit weiterer Intelligenzmerkmale analysieren zu können (zum Beispiel Sprachverständnis, Verarbeitungsgeschwindigkeit). Auch die Berücksichtigung der Arbeitsgedächtniskapazitäten oder gar ihrer einzelnen Komponenten kann wertvolle Hinweise auf die differenzielle Wirksamkeit der Fördermaßnahme geben (Lambert, 2015, S. 85). Mehrdimensionale Intelligenztests sind jedoch im Regelfall in Einzeltestung durchzuführen und beanspruchen eine recht umfangreiche Testdauer. Die Zumutbarkeit der diagnostischen Erhebungen für die Schülerinnen und Schüler stellt ein wichtiges zu beachtendes Kriterium dar, sodass der Ein-

satz eines mehrdimensionalen Intelligenztests nur unter dessen Berücksichtigung erwogen werden sollte. Dennoch böte der Einsatz eines solchen Instruments die Möglichkeit, differenzierte Aussagen zur Wirksamkeit bzgl. verschiedener Facetten kognitiver Fähigkeiten zu leisten.

Bereits diskutiert ist, dass neben der Lehrereinschätzliste für das Lern- und Sozialverhalten auch eine Schülerselbsteinschätzung differenzierte Aussagen zu diesen Aspekten aus Sicht der Schülerinnen und Schüler ermöglichen kann. Der Einsatz des Implementationsfragebogens und auch die Erhebung der Lernfreude zeigen, dass die Schülerinnen und Schüler grundsätzlich dazu bereit sind, Auskunft über ihre Meinung und Haltung zu geben, sodass künftig auch eine Befragung zu den eigenen sozialen Kompetenzen und der Lernfähigkeit erwogen werden könnte. Da die Mathematiklehrkräfte die Schülerinnen und Schüler selbst erst am Anfang des Schuljahres kennenlernen, kommt die Lehrereinschätzliste erst nach Abschluss der Maßnahme zum Einsatz. Zu kritisieren ist hierbei, dass die Fördermaßnahme gerade die Schülerinnen und Schüler mit Schwierigkeiten im Lern- und Sozialverhalten im mathematischen Lernen unterstützt. Dies kann dazu geführt haben, dass einige Schülerinnen und Schüler ein besseres Lern- und Sozialverhalten zeigen als zuvor im regulären Unterrichtssetting und sie den Lehrkräften in dieser Zeitspanne deswegen nicht auffallen.

Es erscheint weniger günstig, die Lernfreude früher, zum Beispiel direkt am ersten Schultag nach dem Übergang in die fünfte Jahrgangsstufe zu erheben, da die Schülerinnen und Schüler sich nach den Sommerferien erst wieder an das Lernumfeld Schule in der Sekundarstufe I gewöhnen müssen, um dann zuverlässige Aussagen zur eigenen Lernfreude tätigen zu können. Zu beachten ist jedoch, dass die Freude an der Maßnahme einen Einfuss auf die Beantwortung der Items der Skala Lernfreude in der Post-Erhebung nehmen kann. Dieser Einfluss wird hier nicht kontrolliert. Da die Skala Aussagen zur allgemeinen Lernfreude umfasst, wäre in einer Replikation denkbar, diese auch zum Beispiel mathematikspezifisch zu erheben.

Bei den beiden Implementationsfragebögen handelt es sich um theoriegeleitete selbstentwickelte Instrumente. Zwar erreichen beide Fragebögen zufriedenstellende Cronbachs'-α-Werte (Schülerinnen und Schüler $\alpha = .924$.; Interventionsleiterinnen und -leiter $\alpha = .849$), doch sind die Gütekriterien nicht in dem Umfang wie bei umfassend geprüften standardisierten Verfahren erfüllt. Die Berücksichtigung der Implementationsqualität von Förderprogrammen und -konzepten ist in Deutschland bisher nicht fest verankert, sodass aktuell keine Verfahren oder Checklisten zur Erhebung von Implementationsqualität vorliegen. Die Zuverlässigkeit der Analyseergebnisse wird durch die unzureichende Erfüllung der Gütekriterien und der kleinen Stichprobe bei den Interventionsleiterinnen und -leiter eingeschränkt, sodass die Ergebnisse lediglich als tendenziell angesehen werden können. Dennoch kann die grundsätzliche Berücksichtigung dieses Aspekts als Stärke der vorliegenden Untersuchung betrachtet werden.

Die Lehrkräfte der Kontrollgruppe füllen zum Zeitpunkt der Post-Erhebung einen einseitigen Methodenfragebogen aus, durch den ermittelt werden soll, inwiefern in der Kontrollgruppe in den vier Wochen vergleichbare Inhalte oder Methoden der Fördermaßnahme zum Einsatz kommen. Der Bogen erweist sich im Nachhinein jedoch als wenig bis gar nicht aussagekräftig. Unklar ist zum Beispiel, ob die Lehrkräfte unter den Begrifflichkeiten auf dem Fragebogen die gleichen Methoden und Inhalte verstehen, wie sie im Rahmen des Projekts verwendet werden. Hier wären kurze Definitionen hilfreich gewesen, um zum Beispiel das gleiche Verständnis des Begriffs „Strategieinstruktion" zu gewährleisten. Zudem wäre die Verdeutlichung der Antworten durch Beispiele oder Erläuterungen seitens der Lehrkräfte unterstützend gewesen, um ihre

Antworten interpretieren zu können. Letztlich provoziert ein Fragebogen wie der Methodenfragebogen das Problem der sozialen Erwünschtheit, sodass die getätigten Aussagen bzgl. ihres Wahrheitsgehalts nicht kontrolliert werden können. Die Ergebnisse des Methodenfragebogens können im Rahmen der vorliegenden Untersuchung nicht zuverlässig ausgewertet werden.

Die vorliegende Studie ist grundsätzlich dem quantitativen Paradigma zuzuordnen. Es zeigt sich jedoch, dass an spezifischen Stellen Teilfragen alleine durch diesen methodischen Zugang nicht beantwortet werden können. Künftig würde sich der zusätzliche Einsatz von strukturierten Interviews zum Beispiel mit Schülerinnen und Schülern (partizipativer Forschungsansatz) oder den Interventionsleiterinnen und -leitern eignen, um tiefergehende Rückmeldung bzgl. der Fördermaßnahme zu erhalten. Hier zeigt die Auswertung der Implementationsfragebögen, dass qualitative Antworten letztlich ausbleiben. Die Verzahnung von quantitativer und qualitativer Forschung ist im Mixed-Methods-Ansatz zu verorten (Döring, 2014b).

Rekrutierung und Zusammensetzung der Stichprobe

Die Randomisierung der Stichprobe erfolgt in der vorliegenden Untersuchung auf Schulebene. Bei einer Randomisierung auf Klassenebene besteht die Schwierigkeit, dass ein schulinterner Austausch zwischen Lehrkräften über die Inhalte der Maßnahme erfolgen und somit eine Beeinflussung (insbesondere) der parallelen Kontrollgruppen-Klassen stattfinden kann. Zudem wird sich durch die Randomisierung auf Schulebene auch eine organisatorische Erleichterung bzgl. der Absprachen zwischen universitärer und schulischer Seite erhofft. Die ungleiche Verteilung der Schülerinnen und Schüler auf die beiden Untersuchungsgruppen ist darin zu begründen, dass im Vorfeld nicht bekannt ist, wie viele Klassen der Jahrgangsstufe fünf an den angeschriebenen Schulen bestehen. Nach der Rekrutierung zeigt sich, dass die Schulen der Kontrollgruppe mehr Klassen umfassen als die der Experimentalgruppe, sodass diese Untersuchungsgruppe insgesamt größer ist. Beide Untersuchungsgruppen weisen den benötigten Stichprobenumfang für die geplanten statistischen Verfahren auf. Die Analyse der Stichprobe (siehe Kapitel 9.1) zeigt, dass in den akademischen Kompetenzen keine Unterschiede zwischen den Untersuchungsgruppen bestehen. Dahingehend kann die Randomisierung als gelungen bezeichnet werden. Unterschiede gibt es jedoch in soziodemografischen Aspekten. In der Experimentalgruppe sprechen weniger Schülerinnen und Schüler ausschließlich Deutsch zu Hause, sind weniger in Deuschland geboren und in dieser Untersuchungsruppe gibt es auch tendenziell eher wenige Bücher zu Hause. Hier scheint es ein Gefälle zwischen den beiden Untersuchungsgruppen zu geben, was mit der Lage der einbezogenen Schulen der Experimentalgruppe in sozialschwachen Stadtteilen zusammenhängen kann. Für eine Folgestudie ist schon zum Zeitpunkt der Randomisierung der Schulen der Sozialindex der Stadtgebiete, in denen sich die Schulen befinden, zu berücksichtigen.

In der vorliegenden Untersuchung werden ausschließlich Ober- und (Integrierte) Gesamtschulen in die Rekrutierung der Stichprobe einbezogen, da Forschungsbefunde zeigen, dass sich in diesen Schulformen ein Großteil der Schülerinnen und Schüler befindet, der Defizite in den mathematischen Kompetenzen zeigt (Humbach, 2008; Mittelberg, 2004; Moser Opitz, 2005; Prenzel et al., 2012). Künftig wäre auch der Einbezug von berufsbildenden Schulen oder Gymnasien zu berücksichtigen. Für Förderschulen könnte sich eine Umsetzung der Maßnahme im Förderschwerpunkt Lernen oder Emotional-soziale Entwicklung anbieten. Da sich die Förderschule Schwerpunkt Lernen im Rahmen der UN-Konvention für Menschen mit Behinderungen

(Vereinte Nationen, 2006) derzeit im Auflösungsprozess befindet, wird eine Umsetzung im inklusiven Settings – wie vorliegend – als zukunftsweisender empfunden.

Die Analyse der Wirksamkeit in Abhängigkeit der Schulform (z. B. OBS vs. IGS) könnte eine weitere sich anschließende Studie darstellen. Die Rekrutierung könnte dann durch das Ziehen einer zweistufigen Klumpenstichprobe erfolgen (Döring & Bortz, 2016c, S. 316), in der dann innerhalb einer Schule einige Klassen die Kontroll- andere hingegen die Experimentalgruppe darstellen.

Wahl und Vorgehen der statistischen Analysen

Zur Überprüfung der kurz- und langfristigen Wirksamkeit der Fördermaßnahme wird die zweifaktorielle Varianzanalyse mit Messwiederholung gewählt. Der Vorteil der Varianzanalyse, die als wichtigstes Verfahren zur Evaluation von Experimenten gilt (Backhaus, Erichson, Plinka & Weiber, 2011, S. 158), ist die Vermeidung der *a*-Fehler-Kumulierung. Grundsätzlich erweist sich der Einsatz der Varianzanalyse als zielführend. Der Fokus der Untersuchung liegt darauf, die kurzfristige Wirksamkeit zum einen und die langfristige Wirksamkeit zum anderen zu analysieren. Aus diesem Grund wird eine Varianzanalyse mit dem Prä- und Post-Messzeitpunkt und eine weitere mit den Post-Follow-Up-I-Follow-Up-II-Messzeitpunkten gewählt. Es liegen somit Haupteffekte und Interaktionseffekte für diese beiden betrachteten Zeiträume vor, was eine differenzierte Beantwortung der Fragestellungen ermöglicht. Da die Varianzanalyse nur die Fälle einschließt, für die zu jedem einbezogenen Messzeitpunkt Daten vorliegen, sind die Stichproben in der Untersuchung der langfristigen Wirksamkeit oftmals kleiner. Alternativ wäre auch eine Varianzanalyse, die alle vier Messzeitpunkte beinhaltet, denkbar. Diese würde jedoch eine weitere Reduktion der Stichprobe beinhalten und zahlreiche Detailanalysen erfordern, um eventuelle Signifikanzen differenziert verorten zu können. Die Vergleichbarkeit der Effektstärke wird dadurch erschwert, dass die Varianzanalyse die Effektstärke in Form des η^2 ausgibt, welches in Folge in das häufiger anzutreffende Cohen's *d* transformiert werden muss. In der vorliegenden Untersuchung wird zur Interpretation der Effektstärken die eher strenge Klassifikation nach Cohen (1988) gewählt, da sie sich im pädagogisch-psychologischen Kontext als bewährt zeigt. Die weitere Ermittlung der Teststärke erweist sich als geeignet, um ggf. die Annahme der Nullhypothesen abzusichern.

Bzgl. der Analyse der Schülerinnen und Schüler mit ungünstigen Lernvoraussetzungen zeigt es sich als schwierig, dass es sich dabei um kaum disjunkte Gruppen handelt. In der Gruppe der Schülerinnen und Schüler mit einem geringen Vorwissen können sich auch welche mit geringen kognitiven Fähigkeiten oder einem geringen Lesevermögen befinden. Die Zusammensetzung der Gruppen der Schülerinnen und Schüler mit ungünstigen Lernvoraussetzungen garantiert lediglich, dass diejenigen einer Gruppe *ein* gleiches Kriterium erfüllen (zum Beispiel geringe kognitive Fähigkeiten). Hier wäre denkbar mittels Clusteranalysen „übergreifende" Gruppen zu identifizieren, die zum Beispiel insgesamt ein eher niedriges, mittleres oder hohes Entwicklungsniveau aufzeigen.

Die gewählten statistischen Verfahren für den Hypothesenkomplex III und IV erweisen sich insgesamt als zielführend.

11 Abschließende Bewertung und Ausblick

Im folgenden Kapitel werden allgemeine Forschungsdesiderata aufgezeigt, eine zusammenfasende Bewertung der vorliegenden Untersuchung vorgenommen, sich daran anschließende Forschungsideen präsentiert und abschließend ein Ausblick geschaffen.

ALLGEMEINE FORSCHUNGSDESIDERATA

Die Abhandlung des Begriffs „mathematische Basiskompetenzen“ (siehe Kapitel 2.1) zeigt, dass es sich hierbei um einen oft verwendeten Terminus handelt, der je nach Autorengruppe und Definition unterschiedliche Aspekte der mathematischen Entwicklung bzw. des Kompetenzerwerbs umfasst. Hier ist eine dringende Aufarbeitung und Systematisierung des Begriffs angezeigt, um zum Beispiel in der Erstellung von Meta-Analysen oder Reviews Fehlerquellen zu vermeiden. Denn letztlich ist es keinesfalls gleich, ob in einer vorliegenden Studie „Vorläuferfertigkeiten“ oder Inhalte des Grundschulcurriculums thematisiert werden und dabei die großen inhaltlichen Unterschiede unter dem Deckmantel des gewählten Kompetenzbegriffs verschwimmen.

Daran anschließend wird festgestellt, dass es in Deutschland im Bereich Mathematik für die Sekundarstufe I kein differenziertes Erwerbs- bzw. Kompetenzstufenmodell gibt. In Fragen des durch Bildungspläne curricular gesteuerten Kompetenzerwerbs in der Sekundarstufe I stellen sich Lehrkräfte oft die Frage, wo – sprich auf welchem Kompetenzniveau – sich die Schülerinnen und Schüler im Vergleich zu anderen befinden und wie sie sich individuell entwickeln. Lohnenswert wäre hier die Entwicklung und empirische Prüfung eines Modells, das sich an ein bestehendes, zum Beispiel an dem von Reiss (2004), anschließt und dies fortführt. Auch die grundsätzliche Unterscheidung zwischen *Erwerbs*- und *Entwicklungs*modellen muss hierbei zukünftig stärker diskutiert werden.

In Fragen der Förderung mathematischer Basiskompetenzen ermöglicht eine intensive Aufarbeitung des Forschungsstandes zu evidenzbasierten Methoden und Maßnahmen, gezielt und fundiert Fördermethoden für die eigene vorliegende Problemstellung zu identifizieren. Dabei erfordert die Abhandlung der Schrittfolge nach Freeman und Sugai (2013) für Lehrkräfte eine zeitintensive Einarbeitung und ein gutes forschungsmethodisches Wissen. Hier ist insbesondere in Deutschland angezeigt, wissenschaftliche Befunde für die Praxis leichter zugänglich zu machen. Als Beispiel dienen hier die Vereinigten Staaten, die dies schon seit vielen Jahren in Form von frei zugänglichen Internetplattformen ermöglichen (siehe Tabelle 17).

In Deutschland ist dabei die „Grüne Liste Prävention“ ein erster Schritt. Bei einer komprimierten Darstellung von Befunden mit einem zusammenfassenden Deklarat „evidenzbasiert“ ist jedoch zu problematisieren, dass die alleinige Implementation einer Fördermethode in den Unterricht oftmals *nicht* die versprochene lernförderliche Wirkung zeigt. Denn wie Hattie und Kollegen (2013) oder auch das umfassende Follow-Through-Programm (Engelmann, 2007; Watkins, 1996) zeigen, ist die *Qualität* der Umsetzung einer Methode oder eines Programms für den Erfolg entscheidend. Auch Whitaker (2009, S. 19) schreibt, „Programme [sind] nie die Lösung und nie das Problem. [...] Kein Programm [führt] von sich aus zu einer [...] Verbesserung“. Zudem gilt gerade im Kontext von Evidenzbasierung, dass eine identifizierte Methode

immer konkret zum vorliegenden Einzelfall passen muss, was im Vorfeld eine intensive Analyse der Situation durch die beteiligten Akteure, wie zum Beispiel der Lehrkräfte, erfordert. Evidenzbasierung ist kein Garant für einen Fördererfolg.

Darüber hinaus wird festgestellt, dass in den Vereinigten Staaten Fördermethoden im Bereich Mathematik stärker standardisiert sind als in Deutschland. Zu nennen sind zum Beispiel das PALS-Programm für peergestütztes Lernen oder das IMPROVE- oder STAD-Konzept für kooperatives Lernen. Hier besteht in Deutschland ein großer Handlungsbedarf, diese evidenzbasierten Methoden für Lehrkräfte standardisiert und unterrichtsnah aufzuarbeiten, damit eine gute Qualität in der Umsetzung der Methoden gewährleistet sein kann. Dies kann insbesondere für den Kompetenzerwerb in inklusiven Lerngruppen eine große Chance darstellen.

Abschließende Bewertung der vorliegenden Untersuchung

Mit der vorliegenden Untersuchung wird in Deutschland erstmalig ein universelles Programm zur Förderung mathematischer Basiskompetenzen zu Beginn der Sekundarstufe I vorgestellt, das sich durch die Implementation evidenzbasierter Fördermethoden kennzeichnet. Die Befunde zeigen, dass sich die Schülerinnen und Schülern mit einem geringen Vorwissen, geringen kognitiven Fähigkeiten, einer geringen Lesekompetenz und einem risikobehafteten Lernverhalten durch die vierwöchige Fördermaßnahme signifikant in ihren mathematischen Basiskompetenzen verbessern. Für die Schülerinnen und Schüler mit geringen Lesekompetenzen kann sogar ein langfristiger Entwicklungsvorsprung bis zum Ende des fünften Schuljahres festgestellt werden. Die Fördermaßnahme führt also bei einem Großteil der Schülerinnen und Schülern – zumindest kurzfristig – zu einer Verbesserung der mathematischen Basiskompetenzen. Die differenzierte Aufarbeitung des aktuellen Forschungsstandes zu evidenzbasierten Fördermethoden anhand der Schrittfolge nach Freeman und Sugai (2013) kann im Rahmen der vorliegenden Untersuchung als lohnenswert bezeichnet werden. Inwiefern durch die Fördermaßnahme auch die Anschlussfähigkeit an den Sekundarstufen-I-Unterricht gewährleistet werden kann, muss im Rahmen einer Replikation in Form einer Längsschnittstudie evaluiert werden. Es zeigt sich darüber hinaus, dass die Lernfreude der Schülerinnen und Schüler ein Prädiktor für die kurzfristige Wirksamkeit der Fördermaßnahme ist. Dieser Befund verdeutlicht die Bedeutsamkeit der Motivation für einen Lernerfolg auch im stark curricular gesteuerten Sekundarstufen-I-Unterricht (siehe das Handlungsmodell nach Matthes, 2006). Zudem zeigen sich sowohl die Interventionsleiterinnen und -leiter als auch die Schülerinnen und Schüler mit der Fördermaßnahme insgesamt zufrieden, was für eine gelungene Umsetzung und eine grundsätzliche Umsetzungsfähigkeit der Fördermaßnahme in den schulischen Alltag spricht.

Um künftig für mehr Schülerinnen und Schüler eine langfristige Wirksamkeit der Fördermaßnahme zu erreichen, sollten Programm-Modifikationen erwogen werden. Zum einen könnten regelmäßige Auffrischungssitzungen im Anschluss an die Fördermaßnahme, in denen die Inhalte mit den aktuellen Unterrichtsthemen verknüpft werden, das langfristige Behalten der Rechenstrategien unterstützen. Auch eine inhaltliche Schwerpunktsetzung (zum Beispiel auf den Bereich Multiplikation) ist im Anschluss an eine Eingangsdiagnostik und die darin identifizierten mathematischen Schwierigkeiten der Schülerinnen und Schüler denkbar. Zudem kann eine aufgearbeitete Bereitstellung der zentralen Fördermethoden mit Handlungsanweisungen für die Mathematiklehrkräfte erwogen werden, sodass sie diese auch weiterhin in ihrem regulären Mathematikunterricht einsetzen können.

Die Ergebnisse der Evaluation zeigen, dass die bisherige Konzeption zur Förderung der mathematischen Basiskompetenzen für die Schülerinnen und Schüler mit einem risikobehafteten Sozialverhalten und mit einem Unterstützungsbedarf im Bereich Lernen nicht ausreicht. Auch die Bewertung der Implementationsqualität durch die Interventionsleiterinnen und -leiter weist auf eine erforderliche stärkere Differenzierung der Inhalte und Materialien gerade für die lernschwachen Schülerinnen und Schüler hin. Die Schülerinnen und Schüler mit massiven Kompetenzdefiziten benötigen auf einer *weiteren* selektiven Ebene im Rahmen des Responsiven Handlungsmodells verstärkte Unterstützung in ihrem Kompetenzerwerb. Für die selektive Ebene ist die Verknüpfung mit bereits vorhandenen evaluierten Förderprogrammen (zum Beispiel „Rechenspiele mit Elfe und Mathis II“ oder „Calcularis“) oder die eigene Weiterentwicklung der hier vorliegenden Fördermaßnahme anzudenken. Im Rahmen einer eigenen Weiterentwicklung wären zum Beispiel fokussierte und differenzierte Kleingruppenarbeiten innerhalb der bestehenden Erarbeitungsphasen der Fördermaßnahme realisierbar. Ebenso können hier die bereits vorhandenen evidenzbasierten Programme der selektiven und indizierten Stufe implementiert werden. Der gemeinsame inklusive Unterricht innerhalb der Lerngruppe ist und bleibt trotz der Erweiterung einer selektiven Stufe erklärtes Ziel. Die Programmmodifikationen müssten in der Folge einer weiteren Evaluation unterzogen werden. Zudem wird angemerkt, dass die Programme „Mathe sicher können“ sowie „Kalkulie“ aufgrund fehlender Evaluationsergebnisse bisher keiner Evidenzstufe zugeordnet werden können.

Weiterführende Forschungsideen

Die vorliegende Arbeit fokussiert insbesondere auf die Wirksamkeit bei Schülerinnen und Schülern, die *ungünstige* Voraussetzungen für den mathematischen Kompetenzerwerb aufweisen. Da aber der Erwerb mathematischer Basiskompetenzen zu Beginn der Sekundarstufe I noch nicht abgeschlossen ist (siehe Kapitel 2.4), kann im Rahmen weiterer Analysen auch die Wirksamkeit bei durchschnittlichen und leistungsstarken Schülerinnen und Schüler differenziert betrachtet werden. Denn auch wenn sich der Blick im Rahmen inklusiver Bildung oftmals auf die schwachen Schülerinnen und Schüler richtet, dürfen die anderen Schülerinnen und Schüler weder randständig betrachtet oder gar vergessen werden. Denn auch sie haben ein Recht auf bestmögliche Förderung. Hier schließt sich auch die Frage nach der Wirksamkeit bzgl. weiterer Schulformen an. Hier könnten zum Beispiel auch berufsbildende Schulen oder Gymnasien in den Blick genommen werden (siehe Kapitel 2.4.2 und 2.4.3).

Ein weiteres Forschungsanliegen betrifft die Analyse der Wirksamkeit der implementierten Methoden. In der vorliegenden Fördermaßnahme werden die Methoden in den Förderstunden „hintereinander“ implementiert, sodass die Wirksamkeit bisher nur auf das Gesamtkonzept zurückgeführt werden kann. Hier ist in weiteren Forschungsarbeiten die differenzierte Wirksamkeit der einzelnen Methoden spezifisch in den Blick zu nehmen, um Hinweise dafür finden zu können, welche Methode bei welcher spezifischen Problemlage den größtmöglichen lernförderlichen Effekt erzielen kann. Die Aufarbeitung des aktuellen Forschungsstandes zeigt zudem, dass zahlreiche Desiderata in der Evaluation der hier implementierten Methoden im Rahmen *inklusiver* Bildung vorliegen (siehe Kapitel 4.3.3). Aussagen zur differenzierten Wirksamkeit der Methoden können im Rahmen dieser Untersuchung nicht getätigt werden.

AUSBLICK

Abschließend stellt sich die Frage, wie die entwickelte Fördermaßnahme entlang des „Implementationspfades“ (Euler & Sloane, 1998, S. 321) Einzug in den schulischen Alltag halten kann. Zahlreiche Befunde und Diskussionen zeigen die Schwierigkeit und die Herausforderungen, wissenschaftliche Erkenntnisse in der schulischen Praxis Realität werden zu lassen (Gräsel & Parchmann, 2004; Križan & Vossen, 2016; Reinmann-Rothmeier & Mandl, 1998). Ein möglicher Schritt in Richtung Implementation wäre, das Förderprogramm dem Fachpublikum im Rahmen von Fort- oder Weiterbildungen sowie in universitärer Lehre vorzustellen. Der unmittelbare Kontakt mit den (angehenden) Lehrkräften ermöglicht es, Diskussionen direkt aufzugreifen und auf (Rück-)Fragen reagieren zu können. Auch die Präsentation auf Tagungen, Kongressen und Fachpraxistagen kann das Bekanntwerden unterstützen. Zudem kann die Fördermaßnahme in praxisnahen Zeitschriften für die Sekundarstufe I („Mathematik 5–10“, „mathematik lehren“) mit exemplarischen Materialien veröffentlicht werden, die einen ersten Einblick ermöglichen.

Durch die vorliegende Untersuchung wird dem Bedarf an (inklusiver) Förderung mathematischer Basiskompetenzen in der Sekundarstufe I begegnet und dabei der verstärkten Forderung des Einsatzes evidenzbasierter Methoden entsprochen. Durch die Fördermaßnahme gelingt es einem Großteil Schülerinnen und Schülern, sich in ihren mathematischen Basiskompetenzen zu verbessern. Ob hiermit auch die langfristige Anschlussfähigkeit an den Sekundarstufen-I-Unterricht gewährleistet wird, kann abschließend nicht beantwortet werden. Dazu sind weitere Längsschnittstudien erforderlich, die unter dem Stichwort „mathematische Handlungsfähigkeit“ auch die Anwendungsorientierung der erworbenen Kompetenzen berücksichtigen. Die Diskussion um evidenzbasierte Förderung im Rahmen inklusiver Bildung aufgreifend, kann abschließend resümiert werden:

> *„Natürlich bleibt das Risiko bestehen, dass eine gut überprüfte Methode bei einem bestimmten Lerner nicht greift. Allerdings ist dieses Risiko deutlich kleiner als bei einem Vorgehen, das nicht auf empirischen Befunden beruht.“*
> *(Kuhl & Euker, 2016, S. 25)*

12 Literaturverzeichnis

Adams, G. & Carnine, D. (2003). Direct Instruction. In H.L. Swanson, K. Harris & S. Graham (Hrsg.), *Handbook of Learning Disabilities* (S. 403–416). New York: Guilford Press.

Ahrbeck, B. (2014). *Inklusion. Eine Kritik* (2. Auflage.). Stuttgart: Kohlhammer.

Alarcón, M., DeFries, J., Gillis Light, J. & Pennington, B. (1997). A Twin Study of Mathematics Disability. *Journal of Learning Disabilities*, *30* (6), 617–623.

Allinder, R.M., Bolling, R.M., Oats, R.G. & Gagnon, W.A. (2000). Effects of Teacher Self-Monitoring on Implementation of Curriculum-Based Measurement and Mathematics Computation Achievement of Students with Disabilities. *Remedial and Special Education*, *21* (4), 219–226.

Alloway, T., Gathercole, S., Adams, A.-M. & Willis, C. (2005). Working Memory abilities in children with special educational needs. *Educational & Child Psychology*, *22* (4), 56–67.

American Psychiatric Association. (2013). *Diagnostischer und statistischer Leitfaden psychischer Störungen V (Diagnostic and Statistical Manual of Mental Disorders V)*.

Anderman, E.M. & Midgley, C. (1997). Changes in achievement goal orientations, perceived academic competence, and grades across the transition to middle-level schools. *Contemporary Educational Psychology*, *22*, 269–298.

Von Aster, M., Käser, T., Kohn, J., Kucian, K., Rauscher, L. & Vögeli, C. (2016). „Calcularis“ - Eine adaptive Lernsoftware zur Matheförderung. In M. Hasselhorn & W. Schneider (Hrsg.), *Förderprogramme für Vor- und Grundschule* (S. 225–247). Hogrefe.

Von Aster, M., Schweiter, M. & Weinhold Zulauf, M. (2007). Rechenstörung bei Kindern. Vorläufer, Prävalenz und psychische Symptome. *Zeitschrift fur Entwicklungspsychologie und Pädagogische Psychologie*, *39* (2), 85–96.

Auer, M., Gruber, G., Wimmer, H. & Mayringer, H. (2005). *Salzburger Lese-Screening für die Klassenstufen 5-8*. Göttingen: Hogrefe.

Ausubel, D.P. (1960). The use of advance organizers in the learning and retention of meaningful verbal material. *Journal of Educational Psychology*, *51* (5), 267–272.

Ayres, P. & Paas, F. (2012). Cognitive Load Theory: New Directions and Challenges. *Applied Cognitive Psychology*, *26*, 827–832.

Bach, H. (1977). *Sonderpädagogik im Grundriss*. Berlin: Marhold.

Backhaus, K., Erichson, B., Plinka, W. & Weiber, R. (2011). *Multivariate Analysemethoden*. Berlin: Springer.

Baddeley, A. (2010a). Short term memory. In A. Baddeley, M.W. Eysenck & C. Anderson (Hrsg.), *Memory* (S. 19–40). Hove: Psychology Press.

Baddeley, A. (2010b). Working Memory. In A. Baddeley, M.W. Eysenck & M.C. Anderson (Hrsg.), *Memory* (S. 41–68). Hove: Psychology Press.

Baker, S., Gersten, R. & Lee, D.-S. (2002). A Synthesis of Empirical Research on Teaching Mathematics to Low-Achieving Students. *The Elementary School Journal*, *103* (1), 51–73.

Balzer, L. & Jäger, R.S. (2001). Fachleistung Mathematik in MARKUS. *Empirische Pädagogik*, *15* (4), 535–551.

Barth, K.H. (2006). *Lernschwächen früh erkennen im Vorschul- und Grundschulalter*.

München: Reinhardt.

Barth, K.H. (2010). Früherkennung und Prävention schulischer Lernstörungen im Übergangsbereich Kindergarten - Grundschule. *Rechenschwäche, Rechenstörung, Dyskalkulie* (S. 52–68). Graz: Leykam.

Basendowski, S. (2013). *Die soziale Frage an (mathematische) Grundbildung. Eine empirische Studie zu dem Wesen, der Funktion und der Relevanz mathematischer Kompetenzen in einfachen Erwerbstätigkeiten sowie Analysen für didaktische Implikationen*. Bad Heilbrunn: Klinkhardt.

Becker, M., Lüdtke, O., Trautwein, U. & Baumert, J. (2006). Leistungszuwachs in Mathematik. Evidenz für einen Schereneffekt im mehrgliedrigen Schulsystem? *Zeitschrift für Pädagogische Psychologie*, *20* (4), 233–242.

Becker, W.C. & Engelmann, S. (1996). Sponsor Findings from Project Follow Through. *Effective School Practices*, *1*, 33–42.

Berk, L.E. (2011). *Entwicklungspsychologie*. München: Pearson.

Berkeley, S., Bender, W.N., Peaster, L.G. & Saunders, L. (2009). Implementation of Response to Intervention. *Journal of Learning Disabilities*, *42* (1), 85–95.

Blackorby, J., Chorost, M., Garza, N. & Guzman, A.-M. (2004). *The Academic Performance of Elementary and Middle School Students with Disabilities*. Verfügbar unter: http://www.seels.net/designdocs/engagement/04_SEELS_outcomes_C4_8-19-04.pdf

Blum, W. (2006). Einführung. *Bildungsstandards Mathematik: konkret. Sekundarstufe I: Aufgabenbeispiele, Unterrichtsanregungen, Fortbildungsideen* (S. 14–32). Berlin: Cornelsen.

Blum, W. & Leiß, D. (2005). Modellieren im Unterricht mit der „Tanken"-Aufgabe. *mathematik lehren*, *128*, 18–21.

Blumenthal, Y. & Mahlau, K. (2015). Effektiv fördern - Wie wähle ich aus? Ein Plädoyer für die Evidenzbasierte Praxis in der schulischen Sonderpädagogik. *Zeitschrift für Heilpädagogik*, *66*, 408–421.

Böck, W. & Migutsch, R. (1974). *Der Übergang von der Grundschule zu weiterführenden Schulen. Ergebnisse einer empirischen Untersuchung in Oberösterreich*. Wien: Ketterl.

Born, A. & Oehler, C. (2013). *Kinder mit Rechenschwäche erfolgreich fördern. Ein Praxishandbuch für Eltern, Lehrer und Therapeuten*. Stuttgart: Kohlhammer.

Bos, W. & Gröhlich, C. (2010). *KESS 8. Kompetenzen und Einstellungen von Schülerinnen und Schülern am Ende der Jahrgangsstufe 8*. Münster: Waxmann.

Böttinger, T. (2015). Bilder lügen nicht – Zur Gefährdnung individuumszentrierter Pädagogik. *Zeitschrift für Heilpädagogik*, *66*, 422–431.

Braun, E. (2011). Rahmenbedingungen für „Übergänge" in der Kulturarbeit mit Menschen mit Behinderung. In W. Bleher (Hrsg.), *Transfer. Band 5. Übergänge im Bildungssystem: biografisch - institutionell - thematisch* (S. 43–63). Baltmannsweiler: Schneider Verlag Hohengehren.

Bremm, M.H. & Kühn, R. (1992). *Rechentest RT 9+*. Göttingen: Hogrefe.

Breuer, H. & Weuffen, M. (2000). *Lernschwierigkeiten am Schulanfang. Schuleingangsdiagnostik zur Früherkennung und Frühförderung*. Weinheim und Basel: Beltz.

Brown-Chidsey, R. & Steege, M.W. (2010). *Response to Intervention. Principles and Strategies for Effective Practice*. New York: The Guilford Press.

Brown, M., Askew, M., Hodgen, J., Rhodes, V. & Wiliam, D. (2003). Individual and chort progression in learning numeracy ages 5-11: results from the Leverhulme 5-year longitudinal study. In F. Lin & J. Guo (Hrsg.), *Proceedings of the International Conference on Science and Mathematics Learning* (S. 81–109). Taipei: National Taiwan Normal University.

Bühner, M. & Ziegler, M. (2009). *Statistik für Psychologen und Sozialwissenschaftler*. Hallbergmoos: Pearson.

Bundesministerium für Bildung und Forschung. (2010). *Der Übergang von der Grundschule in die weiterführende Schule. Leistungsgerechtigkeit und regionale, soziale und ethnisch-kulturelle Disparitäten*. Bonn.

Bundesministerium für Bildung und Forschung. (2016). *Bekanntmachung des Bundesministerium für Bildung und Forschung der Richtlinie zur Förderung der Forschung zur „Qualifizierung der pädagogischen Fachkräfte für inklusive Bildung“*. Bundesanzeiger vom 29.03.2016.

Bundschuh, K. (2014). *Einführung in die sonderpädagogische Diagnostik*. München: UTB.

Burns, M.K., Appleton, J.J. & Stehouwer, J.D. (2005). Meta-Analytic Review of Responsiveness-to-Intervention: Examining Field-Based and Research-Implemented Models. *Journal of Psychoeducational Assessment*, *23*, 381–394.

Busch, J., Oranu, N., Schmidt, C. & Grube, D. (2013). Rechenschwäche im Grundschulalter: Reduzierte Verfügbarkeit basalen arithmetischen Faktenwissens und Belastung des Arbeitsgedächtnisses bei Drittklässlern. *Lernen und Lernstörungen*, *2*, 217–227.

Büttner, G., Warwas, J. & Adl-Amini, K. (2012). Kooperatives Lernen und Peer Tutoring im inklusiven Unterricht. *Zeitschrift für Inklusion online*, *1/2*.

Calhoon, M.B. & Fuchs, L. (2003). The Effects of Peer-Assisted Learning Strategies and Curriculum-Based Measurement on the Mathematics Performance of Secondary Students with Disabilities. *Remedial and Special Education*, *24* (4), 235–245.

Casale, G., Hennemann, T. & Grosche, M. (2015). Zum Beitrag der Verlaufsdiagnostik für eine evidenzbasierte sonderpädagogische Praxis am Beispiel des Förderschwerpunkts der emotionalen und sozialen Entwicklung. *Zeitschrift für Heilpädagogik*, *66* (7), 325–334.

Clarke, D. (2001). *Understanding, Assessing, and Developing Young Children's Mathematical Thinking: Research as a Powerful Tool for Professional Growth*. Sydney.

Clausen-Suhr, K. (2009). *Mit Baldur ordnen, zählen, messen*. Oberursel: Finken.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale: N.J.L. Erlbaum Associates.

Council for Exceptional Children. (2014). Standards for Evidence-Based Practices in Special Education. Arlington VA. Verfügbar unter: https://www.cec.sped.org/~/media/Files/Standards/Evidence based Practices and Practice/EBP FINAL.pdf

Cross, C., Woods, T. & Schweingruber, H. (2010). Mathematics Learning in Early Childhood: Paths Toward Excellence and Equity. Committee on Early Childhood Mathematics (Hrsg.), Washington D. C.: National Academic Press.

van Damme, J., De Fraine, B., Van Landeghem, G., Opdenakker, M.-C. & Onghena, P. (2002). A New Study on Educational Effectiveness in Secondary Schools in Flanders: An Introduction. *School Effectiveness and School Improvement*, *13* (4), 383–397.

Dehaene, S. (1992). Varieties of numerical abilities. *Cognition*, *44*, 1–42.

Dember, W. (1974). Motivation and the cognitive revolution. *American Psychologist*, *29*.

Deno, S.L. (1985). Curriculum-Based Measurement: The Emerging Alternative. *Exceptional Children*, *52* (3), 219–232.

Deno, S.L. (1993). Curriculum-based Measurement. In J.J. Kramer (Hrsg.), *Curriculum-based measurement* (S. 1–23). Lincoln: Buros Institute of Mental Measurements, University of Nebraska-Lincoln.

Deno, S.L. (2003). Curriculum-Based Measures: Development and Perspectives. *The Journal of Special Education*, *37* (3), 184–192.

Denton, K. & West, J. (2002). *Children's Reading and Mathematics Achievement in Kindergarten and First Grade*. Washington D. C.

DeSimone, J.R. & Parmar, R.S. (2006). Issues and Challenges for Middle School Mathematics Teachers in Inclusion Classrooms. *School Science and Mathematics*, *106* (8), 338–348.

Deutsche Gesellschaft für Kinder- und Jugendpsychiatrie, P. und P. e. V. (2011). *Kinder- und Jugendpsychiatrie und Psychotherapie in Deutschland. Die Versorgung von psychisch kranken Kindern, Jugendlichen und ihren Familien*. Stuttgart: Schattauer.

Deutscher, T., Prediger, S. & Selter, C. (2013). Mathe sicher können - Sicherung mathematischer Basiskompetenzen in der unteren Sekundarstufe I. *Beiträge zum Mathematikunterricht*, 252–255.

Diehl, K. & Hartke, B. (2007). Curriculumnahe Lernfortschrittsmessung. *Sonderpädagogik*, *37* (4), 195–211.

Diehl, K. & Hartke, B. (2012). *Inventar zur Erfassung der Lesekompetenz im 1. Schuljahr (IEL-1). Ein curriculumbasiertes Verfahren zur Abbildung des Lernfortschritts*. Göttingen: Hogrefe.

Diehl, K., Hartke, B. & Knopp, E. (2009). Curriculum-Based Measurement & Leerlingonderwijsvolgesysteem - Konzepte zur theoriegeleiteten Lernfortschrittsmessung im Anfangsunterricht Deutsch und Mathematik? *Zeitschrift für Heilpädagogik*, *4*, 122–130.

Diehl, K., Mahlau, K., Voß, S. & Hartke, B. (2010). Das Rügener-Integrations-Modell (RIM). *Gemeinsam leben*, *19* (3), 162–167.

Ditton, H. (2007). *Kompetenzen und Laufbahnen im Schulsystem*. Münster: Waxmann.

Ditton, H. & Krüsken, J. (2006). Der Übergang von der Grundschule in die Sekundarstufe I. *Zeitschrift für Erziehungswissenschaft*, *9* (3), 348–372.

Ditton, H. & Krüsken, J. (2009). Denn wer hat, dem wird gegeben werden? Eine Längsschnittstudie zur Entwicklung schulischer Leistungen und den Effekten der sozialen Herkunft in der Grundschulzeit. *Journal for educational research online*, *1*, 33–61.

Doabler, C.T. & Fien, H. (2013). Explicit Mathematics Instruction: What Teachers Can Do for Teaching Students With Mathematics Difficulties. *Intervention in School and Clinic*, *48* (5), 276–285.

Domitrovich, C.E. & Greenberg, M. (2000). The Study of Implementation: Current Findings From Effective Programs that Prevent Mental Disorders in School-Aged Children. *Journal of Educational and Psychological Consultation*, *11* (2), 193–221.

Donovan, S. & Cross, T. (2002). *Minority students in special and gifted education*. Washington D. C.: National Academic Press.

Döring, N. (2014a). Forschungsethik. In M.A. Wirtz (Hrsg.), *Dorsch – Lexikon der Psychologie* (S. 590). Bern: Verlag Hans Huber. Verfügbar unter:

https://portal.hogrefe.com/dorsch/forschungsethik/

Döring, N. (2014b). Mixed-Methods-Ansatz. In M.A. Wirtz (Hrsg.), *Dorsch – Lexikon der Psychologie* (S. 1101). Bern: Verlag Hans Huber.

Döring, N. & Bortz, J. (2016a). Evaluationsforschung. In N. Döring & J. Bortz (Hrsg.), *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften* (S. 975–1037). Berlin: Springer.

Döring, N. & Bortz, J. (2016b). Datenanalyse. In N. Döring & J. Bortz (Hrsg.), *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften* (S. 597–784). Berlin: Springer.

Döring, N. & Bortz, J. (2016c). Stichprobenziehung. In N. Döring & J. Bortz (Hrsg.), *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften* (S. 291–320). Berlin: Springer.

Döring, N. & Bortz, J. (2016d). Qualitätskriterien in der empirischen Sozialforschung. In N. Döring & J. Bortz (Hrsg.), *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften* (S. 81–120). Berlin: Springer.

Dornheim, D. (2008). *Prädiktion von Rechenleistung und Rechenschwäche. Der Beitrag von Zahlen-Vorwissen und allgemein-kognitiven Fähigkeiten*. Berlin: Logos.

Drüke-Noe, C., Möller, G., Pallack, A., Schmidt, S., Schmidt, U., Sommer, N. et al. (2011). *Basiskompetenzen Mathematik für Alltag und Berufseinstieg am Ende der allgemeinen Schulpflicht*. Berlin: Cornelsen.

DUDEN (2013). Kompetenz. *DUDEN Online-Wörterbuch*. Bibliographisches Institut GmbH.

DUDEN (2016a). Entwicklung. *DUDEN Online-Wörterbuch*. Bibliographisches Institut GmbH.

DUDEN (2016b). Erwerb. *DUDEN Online-Wörterbuch*. Bibliographisches Institut GmbH.

DUDEN (2016c). Evidenzbasiert. *DUDEN Online-Wörterbuch*. Bibliographisches Institut GmbH.

Duncan, G.J., Claessens, A., Huston, A., Pagani, L., Engel, M., Sexton, H. et al. (2007). School Readiness and Later Achievement. *Developmental Psychology*, *43* (6), 1428–1446.

Durlak, J. (1997a). *Successful Prevention Programs for Children and Adolescents*. New York: Plenum Press.

Durlak, J. (1997b). Primary Prevention Programs in Schools. In T.H. Ollendick & R. Prinz (Hrsg.), *Advances in Clinical Child Psychology. Volume 19* (S. 283–314). New York: Plenum Press.

Durlak, J., Weissberg, R.P., Dymnicki, A.B., Taylor, R.D. & Schellinger, K.B. (2011). The Impact of Enhancing Students' Social and Emotional Learning: A Meta-Analysis of School-Based Universal Interventions. *Child Development*, *82* (1), 405–432.

Eder, F. (1995). *Das Befinden von Kindern und Jugendlichen in der Schule. Forschungsbericht im Auftrag des BMUK*. Innsbruck: Studienverlag.

Education Commission of the States. (2013). *Math in the Early Years. A Strong Predictor for Later School Success*. Verfügbar unter: http://www.du.edu/kennedyinstitute/media/documents/math-in-the-early-years.pdf

Ehlert, A., Fritz, A., Arndt, D. & Leutner, D. (2013). Arithmetische Basiskompetenzen von Schülerinnen und Schülern in den Klassen 5 bis 7 der Sekundarstufe. *Journal für Mathematik-Didaktik*, *34* (2), 237–263.

Ehlert, A., Schroeders, U. & Fritz, A. (2012). Kritik am Diskrepanzkriterium in der Diagnostik von Legasthenie und Dyskalkulie. *Lernen und Lernstörungen*, *1* (3), 169–184.

Engelmann, S. (2007). *Follow Through Evaluation. Teaching Needy Kids in Our Backward System*. ADI Press.

Ennemoser, M. (2014). Response to Intervention als schulisches Förderkonzept. In G.W. Lauth, M. Grünke & J.C. Brunstein (Hrsg.), *Interventionen bei Lernstörungen* (S. 505–516). Göttingen: Hogrefe.

Ennemoser, M. & Krajewski, K. (2013). Entwicklungsorientierte Diagnostik mathematischer Basiskompetenzen in den Klassen 5 bis 9. In M. Hasselhorn, A. Heinze, W. Schneider & U. Trautwein (Hrsg.), *Diagnostik mathematischer Kompetenzen* (Test und Trends, Band 11, S. 225–240). Göttingen: Hogrefe.

Ennemoser, M., Krajewski, K. & Schmidt, S. (2011). Entwicklung und Bedeutung von Mengen-Zahlen-Kompetenzen und eines basalen Konventions- und Regelwissens in den Klassen 5 bis 9. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, *43* (4), 228–242.

Etscheidt, S.K. (2007). Progress Monitoring: Legal Issues and Recommendations for IEP Teams. *Teaching Exceptional Children*, *59* (1/2), 56–60.

Euler, D. & Sloane, P.F.E. (1998). Implementation als Problem der Modellversuchsforschung. *Unterrichtswissenschaft*, *4*, 312–326.

Fingerle, M. & Ellinger, S. (2008). *Sonderpädagogische Förderprogramme im Vergleich*. Stuttgart: Kohlhammer.

Fischbach, A., Schuchardt, K., Brandenburg, J., Klesczewski, J., Balke-Melcher, C., Schmidt, C. et al. (2013). Prävalenz von Lernschwächen und Lernstörungen: Zur Bedeutung der Diagnosekriterien. *Lernen und Lernstörungen*, *2* (2), 65–76.

Fitzell, S.G. (2011). *RTI Strategies for Secondary Teachers*. Thousand Oaks: Corwin.

Foegen, A. (2006). Monitoring Student Progress in Algebra. State College: Adequate Yearly Progress and Students with IEPs: A Focus on Improving Student Achievement Through Effective Practices.

Foegen, A., Jiban, C. & Deno, S.L. (2007). Progress Monitoring Measures in Mathematics: A Review of the Literature. *Journal of Special Education*, *41* (121–139).

Forbringer, L.L. & Fuchs, W. (2014). *RtI in Math. Evidence-Based Interventions for Struggling Students*. New York: Routledge.

Freeman, J. & Sugai, G. (2013). Identifying Evidence-based Special Education Interventions From Single-Subject Research. *Teaching Exceptional Children*, *45*, 1–12.

Freesemann, O. (2014). *Schwache Rechnerinnen und Rechner fördern*. Wiesbaden: Springer.

Friedrich, K. & Laag, N. (2015). Das Pilotprojekt „Inklusive Grundschule“: Hintergründe und Zielstellungen. In D.I. für I.P.F. (DIPF) Landesinstitut für Schule und Medien Berlin-Brandenburg (LISUM), Universität Potsdam (Hrsg.), *Inklusives Lernen und Lehren im Land Brandenburg Abschlussbericht zur Begleitforschung des Pilotprojekts „Inklusive Grundschule“* (S. 25–30). Ludwigsfelde: Landesinstitut für Schule und Medien Berlin-Brandenburg (LISUM).

Fritz, A. & Ricken, G. (2008). *Rechenschwäche*. München: Ernst Reinhardt Verlag.

Fritz, A. & Ricken, G. (2009). Grundlagen des Förderkonzeptes „Kalkulie“. In A. Fritz, G. Ricken & S. Schmidt (Hrsg.), *Handbuch Rechenschwäche* (S. 374–395). Weinheim und Basel: Beltz Verlag.

Fritz, A., Ricken, G., Balzer, L., Leutner, D. & Willmes, K. (2012). *Zentrale arithmetische Konzepte mathematischen Wissens im Vorschulalter: Ein integratives Kompetenzstufenmodell*. Unveröffentliches Manuskript.

Fuchs, D. & Fuchs, L. (2006). Introduction to Response to Intervention: What, why, and how valid is it? *Reading Research Quarterly*, *41* (1), 93–99.

Fuchs, L. & Fuchs, D. (1990). The role of skill analysis in curriculum-based measurement in math. *School Psychology Review*, *53* (3), 199–208.

Fuchs, L., Fuchs, D., Hamlett, C. & Appleton, A. (2002). Explicitly Teaching for Transfer: Effects on the Mathematical Problem-Solving Performance of Students with Mathematics Disabilities. *Learning Disabilities Research & Practice*, *17* (2), 90–106.

Fuchs, L., Fuchs, D. & Karns, K. (2001). Enhancing kindergartners' mathematical development: Effects of peer-assisted learning strategies. *The Elementary School Journal*, *101*, 495–510.

Fuchs, L., Fuchs, D., Powell, S.R., Seethaler, P.M., Cirino, P. & Fletcher, J.M. (2008). Intensive Intervention for Students with Mathematics Disabilities: Seven Principles of Effective Practice. *Learning Disability Quarterly*, *31*, 79–92.

Fuchs, L., Fuchs, D., Yazdian, L. & Powell, S.R. (2002). Enhancing first-grade children's mathematical development with Peer-Assisted Learning Strategies. *School Psychology Review*, *31* (4), 569–583.

Fuson, K.C. (1988). *Children's Counting and Concepts of Number*. New York: Springer.

Gaidoschik, M. (2004). Förderung rechenschwacher Kinder: Wege und Irrwege. *Dyskalkulie-Symposium*. Klagenfurt. Verfügbar unter: http://www.recheninstitut.at/wp-content/uploads/2004/10/wege-irrwege.pdf

Gaidoschik, M. (2008). „Rechenschwäche" in der Sekundarstufe: Was tun? *Journal für Mathematikdidaktik*, *29* (3/4), 287–294.

Gaidoschik, M. (2009). *Rechenschwäche verstehen - Kinder gezielt fördern. Ein Leitfaden für die Unterrichtspraxis*. Buxtehude: Persen.

Gaidoschik, M. (2010). *Rechenschwäche vorbeugen. 1. Schuljahr: Vom Zählen zum Rechnen*. Wien: G & G Verlagsgesellschaft mbH.

Galton, M., Comber, C. & Pell, T. (2002). The consequences of transfer for pupils: attitudes and attainment. In L. Hargreaves & M. Galton (Hrsg.), *Transfer from the Primary Classroom. Twenty years on*. (S. 131–158). London: Routledge.

Gathercole, S. & Pickering, S. (2001). Working deficits in children with special educational needS. *British Journal of Special Education*, *28* (2), 89–97.

Gaupp, N., Zoelch, C. & Schumann-Hengsteler, R. (2004). Defizite numerischer Basiskompetenzen bei rechenschwachen Kindern der 3. und 4. Klassenstufe. *Zeitschrift für Pädagogische Psychologie/German Journal of Educational Psychology*, *18* (1), 31–42.

Geary, D. (2000). From infancy to adulthood: the development of numerical abilities. *European Child & Adolescent Psychiatry*, *9* (2), 11–16.

Gebhardt, M., Oelkrug, K. & Tretter, T. (2013). Das mathematische Leistungsspektrum bei Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf in der Sekundarstufe. Ein explorativer Querschnitt der fünften bis neunten Klasse in Münchener Förderschulen. *Empirische Sonderpädagogik*, *2*, 130–143.

Gelman, R. & Gallistel, C.R. (1978). *The Child's Understanding of Number*. Cambridge:

Harvard University Press.

Gerlach, M., Fritz, A. & Leutner, D. (2013). *MARKO-T Mathematik- und Rechenkonzepte im Vor- und Grundschulalter - Training*. Göttingen: Hogrefe.

Gerlach, M., Fritz, A., Ricken, G. & Schmidt, S. (2007). *KALKULIE-T. Diagnose- und Trainingsprogramm für rechenschwache Kinder*. Berlin: Cornelsen.

Gerrig, R.J. (2015). *Psychologie*. Hallbergmoos: Pearson.

Gersten, R., Chard, D.J., Jayanthi, M., Baker, S., Morphy, P. & Flojo, J. (2009). Mathematics Instruction for Students with Learning Disabilities: A Meta-Analysis of Instructional ComponentS. *Review of Educational Research*, *79* (3), 1202–1242.

Gersten, R., Jordan, N. & Flojo, J. (2005). Early Identification and Interventions for Students with Mahematics Difficulties. *Journal of Learning Disabilities*, *38* (4), 293–304.

Gerster, H.-D. (2003). Schwierigkeiten beim Erwerb arithmetischer Konzepte im Anfangsunterricht. In F. Lenart, N. Holzer & H. Schaupp (Hrsg.), *Rechenschwäche, Rechenstörung, Dyskalkulie* (S. 154–160). Graz: Leykam.

Gerster, H.-D. (2009). Probleme und Fehler bei den schriftlichen Rechenverfahren. In A. Fritz, G. Ricken & S. Schmidt (Hrsg.), *Handbuch Rechenschwäche* (S. 269–284). Weinheim: Beltz.

Gerster, H.-D. & Schulz, R. (2004). *Schwierigkeiten beim Erwerb mathematischer Konzepte im Anfangsunterricht. Bericht zum Forschungsprojekt Rechenschwäche - Erkennen, Beheben, Vorbeugen. Erweiterte Ausgabe*. Freiburg: Pädagogische Hochschule Freiburg. Institut für Mathematik und Informatik und ihre Didaktiken.

Gold, A. (2011). *Lernschwierigkeiten. Ursachen, Diagnostik, Intervention*. Stuttgart: Kohlhammer.

Gold, A. (2015). Lernschwierigkeiten. Wie man einen pädagogisch-psychologischen Dauerbrenner immer wieder aufs Neue befeuern kann. *Zeitschrift für Pädagogische Psychologie*, *29* (3–4), 123–132.

Götz, L., Lingel, K. & Schneider, W. (2013). *Deutscher Mathematiktest für fünfte Klassen. DEMAT 5+*. Göttingen: Hogrefe.

Graney, S.B. & Shinn, M.R. (2005). Effetcs of Reading Curriculum-Based Measurement (R-CBM) Teacher Feedback in General Education Classrooms. *School Psychology Review*, *34* (2), 184–201.

Gräsel, C. & Parchmann, I. (2004). Implementationsforschung - oder: der steinige Weg, Unterricht zu verändern. *Unterrichtswissenschaft*, *32* (3), 196–214.

Griebel, W. & Niesel, R. (2011). *Beiträge zur Bildungsqualität: Übergänge verstehen und begleiten: Transitionen in der Bildungslaufbahn von Kindern*. Berlin: Cornelsen.

Groeger-Roth, F. (2014). *Die „Grüne Liste Prävention“ und der „Wegweiser Prävention“*. Karlsruhe: 19. Deutscher Präventionstag, 12.05.2014. Verfügbar unter: http://www.praeventionstag.de/dokumentation/download.cms?id=1833

Gross, M., Jäncke, L., Von Aster, M. & Meyer, M. (2015). *Calcularis*. Zürich: Dybuster. Verfügbar unter: http://www.dybuster.com/calcularis/was-ist-es

Grube, D. (2005). Entwicklung des Rechnens im Grundschulalter. In M. Hasselhorn, H. Marx & W. Schneider (Hrsg.), *Diagnostik von Mathematikleistungen. Test und Trends N. F. Band 4* (S. 105–124). Göttingen: Hogrefe.

Grube, D. (2006). *Entwicklung des Rechnens im Grundschulalter*. Münster: Waxmann.

Grube, D. (2008). Rechenschwäche. In W. Schneider & M. Hasselhorn (Hrsg.), *Handbuch der Pädagogischen Psychologie* (S. 642–652). Göttingen: Hogrefe.

Grünke, M. (2006). Zur Effektivität von Fördermethoden bei Kindern und Jugendlichen mit Lernstörungen. Eine Synopse vorliegender Metaanalysen. *Kindheit und Entwicklung*, *15* (4), 239–254.

Grünke, M. & Grosche, M. (2014). Lernbehinderung. *Interventionen bei Lernstörungen* (S. 76–89). Göttingen: Hogrefe.

Gugel, G. (2006). *Methoden-Manual „Neues Lernen"*. Weinheim und Basel: Beltz.

Haag, L. (2014). Tutorielles Lernen. *Interventionen bei Lernstörungen* (S. 465–471). Göttingen: Hogrefe.

Haag, L. & Streber, D. (2011). Tutorielles Lernen. *Empirische Pädagogik*, *25* (3), 358–369.

Haffner, J., Baro, K., Parzer, P. & Resch, F. (2005). *Heidelberger Rechentest. HRT 1–4*. Göttingen: Hogrefe.

Hagenauer, G. (2010). Kurzzeitintervention versus Langzeitintervention. *Pädagogische Interventionsforschung* (S. 243–251). Weinheim: Juventa.

Hammer, S., Reiss, K., Lehner, M.C., Heine, J.-H., Sälzer, C. & Heinze, A. (2016). Mathematische Kompetenz in PISA 2015: Ergebnisse, Veränderungen und Perspektiven. *PISA 2015* (S. 219–248). Münster: Waxmann.

Hartke, B. (2005). Schulische Prävention - welche Maßnahmen haben sich bewährt? *Zeitschrift für Heilpädagogik*, *56* (12), 470–481.

Harwell, J.M. & Williams Jackson, R. (2008). *The Complete Learning Disabilities Handbook. Ready-to-Use Strategies & Activities for Teaching Students with Learning Disabilities*. San Francisco: Jossey-Bass.

Hascher, T. & Schmitz, B. (2010). *Pädagogische Interventionsforschung*. Weinheim: Juventa.

Hasemann, K. (2001). Results of an empirical study with 5 to 7 year old children. *Developments in Mathematics Education in German-speaking Countries. Seletced Papers from the conference of didactics of mathematics* (S. 31–40). Hildesheim.

Hasemann, K. & Gasteiger, H. (2014). *Anfangsunterricht Mathematik*. Heidelberg, Berlin: Springer.

Hasselhorn, M., Schneider, W. & Trautwein, U. (2014). *Mathematisches Strategiewissen für 5. und 6. Klassen. MAESTRA 5-6+*. Göttingen: Hogrefe.

Hasselhorn, M. & Schuchardt, K. (2006). Epidemiologie Lernstörungen. Eine kritische Skizze zur Epidemiologie. *Kindheit und Entwicklung*, *15* (4), 208–215.

Hattie, J. (2014). *Lernen sichtbar machen für Lehrpersonen*. Hohengehren: Schneider.

Hattie, J., Beywl, W. & Zierer, K. (2013). *Lernen sichtbar machen*. Baltmannsweiler: Schneider Verlag Hohengehren.

Hattie, J. & Timperley, H. (2007). The Power of Feedback. *Review of Educational Research*, *77* (1), 81–112.

Von Hehn, U. (2014). Korrelationskoeffizient nach Pearson. Medistat Medizinische Statistik. Verfügbar unter: https://www.medistat.de/glossar/korrelation-assoziation/korrelationskoeffizient-nach-pearson/

Heimlich, U. (2009). *Lernschwierigkeiten*. Bad Heilbrunn: Klinkhardt.

Heimlich, U. & Wember, F. (2014). Lernbehinderung, Lernbeeinträchtigung, Lernstörung. In

F. Wember, R. Stein & U. Heimlich (Hrsg.), *Handlexikon Lernschwierigkeiten und Verhaltensstörungen* (S. 51–54). Stuttgart: Kohlhammer.

Hein, J., Bzufka, M.W. & Neumärker, K.J. (2000). The specific disorder of arithmetic skills. Prevalence studies in a rural and an urban population sample and their clinico-neuropsychological validation. *European Child & Adolescent Psychiatry*, *9* (9), 87–101.

Heinecke-Müller, M. (2014). Lernschwierigkeiten. In M.A. Wirtz (Hrsg.), *DORSCH Lexikon der Psychologie* (S. 1010). Bern: Verlag Hans Huber.

Hellmich, F. (2007). Möglichkeiten der Förderung mathematischer Vorläuferfertigkeiten im vorschulischen Bereich. *Bildungsforschung*, *4* (1), 1–16.

Helmke, A., Jäger, R.S., Balzer, L., Hahl, A., Hosenfeld, Ridder, A. et al. (2000). *Erster Ergebnisbericht MARKUS*. Landau: Fachbereich Psychologie und Zentrum für empirische pädagogische Forschung der Universität Koblenz-Landau.

Helmke, A., Jäger, R.S., Balzer, L., Hosenfeld, I., Ridder, A. & Schrader, F.-W. (2002). *Das Projekt MARKUS. Mathematik-Gesamterhebung Rheinland-Pfalz: Kompetenzen, Unterrichtsmerkmale, Schulkontext*. (Ministerium für Bildung und Frauen und Jugend Rheinland-Pfalz, Hrsg.). Landau: Verlag Empirische Pädagogik.

Hesse, I. & Latzko, B. (2011). *Diagnostik für Lehrkräfte. 2. Auflage*. Opladen: UTB.

Hillenbrand, C. (2014). Evidenzbasierte Praxis im Förderschwerpunkt emotional-soziale Entwicklung. In R. Stein (Hrsg.), *Inklusion im Förderschwerpunkt emotionale und soziale Entwicklung* (S. 170–215). Stuttgart: Kohlhammer.

Hillenbrand, C. (2015). Evidenzbasierung sonderpädagogischer Praxis: Widerspruch oder Gelingensbedingung. *Zeitschrift für Heilpädagogik*, *66* (7), 312–324.

Hillenbrand, C., Hennemann, T., Hens, S. & Hövel, D. (2015). *„Lubo aus dem All!“ – 1. und 2. Klasse*. München: Reinhardt.

Hintz, A.-M., Melzer, C., Scheer, D., Paal, M. & Hillenbrand, C. (2015). Bridging the Gap Between Research and Practice - First Results of an Online Survey Focusing Teachers' Knowledge and Professional Needs in the Context of Inclusive Education in School. *Learning Disabilities Worldwide Conference in Venice, 08.-09.09.2015*. Venice.

Hinz, A. (2013). Inklusion – von der Unkenntnis zur Unkenntlichkeit!? - Kritische Anmerkungen zu einem Jahrzehnt Diskurs über schulische Inklusion in Deutschland. *Zeitschrift für Inklusion online*, *1*.

Vom Hofe, R., Hafner, T., Blum, W. & Pekrun, R. (2009). Die Entwicklung mathematischer Kompetenzen in der Sekundarstufe - Ergebnisse der Längsschnittstudie PALMA. *Mathematiklernen vom Kindergarten bis zum Studium. Kontinuität und Kohärenz als Herausforderung für den Mathematikunterricht*. (S. 125–146). Münster: Waxmann.

Hoffmann, E. (2009). Schulische Fachleistungen. In R. Lehmann & E. Hoffmann (Hrsg.), *BELLA. Berliner Erhebung arbeitsrelevanter Basiskompetenzen von Schülerinnen und Schülern mit Förderbedarf „Lernen“* (S. 31–63). Münster: Waxmann.

Holden-Pitt, L. (2007). *SEELS. What makes a difference? Influences on outcomes for students with disabilities*. U.S. Department of Education.

Hosp, M., Hosp, J.L. & Howell, K.W. (2007). *The ABCs of CBM*. New York: Guilford Press.

Huber, C. & Grosche, M. (2012). Das response-to-intervention-Modell als Grundlage für einen inklusiven Paradigmenwechsel in der Sonderpädagogik. *Zeitschrift für Heilpädagogik*, *8*, 312–322.

Huber, C., Grosche, M. & Schütterle, P. (2013). Inklusive Schulentwicklung durch response-

to-intervention (RTI) - Realisierungsmöglichkeiten des RTI-Konzepts im Förderbereich Lesen. *Gemeinsam leben*, *2*, 79–90.

Huber, H.D. (2004). Im Dschungel der Kompetenzen. In M. Huber, Hans Dieter; Lockemann, Bettina; Scheibel (Hrsg.), *Visuelle Netze. Wissensräume in der Kunst*. (S. 15–29). Ostfildern-Ruit: Hatje Cant.

Humbach, M. (2008). *Arithmetische Basiskompetenzen in der Klasse 10. Quantitative und qualitative Analysen*. Berlin: Verlag Dr. Köster.

Humphreys, C. & Parker, R. (2015). *Making Number Talks Matter*. Portland: Stenhouse.

Hussy, W., Schreier, M. & Echterhoff, G. (2010). *Forschungsmethoden*. Berlin: Springer.

Institut zur Qualitätsentwicklung im Bildungswesen. (2009). *Handreichung Vera 8 Mathematik*. Berlin: Humboldt Universität zu Berlin.

Isaacs, E.B., Edmonds, C.J., Lucas, A. & Gadian, D.G. (2001). Calculation difficulties in children of very low birthweight: a neural correlate. *Brain*, *124* (9), 1701–1707.

Isakson, K. & Jarvis, P. (1999). The adjustment of adolescents during the transition into high school: A short-term longitudinal study. *Journal of Youth and Adolescence*, *28*, 1–26.

Ise, E., Dolle, K., Pixner, S. & Schulte-Körne, G. (2012). Effektive Förderung rechenschwacher Kinder. Eine Metaanalyse. *Kindheit und Entwicklung*, *21* (3), 181–192.

Ise, E., Engel, R. & Schulte-Körne, G. (2012). Was hilft bei der Lese-Rechtschreibstörung? Ergebnisse einer Metaanalyse zur Wirksamkeit deutschsprachiger Förderansätze. *Kindheit und Entwicklung*, *21* (2), 122–136.

Jacobs, C. & Petermann, F. (2003). Dyskalkulie - Forschungsstand und Perspektiven. *Kindheit und Entwicklung*, *12* (4), 197–211.

Jacobs, C. & Petermann, F. (2007). *Rechenstörungen*. Göttingen: Hogrefe.

Jacobs, C. & Petermann, F. (2012). *Diagnostik von Rechenstörungen. Kompendien Psychologische Diagnostik. Band 7*. Göttingen: Hogrefe.

Jacobs, C. & Petermann, F. (2014). *Rechenfertigkeiten- und Zahlenverarbeitungs-Diagnostikum für die 2. bis 6. Klasse. RZD 2-6 (2. überarbeitete und erweiterte Auflage)*. Göttingen: Hogrefe.

Jank, W. & Meyer, H. (1994). *Didaktische Modelle*. Frankfurt a. M.: Cornelsen.

Johnson, D.W., Johnson, R.T. & Stanne, M.E. (2000). *Cooperative Learning Methods: A meta-analysis*. Minneapolis: University of Minnesota, Cooperative Learning Center.

Jonkmann, K., Maaz, K., Neumann, M. & Gresch, C. (2010). Übergangsquoten und Zusammenhänge zu familiärem Hintergrund und schulischen Leistungen: Deskriptive Befunde. In Bundesministerium für Bildung und Forschung (Hrsg.), *Der Übergang von der Grundschule in die weiterführende Schule. Leistungsgerechtigkeit und regionale, soziale und ethnisch-kulturelle Disparitäten* (S. 123–150). Berlin: Bundesministerium für Bildung und Forschung - Referat Bildungsforschung.

Jordan, N. (2010). Early Predictors of Mathematics Achievement and Mathematics Learning Difficulties. Verfügbar unter: http://www.child-encyclopedia.com/sites/default/files/textes-experts/en/784/early-predictors-of-mathematics-achievement-and-mathematics-learning-difficulties.pdf

Jordan, N., Hanich, L. & Kaplan, D. (2003). A longitudinal study of mathematical competencies with specific mathematics difficulties versus children with comorbid mathematics and reading difficulties. *Child Development*, *74*, 834–850.

Jordan, R. & Stein, R. (2011). *Test zum mathematischen Textverständnis. TeMaTex*. Göttingen: Hogrefe.

Kame'Enui, E., Fien, H. & Korgesaar, J. (2014). Direct Instruction as Eo nomine and Contronym. Why the right Words and the Details Matter. In H.L. Swanson, K. Harris & S. Graham (Hrsg.), *Handbook of Learning Disabilities* (2nd Edition, S. 489–506).

Kanter, G.O. (1977). Lernbehinderungen und die Personengruppe der Lernbehinderten. In G.O. Kanter & O. Speck (Hrsg.), *Pädagogik der Lernbehinderten, Handbuch der Sonderpädagogik* (S. 34–64). Berlin: Marhold.

Käser, T., Baschera, G.-M., Kohn, J., Kucian, K., Richtmann, V., Grond, U. et al. (2013). Design and evaluation of the computer-based training program Calcularis for enhancing numerical cognition. *Frontiers in Psychology*, *4*, 1–13.

Käser, T., Busetoo, A.G., Baschera, G.-M., Kohn, J., Kucian, K., Von Aster, M. et al. (2012). Modelling and Optimizing the Process of Learning Mathematics. Proceedings of ITS. Chania, Greece.

Käter, C. (2014). *Entwicklung eines curriculumbasierten Messverfahrens (CBM) und dessen Überprüfung hinsichtlich des Testgütekriteriums Reliabilität. Unveröffentlichte Masterarbeit*. Oldenburg: Carl von Ossietzky Universität Oldenburg.

Käter, C., Käter, T., Martenstein, R. & Hillenbrand, C. (2016). Leitfadengestützte Konstruktion eines Instruments der Lernverlaufsdiagnostik (CBM) im Bereich Lesen. *Zeitschrift für Heilpädagogik*, *66*, 168–179.

Kaufman, P. & Bradbury, D. (1992). *Characteristics of AT-Risk Students in NELS:88. Contract Report*. Berkeley: National Center for Education Statistics. Verfügbar unter: http://nces.ed.gov/pubs92/92042.pdf

Keeley, P. & Tobey, C.R. (2011). *Mathematics Formative Assessment*. Thousand Oaks: Corwin.

Klauer, K.J. (2006). Erfassung des Lernfortschritts durch curriculumbasierte Messung. *Heilpädagogische Forschung*, *32* (1), 16–26.

Klauer, K.J. (2011a). Lernverlaufsdiagnostik - Konzept, Schwierigkeiten und Möglichkeiten. *Empirische Sonderpädagogik*, *3*, 207–224.

Klauer, K.J. (2011b). *Transfer des Lernens*. Stuttgart: Kohlhammer.

Klauer, K.J. (2014). Formative Leistungsdiagnostik: Historischer Hintergrund und Weiterentwicklung zur Lernverlaufsdiagnostik. In M. Hasselhorn, W. Schneider & U. Trautwein (Hrsg.), *Lernverlaufsdiagnostik (Bd. 12)* (S. 1–17). Göttingen: Hogrefe.

Klicpera, C., Schabmann, A. & Gasteiger-Klicpera, B. (2013). *Legasthenie - LRS* (4. Auflage.). Weimar: UTB.

Klieme, E. (2004). Was sind Kompetenzen und wie lassen sie sich messen? *Pädagogik*, *6*, 10–13.

Kluge, F. (2012). *Etymologisches Wörterbuch der deutschen Sprache*. Berlin: De Gruyter.

Koch, K. (2001). *Von der Grundschule in die Sekundarstufe. Band 2: Der Übergang aus Sicht der Lehrerinnen und Lehrer*. Opladen: Leske & Budrich.

Koch, K. (2008). Förderung mathematischer Kompetenzen. In M. Fingerle & S. Ellinger (Hrsg.), *Sonderpädagogische Förderprogramme im Vergleich* (S. 85–108). Stuttgart: Kohlhammer.

Koumoula, A., Tsironi, V., Stamouli, V., Bardani, I., Siapati, S., Graham, A. et al. (2004). An Epidemiological Study of Number Processing and Mental Calculation in Greek

Schoolchildren. *Learning Disabilities*, *37*, 377–388.

Krajewski, K. (2008). *Vorhersage von Rechenschwäche in der Grundschule*. Hamburg: Verlag Dr. Kovač.

Krajewski, K. (2013). Förderung des Zahlenverständnisses. In G. Lauth, M. Grünke & J. Brunstein (Hrsg.), *Interventionen bei Lernstörungen* (S. 199–208). Göttingen: Hogrefe.

Krajewski, K. & Ennemoser, M. (2010). Entwicklung mathematischer Basiskompetenzen in der Sekundarstufe. *Empirische Pädagogik*, *24* (4), 353–370.

Krajewski, K. & Ennemoser, M. (2013). Entwicklung und Diagnostik der Zahl-Größen-Verknüpfung zwischen 3 und 8 Jahren. In M. Hasselhorn, A. Heinze, W. Schneider & U. Trautwein (Hrsg.), *Diagnostik mathematischer Kompetenzen* (S. 41–65). Göttingen: Hogrefe.

Krajewski, K., Grüßing, M. & Peter-Kopp, A. (2009). Die Entwicklung mathematischer Kompetenzen bis zum Beginn der Grundschulzeit. In A. Heinze & M. Grüßing (Hrsg.), *Mathematiklernen vom Kindergarten bis zum Studium. Kontinuität und Kohärenz als Herausforderung für den Mathematikunterricht*. (S. 17–34). Münster: Waxmann.

Krajewski, K., Nieding, G. & Schneider, W. (2013). *MZZ Mengen, zählen, Zahlen. Die Welt der Mathematik verstehen*. Berlin: Cornelsen.

Krajewski, K. & Schneider, W. (2006). Mathematische Vorläuferfertigkeiten im Vorschulalter und ihre Vorhersagekraft für die Mathematikleistungen bis zum Ende der Grundschulzeit. *Psychologie in Erziehung und Unterricht*, *53*, 246–262.

Krause, C. (1987). Zur Genese des Selbstbildes bei neun- bis dreizehnjährigen Schülern. *Psychologische Praxis, Ergänzungsheft*, 57–65.

Krauthausen, G. (2009). Entwicklung arithmetischer Fertigkeiten und Strategien - Kopfrechnen und halbschriftliches Rechnen. In A. Fritz, G. Ricken & S. Schmidt (Hrsg.), *Handbuch Rechenschwäche* (S. 100–117). Weinheim: Beltz.

Krauthausen, G. & Scherer, P. (2008). *Einführung in die Mathematikdidaktik*. Heidelberg: Spektrum Akademischer Verlag.

Krautz, J. (2009). Bildung als Anpassung ? Das Kompetenz-Konzept im Kontext einer ökonomisierten Bildung. *Fromm Forum*, *13*, 87–100.

Kristen, C., Römmer, A., Müller, W. & Kalter, F. (2005). *Longitudinal Studies for Education Reports: European and North American Examples. Education Reform Volume 10*. (Federal Ministry of Education and Research, Hrsg.). Bonn.

Križan, A. & Vossen, A. (2016). Evidenzbasierung in Schulen durch Verzahnung von Wissenschaft und Praxis erreichen. *Zeitschrift für Heilpädagogik*, *67*, 79–90.

Krizinger, H. & Günther, T. (2013). Lesen, Schreiben, Rechnen - gibt es Unterschiede zwischen den Geschlechtern? *Lernen und Lernstörungen*, *2* (1), 35–49.

Kroesbergen, E.H. & Van Luit, J.E.H. (2003). Mathematics Interventions for Children with Special Educational Needs. *Remedial and Special Education*, *24* (2), 97–114.

Krüger, H.-H., Köhler, S.-M., Pfaff, N. & Zschach, M. (2007). Die Bedeutung des Übergangs von der Grundschule in die Sekundarstufe I für Freundschaftsbeziehungen von Kindern. *Zeitschrift für Pädagogik*, *53* (4), 509–521.

Kubinger, K.D. & Holocher-Ertl, S. (2014). *Adaptives Intelligenzdiagnostikum 3. AID-3*. Göttingen: Hogrefe.

Kucian, K. & Von Aster, M. (2013). Dem Gehirn beim Rechnen zuschauen. *Rechenstörungen bei Kindern. Neurowissenschaft, Psychologie und Pädagogik* (S. 59–77). Göttingen:

Vandenhoeck & Ruprecht.

Kucian, K., Loenneker, T., Dietrich, T., Dosch, M., Martin, E. & Von Aster, M. (2006). Impaired neural networks for approximate calculation in dyscalculic children: a functional MRI study. *Behavioral and Brain Functions*, *5*, 2–31.

Kuhl, J. & Euker, N. (2016). Evidenzbasierte Förderung von Kindern und Jugendlichen mit intellektueller Beeinträchtigung - Chancen und Grenzen des Konzepts. In J. Kuhl & N. Euker (Hrsg.), *Evidenzbasierte Diagnostik und Förderung von Kindern und Jugendlichen mit intellektueller Beeinträchtigung* (S. 19–38). Göttingen: Hogrefe.

Kuhl, J. & Hecht, T. (2014). Prävention von Lernschwierigkeiten durch die Implementierung von Diagnostik und Förderung. *Zeitschrift für Heilpädagogik*, *11*, 406–415.

Kuhn, J.-T., Holling, H., Raddatz, J. & Dobel, C. (2016). *Meister CODY: Ein computergestützter Test und Training für Kinder mit Dyskalkulie*. Westfälische Wilhelms-Universität Münster & kaasa health. Verfügbar unter: https://cdn.meistercody.com/files/dyskalkulie_test_und_training_wissenschaft_de1.pdf

Kultusministerkonferenz. (2002). *PISA 2000 – Zentrale Handlungsfelder*. Verfügbar unter: https://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/2002/2002_10_07-Pisa-2000-Zentrale-Handlungsfelder.pdf

Kultusministerkonferenz. (2003). *Bildungsstandards im Fach Mathematik für den Mittleren Schulabschluss*. München: Luchterhand.

Kultusministerkonferenz. (2004). *Bildungsstandards im Fach Mathematik für den Primarbereich*. München: Luchterhand.

Kunsch, C.A., Jitendra, A.K. & Sood, S. (2007). The Effects of Peer-Mediated Instruction in Mathematics for Students with Learning Problems: A Research Synthesis. *Learning Disabilities Research & Practice*, *22* (1), 1–12.

Kunter, M. & Trautwein, U. (2013). *Psychologie des Unterrichts*. Paderborn: Schöningh.

Lambert, K. (2015). *Rechenschwäche*. Göttingen: Hogrefe.

Lambert, R., Algozzine, B. & Mc Gee, J. (2014). Effects of Progress Monitoring on Math Performance of At-Risk Students. *British Journal of Education, Society & Behavioural Science*, *4* (4), 527–540.

Landerl, K. & Fussenegger, B. (2006). Dyskalkulie und Legasthenie: Same or different? *Sprache - Stimme - Gehör*, *30*, 165–170.

Landerl, K. & Kaufmannn, L. (2013). *Dyskalkulie*. München: Ernst Reinhardt Verlag.

Lauth, G. (2014). Allgemeine Lernschwäche. In G.W. Lauth, M. Grünke & J.C. Brunstein (Hrsg.), *Interventionen bei Lernstörungen* (S. 66–76). Göttingen: Hogrefe.

Lauth, G., Brunstein, J.C. & Grünke, M. (2014). Lernstörungen im Überblick. In G.W. Lauth, M. Grünke & J.C. Brunstein (Hrsg.), *Interventionen bei Lernstörungen* (S. 17–31). Göttingen: Hogrefe.

Lauth, G. & Mackowiak, K. (2006). Lernstörungen. *Kindheit und Entwicklung*, *15* (4), 199–207.

Lebens, M. & Lauth, G. (2014). Direkte Instruktion. In G.W. Lauth, M. Grünke & J.C. Brunstein (Hrsg.), *Interventionen bei Lernstörungen* (S. 418–428). Göttingen: Hogrefe.

Lehmann, R., Husfeldt, V. & Peek, R. (2001). Lernstände und Entwicklungen im Fach Mathematik - Ergebnisse der Hamburger Untersuchung (LAU) in den Jahrgangsstufen 5 und 6. In G. Kaiser, N. Knoche, D. Lind & W. Zillmer (Hrsg.), *Leistungsvergleiche im Mathematikunterricht. Ein Überblick über aktuelle nationale Studien* (S. 29–50).

Hildesheim: franzbecker.

Lehmann, R. & Lenkeit, J. (2008). *ELEMENT. Erhebung zum Lese- und Mathematikverständnis in den Jahrgangsstufen 4 bis 6*. Berlin: Humboldt Universität zu Berlin. Verfügbar unter: http://www.berlin.de/imperia/md/content/sen-bildung/schulqualitaet/element6_bericht_komplett.pdf?start&ts=1210843547&file=element6_bericht_komplett.pdf

Lehmann, R., Peek, R., Gänsfuß, R., Lutkat, S., Mücke, S. & Barth, I. (2001). *QuaSUM. Qualitätsuntersuchung an Schulen zum Unterricht in Mathematik. Ergebnisse einer repräsentativen Befragung*. Potsdam: Ministerium für Bildung, Jugend und Sport des Landes Brandenburg.

Leiß, D. & Blum, W. (2006). Beschreibung zentraler mathematischer Kompetenzen. In W. Blum, C. Drüke-Noe, R. Hartung & O. Köller (Hrsg.), *Bildungsstandards Mathematik: konkret. Sekundarstufe I: Aufgabenbeispiele, Unterrichtsanregungen, Fortbildungsideen* (S. 33–50). Berlin: Cornelsen.

Lembke, E.S., McMaster, K. & Stecker, P.M. (2010). The Prevention Science of Reading Research within a Response-to-Intervention Model. *Psychology in the Schools*, *47* (1).

Lembke, E.S. & Stecker, P.M. (2007). *Curriculum-Based Measurement in Mathematics. An Ecidence-Based Formative Assessment Procedure*. Portsmouth: RMC Research Corporation Center on Instruction.

Lembke, E.S. & Stormont, M. (2005). Using Research-Based Practices To Support Students With Diverse Needs in General Education Settings. *Psychology in the Schools*, *42* (8), 761–763.

Lenhard, A. & Lenhard, W. (2011). Evidenzbasierte Mathematikförderung: Evaluation der Rechenspiele mit Elfe und Mathis. Verfügbar unter: http://www.psychometrica.de/rechenspiele_evaluation.html

Lenhard, A., Lenhard, W., Schug, M. & Kowalski, A. (2011). Computerbasierte Mathematikförderung mit den „Rechenspielen mit Elfe und Mathis I“ - Vorstellung und Evaluation eines Computerprogramms für Erst- bis Drittklässler. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, *43* (2), 79–88.

Lenhard, W. & Lenhard, A. (2016a). Evidenzbasierte Förderung schulischer Fertigkeiten am Computer: Lernspiele mit Elfe und Mathis. In M. Hasselhorn & W. Schneider (Hrsg.), *Förderprogramme für Vor- und Grundschule* (S. 87–113). Göttingen: Hogrefe.

Lenhard, W. & Lenhard, A. (2016b). Berechnung von Effektstärken. Verfügbar unter: http://www.psychometrica.de/effektstaerke.html

Lenhard, W., Lenhard, A. & Lingel, K. (2010). *Rechenspiele mit Elfe und Mathis II*. Göttingen: Hogrefe.

Lenz, B.K. & Deshler, D.D. (2008). Jugendliche mit Lernstörungen: Auf den Spuren von The Educator’s Enigma. In B. Y. L. Wong (Hrsg.), *Lernstörungen verstehen* (S. 439–467). Berlin: Spektrum Akademischer Verlag.

Lerkkanen, M.-K., Rasku-Puttonen, H. & Nurmi, J.-E. (2005). Mathematical performance predicts progress in reading comprehension among 7-year olds. *European Journal of Psychology of Education*, *20* (2), 121–137.

Liegmann, A. (2014). Der Grundschulübergang als komplexer Prozess. Empirische Forschung zum Übergang von der Primarstufe in die Sekundarstufe. *Facetten von Übergängen im Bildungssystem* (S. 35–48).

Linderkamp, F. & Grünke, M. (2007). Lern- und Verhaltensstörungen: Klassifikation,

Prävalenz und Prognostik. In F. Linderkamp & M. Grünke (Hrsg.), *Lern- und Verhaltensstörungen – Genese, Diagnostik und Intervention* (S. 14–28). Weinheim: Psychologie Verlags Union.

Lobeck, A. (1992). *Rechenschwäche. Geschichtlicher Rückblick, Theorie und Therapie*. Luzern: Schweizerische Zentralstelle für Heilpädagogik SZH.

Lobeck, A., Frei, M. & Blöchlinger, R. (1990). *Schweizer Rechentest 4.-6. Klasse. SR 4-6*. Göttingen: Hogrefe.

Lorenz, J.H. (1993). Eine Rechenstörung früh erkennen. *Grundschule*, *26* (6), 8–9.

Lorenz, J.H. (2014). Rechenschwäche. In G.W. Lauth, M. Grünke & J.C. Brunstein (Hrsg.), *Interventionen bei Lernstörungen* (S. 43–55). Göttingen: Hogrefe.

Lossau, N. (2016, März). Völlig verrechnet. *DIE ZEIT*, 18–19.

Lowenthal, A.B. & Bassett, D. (2012). Transition Assessment: Using Gap Analysis to Enhance Effective Transition Planning. *Intervention in School and Clinic*, *48* (1), 30–37.

Maaz, K. & Nagy, G. (2010). Der Übergang von der Grundschule in die weiterführenden Schulen des Sekundarschulensystems: Definition, Spezifikation und Quantifizierung primärer und sekundärer Herkunftseffekte. In Bundeministerium für Bildung und Forschung (Hrsg.), *Der Übergang von der Grundschule in die weiterführende Schule. Leistungsgerechtigkeit und regionale, soziale und ethnisch-kulturelle Disparitäten* (S. 151–180). Bonn: Bundeministerium für Bildung und Forschung.

Mähler, C. (2007). Arbeitsgedächtnisfunktionen bei lernbehinderten Kindern und Jugendlichen. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, *39* (2), 97–106.

Mähler, C. (2014). Diskrepanzkriterium. In M.A. Wirtz (Hrsg.), *DORSCH Lexikon der Psychologie*. Berlin: Verlag Hans Huber. Verfügbar unter: https://portal.hogrefe.com/dorsch/diskrepanzkriterium/

Martenstein, R., Melzer, C. & Reinck, C. (2013). Der Advance Organizer in der Grundschule. Grundlagen und Empfehlungen zum Einsatz. In H. Habeck (Hrsg.), *Attraktive Grundschule. Jedes Kind fordern und fördern* (S. 1–18). Stuttgart: Dr. Josef Raabe Verlags-GmbH.

Mastropieri, M.A., Scruggs, T.E., Davidson, T. & Rana, R.K. (2008). Unterrichtsinterventionen in Mathematik für Schüler mit Lernstörungen. In B.Y.L. Wong (Hrsg.), *Lernstörungen verstehen* (S. 309–311). Berlin: Springer.

Matthes, G. (2006). *Individuelle Lernförderung bei Lernstörungen. Potsdamer Studientexte - Sonderpädagogik*. Potsdam: Universität Potsdam.

Mayringer, H. & Wimmer, H. (2003). *Salzburger Lese-Screening für die Klassenstufen 1-4*. Göttingen: Hogrefe.

Mazzocco, M.M.M. & Myers, G.F. (2003). Complexities in identifying and defining mathematics learning disability in the primary school-age years. *Annals of Dyslexia*, *53*, 218–253.

McGee, C., Ward, R., Gibbons, J. & Harlow, A. (2003). *Transition to Secondary School: A literatur Review*. University of Waikato.

McMaster, K. & Fuchs, D. (2002). Effects of Cooperative Learning on the Academic Achievement of Students with Learning Disabilities: An Update of Tateyama-Sniezek's Review. *Learning Disabilities Research & Practice*, *17* (2), 107–117.

Methe, S., Kilgus, S.P., Neiman, C. & Riley-Tillman, T.C. (2012). Meta-Analysis of

Interventions for Basic Mathematics Computation in Single-case Research. *Journal of Behavioural Education*, *21*, 230–253.

Meyer, H. (2002). Unterrichtsmethoden. In H. Kiper, H. Meyer & W. Topsch (Hrsg.), *Einführung in die Schulpädagogik* (S. 109–121). Berlin: Cornelsen.

Meyer, H. (2010). *Leitfaden Unterrichtsvorbereitung*. Berlin: Cornelsen.

Meyer, H. & Klapper, A. (2006). Unterrichtsstandards für ein kompetenzorientiertes Lernen und Lehren. *Auf den Anfang kommt es an: Kompetenzen entwickeln - Kompetenzen stärken* (S. 89–108). Wiesbaden: VS Verlag für Sozialwissenschaften.

Meyer, S. (1993). *Was sagst du zur Rechenschwäche, Sokrates?* Luzern: Schweizerische Zentralstelle für Heilpädagogik SZH.

Mietzel, G., Willenberg, H., Poerschke, J. & Peek, R. (2001). *Hamburger Schulleistungstest für vierte und fünfte Klassen. HST 4/5*. Göttingen: Hogrefe.

Mihalic, S. (2004). The Importance of Implementation Fidelity. *Emotional and Behavioral Disorder in Youth*, 83–105.

Mittag, W. & Bieg, S. (2010). Die Bedeutung und Funktion pädagogischer Interventionsforschung und deren grundlegende Qualitätskriterien. In T. Hascher & B. Schmitz (Hrsg.), *Pädagogische Interventionsforschung* (S. 31–47). Weinheim: Juventa.

Mittelberg, A. (2004). *Rechenschwächen in der Hauptschule - Eine Studie zu den Rechenleistungen in den Klassen 7 und 8. Unveröffentlichte Dissertation*. Hannover: Universität Hannover.

Mitzlaff, H. & Wiederhold, K. (1989). Gibt es überhaupt ein Übergangsproblem? Erste Ergebnisse aus einem Forschungsprojekt. *Übergänge nach der Grundschule* (S. 12–41). Hemsbach: Beltz.

Molko, N., Cachia, A., Riviere, D., Mangin, J.-F., Bruandet, M., Le Bihan, D. et al. (2003). Functional and Structural Alterations of the Intraparietal Sulcus in a Developmental Dyscalculia of Genetic Origin. *Neuron*, *40*, 847–858.

Moll, K. & Landerl, K. (2011). Assoziationen und Dissoziationen von Störungen des Lesens und Rechtschreibens. *Spektrum Patholinguistik*, *4*, 47–74.

Moog, W. & Schulz, A. (2005). *Zahlen begreifen. Diagnose und Förderung bei Kindern mit Rechenschwäche. Mit Test und Trainigsverfahren*. Weinheim und Basel: Beltz.

Morgan, P.L., Farkas, G. & Maczuga, S. (2014). Which Instructional Practices Most Help First-Grade Students With and Without Mathematics Difficulties? *Educational Evaluation and Policy Analysis*, 1–22.

Moser Opitz, E. (2005). Lernschwierigkeiten Mathematik in Klasse 5 bis 8. Eine empirische Untersuchung zu fehlenden mathematischen Basiskompetenzen. *Vierteljahresschrift für Heilpädagogik und ihre Nachbargebiete*, *74*, 113–128.

Moser Opitz, E. (2007a). *Rechenschwäche/Dyskalkulie*. Bern: Haupt.

Moser Opitz, E. (2007b). Lernbereich Mathematik. *Didaktik des Unterrichts im Förderschwerpunkt Lernen* (S. 253–279). Stuttgart: Kohlhammer.

Moser Opitz, E., Reusser, L., Moeri Müller, M., Anliker, B., Wittich, C., Freesemann, O. et al. (2010). *Basisdiagnostik Mathematik für die Klassen 4–8 Basis-Math 4-8*. Göttingen: Hogrefe.

Mugerauer, R. (2012). *Kompetenzen als Bildung?* Marburg: Tectum Verlag.

Mulcahy, C.A., Krezmien, M.P. & Travers, J. (2012). Improving Mathematics Performance

Among Secondary Students With EBD: A Methodological Review. *Remedial and Special Education*, *37* (2), 113–128.

Müller-Ruckwitt, A. (2008). *„Kompetenz" – Bildungstheoretische Untersuchungen zu einem aktuellen Begriff*. Würzburg: ERGON Verlag.

Müller, K. & Ehmke, T. (2013). Soziale Herkunft als Bedingung der Kompetenzentwicklung. *PISA 2012. Fortschritte und Herausforderungen in Deutschland* (S. 245–274). Münster: Waxmann.

Müller, R. (2003). *Diagnostischer Rechtschreibtest für 1. Klassen. DRT 1*. Göttingen: Hogrefe.

Murphy, E., Grey, I.M. & Honan, R. (2005). Co-operative learning for students with difficultiesin learning: a description of models and guidelines for implementation. *British Journal of Special Education*, *32* (3), 157–164.

Myschker, N. (2005). *Verhaltensstörungen bei Kindern und Jugendlichen*. Stuttgart: Kohlhammer.

National Center on Response to Intervention. (2010). *Essential Components of RTI – A Closer Look at Response to Intervention*. US Office of Special Education Programs.

National Council of Teachers of Mathematics. (2000). *Principles and Standards for School Mathematics*. Portland: Graphic Arts Center.

Neukäter, D. & David, D. (1996). Rechenleistungen von Schülern mit und ohne Verhaltensstörungen. *Sonderpädagogik*, *26* (3), 154–159.

Niedersächsische Landesschulbehörde. (2016). Stundentafel Integrierte Gesamtschule, Jahrgang 5–10. Verfügbar unter: https://www.landesschulbehoerde-niedersachsen.de/themen/schueler/klassenbildung-stundentafeln/stundentafel-integrierte-gesamtschule-jahrgang-5-10

Niedersächsisches Kultusministerium. (2005). *Orientierungspläne für Bildung und Erziehung im Elementarbereich niedersächsischer Tageseinrichtungen*. Langenhagen: gutenberg beuys feindruckerei GmbH. Verfügbar unter: file:///Users/carolinakaeter/Downloads/2014.11.13_Orientierungsplan_web.pdf

Niklas, F. & Schneider, W. (2012). Einfluss von „Home Numeracy Environment" auf die mathematische Kompetenzentwicklung vom Vorschulalter bis Ende des 1. Schuljahres. *Zeitschrift für Familienforschung*, *2*, 134–147.

Nußbeck, S. (2007). Evidenzbasierte Praxis – ein Konzept für sonderpädagogisches Handeln? *Sonderpädagogik*, *37* (2/3), 146–155.

O'Connor, R.E. & Jenkins, J.R. (2013). Cooperative Learning for Students with Learning Disabilities. In H.L. Swanson, K. Harris & S. Graham (Hrsg.), *Handbook of Learning Disabilities* (S. 507–525). New York: Guilford Press.

O'Melia, M.C. & Rosenberg, M.S. (1994). Effects of Cooperative Homework Teams on the Acquisition of Mathematics Skills by Secondary Students with Mild Disabilities. *Exceptional Children*, *60* (6), 538–548.

OECD. (2013). *PISA 2012 Assessment and Analytical Framework. Mathematics, Reading, Science, Problem Solving and Financial Literacy* (OECD Progr.). OECD Publishing.

Opdenakker, M.-C. & Van Damme, J. (2006). Differences between Secondary Schools: A Study about school context, group composition, school pratice, and school effects with special attention to public and Catholic schools and types of schools. *School Effectiveness and School Improvement*, *17* (1), 87–117.

Opdenakker, M.-C. & Van Damme, J. (2007). Do school context, student composition and

school leadership affect school practice and outcomes in secondary education? *British Educational Research Journal*, *33* (2), 179–206.

Van Ophuysen, S. & Wendt, H. (2009). Zur Veränderung der Mathematikleistung von Klasse 4 bis 6. Welchen Einfluss haben Kompositions- und Unterrichtsmerkmale? *Zeitschrift für Erziehungswissenschaft, Sonderheft Bildungsentscheidungen*, *12*, 302–327.

Ortner, A. & Ortner, R. (2000). *Verhaltens- und Lernschwierigkeiten*. Weinheim: Beltz.

Oswald, E., Valkanover, S. & Conzelmann, A. (2013). Die Implementation einer Intervention zur Veränderung des Selbstkonzepts im Sportunterricht. *Schweizerische Zeitschrift für Bildungswissenschaften*, *35* (2), 255–273.

Owen, R.L. & Fuchs, L. (2002). Mathematical Problem-Solving Strategy Instruction for Third-Grade Students with Learning Disabilities. *Remedial and Special Education*, *23* (5), 268–278.

Pant, H.A., Böhme, K. & Köller, O. (2012). Das Kompetenzkonzept der Bildungsstandards und die Entwicklung von Kompetenzstufenmodellen. In P. Stanat, H.A. Pant, K. Böhme & D. Richter (Hrsg.), *Kompetenzen von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe in den Fächern Deutsch und Mathematik. Ergebnisse des IQB-Ländervergleichs* (S. 49–55). Münster: Waxmann.

Peard, R. (2010). Dyscalculia: What is its prevalence? Research evidence from case studies. *Procedia Social and Behavioral Sciences*, *8*, 106–113.

Pekrun, R., Vom Hofe, R., Blum, W., Götz, S., Wartha, S., Frenzel, A. et al. (2006). Projekt zur Analyse der Leistungsentwicklung in Mathematik (PALMA). Entwicklungsverläufe, Schülervoraussetzungen und Kontextbedingungen von Mathematikleistung in der Sekundarstufe I. In M. Prenzel & L. Allolio-Näcke (Hrsg.), *Untersuchungen zur Bildungsqualität von Schule: Abschlussbericht des DFG-Schwerpunktprogramms* (S. 21–53). Münster: Waxmann.

Petermann, F. (2015a). Alltagsintegrierte Förderung oder Förderprogramme im Vergleich? *Frühe Bildung*, *4* (3), 161–164.

Petermann, F. (2015b). Implementationsforschung: Grundbegriffe und Konzepte. *Psychologische Rundschau*, *65* (3), 122–128.

Petermann, F. & Petermann, U. (2010). *Hamburg-Wechsler-Intelligenztest für Kinder - IV. HAWIK-IV*. Göttingen: Hogrefe.

Petermann, U. & Petermann, F. (2013). *Lehrereinschätzliste für Sozial- und Lernverhalten*. Göttingen: Hogrefe.

Petermann, U. & Petermann, F. (2014). *Schülereinschätzliste für Sozial- und Lernverhalten*. Göttingen: Hogrefe.

Poplin, M., Rivera, J., Durish, D., Hoff, L., Kawell, S., Pawlak, P. et al. (2011). She's Strict for a Good Reason. Highly Effective Teachers in Low-Performing Urban Schools. *Kappan*, *92* (5).

Potter, N. (2001). *Looking forward to teaching: A team approach to teaching in Years 9 and 10*. Wellington: New Zealand Council for Educational Research.

Prediger, S., Selter, C., Hußmann, S. & Nührenbörger, M. (2014). *Mathe sicher können. Band 2: Brüche, Prozente und Dezimalzahlen*. Berlin: Cornelsen.

Prenzel, M., Sälzer, C., Klieme, E. & Köller, O. (2012). *PISA 2012. Fortschritte und Herausforderungen in Deutschland*. Münster: Waxmann.

Preuss-Lausitz, U. (2016). Throughput instead of input: Herausforderungen beim Wegfall der

Feststellungsdiagnostik in den Bereichen Lernen, Emotionale und soziale Entwicklung und Sprache. *Zeitschrift für Heilpädagogik*, *67*, 204–214.

Przychodzin, A.M., Marchand-Martella, N.E., Martella, R.C. & Azim, D. (2004). Direct Instruction: Mathematics Programs: An Overview and Research Summary. *Journal of Direct Instruction*, *4* (1), 53–84.

Racherbäumer, K. (2014). Individualisiertes Lehren im Übergang?! Ergebnisse einer Lehrer/innenbefragung zur Unterrichtsgestaltung nach dem Übergang in die Primar- bzw. Sekundarstufe. In A. Liegmann, I. Mammes & K. Racherbäumer (Hrsg.), *Facetten von Übergängen im Bildungssystem* (S. 65–78). Münster: Waxmann.

Racherbäumer, K. & Kohnen, M. (2014). Schulstufenübergreifende Kooperationen von Lehrkräften der Primar- und Sekundarstufe zur Gestaltung anschlussfähiger Bildungsprozesse am Beispiel Naturwissenschaften und Mathematik. In A. Liegmann, I. Mammes & K. Racherbäumer (Hrsg.), *Facetten von Übergängen im Bildungssystem* (S. 95–109). Münster: Waxmann.

Ramaa, S. & Gowramma, I.P. (2002). A systematic procedure for identifying and classifiying children with dyscalculia among primary school children in India. *Dyslexia*, *8*, 67–85.

Rasch, B., Friese, M., Hofmann, W. & Naumann, E. (2006a). *Quantitative Methoden 2*. Heidelberg: Springer Medizin.

Rasch, B., Friese, M., Hofmann, W. & Naumann, E. (2006b). *Quantitative Methoden 1*. Heidelberg: Springer Medizin.

Rauer, W. & Schuck, K.-D. (2003). *Fragebogen zur Erfassung emotionaler und sozialer Schulerfahrungen von Grundschulkindern dritter und vierter Klassen. FEESS 3-4*. Göttingen: Hogrefe.

Reichenbach, R. (2010). Der Matthäus-Effekt und der andere Sinn von Schule. *Heterogenität bejahen: Bildungsgerechtigkeit als Auftrag und Herausforderung für evangelische Schulen*. (S. 33–48). Münster: Waxmann.

Reinck, C. (2016). *Lernförderung im Mathematikunterricht. Eine quantitativ-empirische Erhebung zur Untersuchung der Wirksamkeit eines Advance Organizers für heterogene Lerngruppen im Mathematikunterricht der 3. Jahrgangsstufe. Dissertation*. Carl von Ossietzky Universität Oldenburg.

Reinmann-Rothmeier, G. & Mandl, H. (1998). Wenn kreative Ansätze versanden: Implementation als verkannte Aufgabe. *Unterrichtswissenschaft*, *26* (4), 292–311.

Reiss, K. (2004). Bildungsstandards und die Rolle der Fachdidaktik am Beispiel der Mathematik. *Zeitschrift für Pädagogik*, *50*, 635–649.

Reiss, K. (2009). Mathematische Kompetenz zwischen Grundschule und Sekundarstufe: Zusammenfassung und Forschungsdesiderata. In A. Heine & M. Grüßing (Hrsg.), *Mathematiklernen vom Kindergarten bis zum Studium. Kontinuität und Kohärenz als Herausforderung für den Mathematikunterricht*. (S. 117–124). Münster: Waxmann.

Reiss, K., Heinze, A. & Pekrun, R. (2007). Mathematische Kompetenz und ihre Entwicklung in der Grundschule. *Zeitschrift für Erziehungswissenchaft, Sonderheft Kompetenzdiagnostik*, *10* (8), 107–127.

Reiss, K., Hellmich, F. & Joachim, T. (2002). Individuelle und schulische Bedingungsfaktoren für Argumentationen und Beweise im Mathematikunterricht. In M. Prenzel & J. Doll (Hrsg.), *Bildungsqualität von Schule: Schulische und außerschulische Bedingungen mathematischer, naturwissenschaftlicher und überfachlicher Kompetenzen* (S. 51–64). Weinheim und Basel: Beltz.

Reiss, K., Roppelt, A., Haag, N., Pant, H.A. & Köller, O. (2012). Kompetenzstufenmodelle im Fach Mathematik. In P. Stanat, H.A. Pant, K. Böhme & D. Richter (Hrsg.), *Kompetenzen von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe in den Fächern Deutsch und Mathematik. Ergebnisse des IQB-Ländervergleichs* (S. 72–84). Münster: Waxmann.

Reiss, K. & Winkelmann, H. (2008). Step by step. Ein Kompetenzstufenmodell für das Fach Mathematik. *Grundschule*, *40* (10), 34–37.

Reiss, K. & Winkelmann, H. (2009). Kompetenzstufenmodelle für das Fach Mathematik im Primarbereich. In D. Granzer, O. Köller, A. Bremerich-Vos, M. van den Heuvel-Panhuizen, K. Reiss & G. Walther (Hrsg.), *Bildungsstandards Deutsch und Mathematik. Leistungsmessung in der Grundschule* (S. 120–141). Weinheim und Basel: Beltz.

Rensing, J. (2014). *Untersuchung der Änderungssensibilität eines selbst entwickelten curriculumbasierten Messverfahrens zu mathematischen Basiskompetenzen in der 4. Klasse. Unveröffentlichte Masterarbeit*. Oldenburg: Carl von Ossietzky Universität Oldenburg.

Rensing, J., Käter, C., Käter, T. & Hillenbrand, C. (2017). Konstruktion und Überprüfung eines curiculumbasierten Testverfahrens im Fach Mathematik für die vierte Klasse. *Empirische Sonderpädagogik*, *4* (8), 346–366.

Resnick, L. (1983). A developmental theory of number understanding. *The development of mathematical thinking* (S. 109–151). New York: Academic Press.

Rey, G.D. & Nieding, G. (2010). Arbeitsgedächtnis und Cognitive Load - Formen, Messmethoden und Anwendungen. In H.-P. Trolldenier, W. Lenhard & P. Marx (Hrsg.), *Brennpunkte der Gedächtnisforschung* (S. 67–85). Göttingen: Hogrefe.

Riccomini, P.J. & Witzel, B.S. (2010). *Response to Intervention in Math*. Thousand Oaks: Corwin.

Ricken, G. (2014). Rechenschwierigkeiten. In F. Wember, R. Stein & U. Heimlich (Hrsg.), *Handlexikon Lernschwierigkeiten und Verhaltensstörungen* (S. 69–72). Stuttgart: Kohlhammer.

Ricken, G., Fritz, A. & Balzer, L. (2013). *MARKO-D Mathematik- und Rechenkonzepte im Vorschulalter - Diagnose*. Göttingen: Hogrefe.

Rittle-Johnson, B. & Siegler, R.S. (1998). The relation between conceptual and procedural knowledge in learning mathematics: A review. In C. Donlan (Hrsg.), *The development of mathematical skills. Studies in developmental psychology* (S. 75–110). Hove: Psychology Press/Taylor & Francis.

Rohrbeck, C.A., Ginsburg-Block, M.D., Fantuzzo, J.W. & Miller, T.R. (2003). Peer-Assisted Learning Interventions With Elementary School Students: A Meta-Analytic Review. *Journal of Educational Psychology*, *95* (2), 240–257.

Rolus-Borgward, S. (1997). Schulleistungen von Kindern und Jugendlichen mit Verhaltensstörungen – eine Literaturanalyse. *Sonderpädagogik*, *27* (4), 194–201.

Rost, D. (2013). *Interpretation und Bewertung pädagogisch-psychologischer Studien*. Bad Heilbrunn: Klinkhardt.

Sälzer, C., Reiss, K., Schiepe-Tiska, A., Prenzel, M. & Heinze, A. (2013). Zwischen Grundlagenwissen und Anwendungsbezug: Mathematische Kompetenz im internationalen Vergleich. In M. Prenzel, C. Sälzer, E. Klieme & O. Köller (Hrsg.), *PISA 2012. Fortschritte und Herausforderungen in Deutschland* (S. 47–97). Münster: Waxmann.

Schad, G. (2015). Evidenzbasierte Erziehung? *Zeitschrift für Heilpädagogik*, *66* (7), 335–344.

Schaupp, H., Holzer, N. & Lenart, F. (2010). *Eggenberger Rechentest 4+. ERT 4+*. Göttingen: Hogrefe.

Schell, A. (2011). *Die Förderung emotionaler und sozialer Kompetenzen bei Kindern im Vorschulalter. „Lubo aus dem All" - Entwicklung, Implementierung und Evaluation eines Trainingsprogramms zur Prävention von Gefühls- und Verhaltensstörungen*. Bad Heilbrunn: Klinkhardt.

Schels, L., Janz, F. & Hintermair, M. (2015). Zur Rolle stressverstärkender Gedanken für das Stresserleben von Sonderpädagoginnen und Sonderpädagogen. *Zeitschrift für Heilpädagogik*, *66*, 552–563.

Scherer, P. (2002). *Produktives Lernen für Kinder mit Lernschwächen: Fördern durch Fordern*. Hamburg: Persen.

Schipper, W. (2009). Schriftliches Rechnen als neue Chance für rechenschwache Kinder. In A. Fritz, G. Ricken & S. Schmidt (Hrsg.), *Handbuch Rechenschwäche* (S. 118–134). Weinheim: Beltz.

Schipper, W. (2010). Thesen und Empfehlungen zum schulischen und außerschulischen Umgang mit Rechenstörungen. *Rechenschwäche, Rechenstörung, Dyskalkulie* (S. 103–121). Graz: Leykam.

Schmidt-Atzert, L. & Amelang, M. (2012). *Psychologische Diagnostik und Intervention*. Berlin: Springer.

Schmidt, F. (2014). *Eine Pilotstudie zur Förderung mathematischer Basiskompetenzen in der fünften Klasse. Unveröffentlichte Masterarbeit*. Oldenburg: Carl von Ossietzky Universität Oldenburg.

Schmitt, R. (2001). *Grundlegende Bildung in und für Europa*. Frankfurt a. M.: Grunschulverband - Arbeitskreis Grunschule e. V.

Schmitz, G.S. & Schwarzer, R. (2002). Individuelle und kollektive Selbstwirksamkeitserwartung von Lehrern. *Zeitschrift für Pädagogik*, *44*, 192–214.

Schneider, W. & Hasselhorn, M. (2016). *Förderprogramme für Vor- und Grundschule*. Göttingen: Hogrefe.

Schneider, W., Küspert, P. & Krajewski, K. (2013). *Die Entwicklung mathematischer Kompetenzen*.

Schnell, R., Hill, P. & Esser, E. (2011). *Methoden der empirischen Sozialforschung*. München: Oldenbourg.

Schröder, U. (2005). *Lernbehindertenpädagogik*. Stuttgart: Kohlhammer.

Schröder, U. & Wittrock, M. (2002). *Lernbeeinträchtigung und Verhaltensstörung. Konvergenzen in Theorie und Praxis*. Stuttgart: Kohlhammer.

Schuchardt, K., Gebhardt, M. & Mähler, C. (2010). Working Memory functions in children with different degrees of intellectual disability. *Journal of Intellectual Disability Research*, *54* (4), 346–353.

Schuchardt, K. & Mähler, C. (2010). Unterscheiden sich Subgruppen rechengestörter Kinder in ihrer Arbeitsgedächtniskapazität, im basalen und arithmetischen Faktenwissen und in den numerischen Basiskompetenzen? *Zeitschrift fur Entwicklungspsychologie und Padagogische Psychologie*, *42* (2), 217–225.

Schulte-Körne, G. (2010). Diagnostik und Therapie der Lese-Reschtschreibstörung. *Deutsches Ärzteblatt*, *107* (41), 718–727.

Schwenk, C. & Schneider, W. (2003). Einflussfaktoren für den Zusammenhang von Rechen- und Schriftsprachleistungen im frühen Grundschulalter. *Kindheit und Entwicklung*, *12* (4), 212–221.

Scriven, M. (1967). *The Methodology of Evaluation*. Washington D. C.: American Educational Research Association.

Sedlmeier, P. & Renkewitz, F. (2013). *Forschungsmethoden und Statistik*. Hallbergmoos: Pearson.

Selter, C., Prediger, S., Nührenbörger, C. & Hußmann, S. (2014). *Mathe sicher können. Band 1: Natürliche Zahlen. Handreichung für ein Diagnose- und Förderkonzept zur Sicherung mathematischer Basiskompetenzen*. Berlin: Cornelsen.

Selter, C., Walter, D., Walther, G. & Wendt, H. (2016). Mathematische Kompetenzen im internationalen Vergleich: Testkonzeption und Ergebnisse. In H. Wendt, W. Bos, C. Selter, O. Köller, K. Schwippert & D. Kasper (Hrsg.), *TIMSS 2015. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 79–136). Münster: Waxmann.

Selter, C., Walther, G., Wessels, J. & Wendt, H. (2012). Mathematische Kompetenzen im internationalen Vergleich: Testkonzeption und Ergebnisse. *TIMSS 2011. Mathematische und naturwissenschaftliche Kompetenz von Grundschulkindern in Deutschland im internationalen Vergleich*. (S. 69–122). Münster: Waxmann.

Senate and House of Representatives of the USA. (2002). *No Child Left Behind Act*. United States of America. Verfügbar unter: http://www2.ed.gov/policy/elsec/leg/esea02/107-110.pdf

Senate and House of Representatives of the USA. (2004). *Individuals with Disabilities Education Improvement Act*. United States of America.

Shapiro, E.S., Edwards, L. & Zigmond, N. (2005). Progress Monitoring of Mathematics Among Students with Learning Disabilities. *Assessment for Effective Intervention*, *30* (2), 15–32.

Siegler, R.S. (1987). The Perils of Averaging Data over Strategies: An Example From Children's Addition. *Journal of Experimental Psychology*, *116* (3), 250–264.

Siegler, R.S., Duncan, G.J., Davis-Kean, P.E., Duckworth, K., Claessens, A., Engel, M. et al. (2012). Early Predictors of High School Mathematics Achievement. *Psychological Science*, *23* (7), 691–697.

Simon, H. (2005). *Dyskalkulie - Kindern mit Rechenschwäche wirksam helfen*. Stuttgart: Klett-Cotta.

Simon, H. & Grünke, M. (2010). *Förderung bei Rechenschwäche*. Stuttgart: Kohlhammer.

Sinner, D. & Kuhl, J. (2010). Förderung mathematischer Basiskompetenzen in der Grundstufe der Schule für Lernhilfe. *Zeitschrift fur Entwicklungspsychologie und Pädagogische Psychologie*, *42* (4), 241–251.

Sirsch, U. (2000). *Probleme beim Schulwechsel*. Münster: Waxmann.

Slavin, R.E., Lake, C. & Groff, C. (2009). Effective Programs in Middle and High School Mathematics: A Best-Evidence Synthesis. *Review of Educational Research*, *79* (2), 839–911.

Smith, T.J., McKenna, C.M. & Hines, E. (2014). Association of group learning with mathematics achievement and mathematics attitude among eith-grade students in the U.S. *Learning Enovironments Research*, *17*, 229–241.

Snijders, J.T., Tellegen, P.J. & Laros, J.A. (2005). *Non-verbaler Intelligenztest (SON-R 5 1/2-*

7). Göttingen: Hogrefe.

South Dakota Department of Education. (2012). *Response to Intervention Implementation Guide: The South Dakota Model*. Verfügbar unter: https://doe.sd.gov/oess/documents/sped_RtI_ImplementationGuide.pdf

Souvignier, E. (2007). Kooperatives Lernen. In U. Heimlich & F. Wember (Hrsg.), *Didaktik des Unterrichts im Förderschwerpunkt Lernen* (S. 138–148). Stuttgart: Kohlhammer.

Souvignier, E. (2008). Lernbehinderung. In W. Schneider & M. Hasselhorn (Hrsg.), *Handbuch der Pädagogischen Psychologie* (S. 663–671). Göttingen: Hogrefe.

Spicuzza, R., Ysseldyke, J., Lemkuil, A., Kosciolek, S., Boys, C. & Teelucksingh, E. (2001). Effects of curriculum-based monitoring on classroom instruction and math achievement. *Journal of School Psychology*, *39*, 521–542.

Spörer, N. (2009). Festigung mathematischer Basiskompetenzen durch Peer-gestütztes Lernen: Ergebnisse einer Trainingsstudie in der Grundschule. *Empirische Pädagogik*, *23* (1), 75–94.

Stecken, T. & Stein, M. (2015). *Test zum mathematischen Diagrammverständnis. TeMaDi*. Göttingen: Hogrefe.

Stecker, P.M. & Fuchs, L. (2000). Effecting Superior Achievement Using Curriculum-Based Measurement: The Importance of Individual Progress Monitoring. *Learning Disabilities Research & Practice*, *15* (3), 128–134.

Stecker, P.M., Fuchs, L. & Fuchs, D. (2005). Using Curriculum-Based Measurement to Improve Student Achievement: Review of Research. *Psychology in the Schools*, *42* (8).

Stein, R. (2014). Didaktik des Unterrichts im Förderschwerpunkt emotionale und soziale Entwicklung. In F. Wember, R. Stein & U. Heimlich (Hrsg.), *Handlexikon Lernschwierigkeiten und Verhaltensstörungen* (S. 110–114). Stuttgart: Kohlhammer.

Stenhoff, D.M. & Linguaris-Kraft, B. (2007). A Review of the Effects of Peer Tutoring on Students with Mild Disabilities in Secondary Settings. *Exceptional Children*, *74* (1), 8–30.

Stern, E. (1997). Erwerb mathematischer Kompetenzen: Ergebnisse aus dem SCHOLASTIK-Projekt. In F.E. Weinert & A. Helmke (Hrsg.), *Entwicklungen im Grundschulalter* (S. 157–170). Weinheim und Basel: Beltz.

Stern, E. (2003). Früh übt sich - Neuere Ergebnisse aus der LOGIK-Studie zum Lösen mathematischer Textaufgaben. In A. Fritz, G. Ricken & S. Schmidt (Hrsg.), *Handbuch Rechenschwäche* (S. 116–130). Weinheim und Basel: Beltz.

Stern, E. (2008). Verpasste Chancen? Was wir aus der LOGIK-Studie über den Mathematikunterricht lernen können. In W. Schneider (Hrsg.), *Entwicklung von der Kindheit bis zum Erwachsenenalter. Befunde der Münchener Längsschnittstudie LOGIK*. (S. 187–202). Weinheim und Basel: Beltz.

Stern, E. (2013). Kognitive Entwicklungspsychologie des mathematischen Denkens. In M. van Aster & J.H. Lorenz (Hrsg.), *Rechenstörungen bei Kindern. Neurowissenschaft, Psychologie und Pädagogik* (S. 141–154). Göttingen: Vandenhoeck & Ruprecht.

Stevens, R.J. & Slavin, R.E. (1995). The Cooperative Elementary School: Effects on Students' Achievement, Attitudes, and Social Relations. *American Educational Research Journal*, *32* (2), 321–351.

Stiftung Rechnen. (2013). *Bürgerkompetenz Rechnen. Ergebnisbericht*. Verfügbar unter: http://zeus.zeit.de/wissen/2013-05/13-05-29_Ergebnisbericht_Studie_Rechnen.pdf

Strahm, P. (2008). *Qualität durch systematisches Feedback. Grundlagen, Einblicke, Werkzeug*.

Bern: Schulverlag.

Strathmann, A.M. & Klauer, K.J. (2010). Lernverlaufsdiagnostik: Ein Ansatz zur längerfristigen Lernfortschrittsmessung. *Zeitschrift für Entwicklungspsychologie und pädagogische Psychologie*, *43*, 111–122.

Strathmann, A.M. & Klauer, K.J. (2012). *Lernverlaufsdiagnostik-Mathematik für zweite bis vierte Klassen (LVD-M 2-4)*. Göttingen: Hogrefe.

Swanson, H.L., Cooney, J.B. & McNamara, J.K. (2008). Lernstörungen und Gedächtnis. *Lernstörungen verstehen* (S. 39–90). Berlin: Spektrum Akademischer Verlag.

Swanson, H.L., Harris, K. & Graham, S. (2006). Handbook of Learning Disabilities. New York: The Guilford Press.

Swanson, H.L. & Sachse-Lee, C. (2000). A Meta-Analysis of Single-Subject-Design Intervention Research for Students with LD. *Journal of Learning Disabilities*, *33* (2), 114–136.

Tennessee Department of Education. (2013). *Response to Instruction & Intervention Framework*.

Tenorth, H.-E. & Tippelt, R. (2007). Implementation. *BELTZ Lexikon Pädagogik*. Beltz Verlag.

Thal, J. & Ebert, U. (2001). *Methodenvielfalt im Unterricht*. Kriftel: Hermann Luchterhand.

Topping, K.J. (2005). Trends in Peer Learning. *Educational Pychology*, *25* (6), 631–645.

Tournaki, N. (2003). The Differential Effects of Teaching Addition Through Strategy Instruction Versus Drill and Practice to Students With and Without Leaning Disabilities. *Journal of Learning Disabilities*, *36* (5), 449–458.

Tran, L., Sanchez, T., Arellano, B. & Swanson, H.L. (2011). A Meta-Analysis of the RTI Literature for Children at Risk for Reading Disabilities. *Journal of Learning Disabilities*, *44* (3), 283–295.

Tröster, H. (2008). *Früherkennung im Kindes- und Jugendalter*. Göttingen: Hogrefe.

Tsuei, M. (2014). Mathematics Synchronous Peer Tutoring System for Students with Learning Disabilities. *Educational Technology & Society*, *17* (1), 115–127.

Ufer, S. (2009). Der Übergang von der Primarstufe in die Sekundarstufe. In A. Heinze & M. Grüßing (Hrsg.), *Mathematiklernen vom Kindergarten bis zum Studium. Kontinuität und Kohärenz als Herausforderung für den Mathematikunterricht* (S. 88–103). Münster: Waxmann.

Ufer, S., Reiss, K. & Heinze, A. (2009). BIGMATH - Ergebnisse zur Entwicklung mathematischer Kompetenz in der Primarstufe. In A. Heinze & M. Grüßing (Hrsg.), *Mathematiklernen vom Kindergarten bis zum Studium. Kontinuität und Kohärenz als Herausforderung für den Mathematikunterricht* (S. 62–85). Münster: Waxmann.

Vaughn, S. & Fuchs, L. (2003). Redefining Learning Disabilities as Inadequate Response to Instruction: The Promise and Potential Problems. *Learning Disabilities Research & Practice*, *18* (3), 137–146.

Vereinte Nationen. (2006). *Übereinkommen über die Rechte von Menschen mit Behinderungen*. Bundesgesetzblatt 2008 II. Verfügbar unter: http://www.institut-fuer-menschenrechte.de/fileadmin/user_upload/PDF-Dateien/Pakte_Konventionen/CRPD_behindertenrechtskonvention/crpd_b_de.pdf

Vollrath, H.-J. & Roth, J. (2012). *Grundlagen des Mathematikunterrichts in der Sekundarstufe*. Heidelberg: Spektrum Akademischer Verlag.

Voß, S. (2013). *Curriculumbasierte Messverfahren im mathematischen Erstunterricht. Zur Güte und Anwendbarkeit einer Adaption US-amerikanischer Verfahren im deutschen Schulsystem*. Saarbrücken: SVH.

Voß, S., Blumenthal, Y., Mahlau, K., Marten, K., Diehl, K., Sikora, S. et al. (2016). *Der Response-to-Intervention-Ansatz in der Praxis. Evaluationsergebnisse zum Rügener Inklusionsmodell*. Münster: Waxmann.

Voß, S. & Hartke, B. (2014). Curriculumbasierte Messverfahren (CBM) als Methode der formativen Leistungsdiagnostik im RTI-Ansatz. *Lernverlaufsdiagnostik (Bd. 12)* (S. 83–99). Göttingen: Hogrefe.

Voß, S., Marten, K., Diehl, K., Mahlau, K., Sikora, S., Blumenthal, Y. et al. (2016). Evaluationsergebnisse des Projekts „Rügener Inklusionsmodell (RIM) - Präventive und Integrative Schule auf Rügen (PISaR)“ nach vier Jahren. *Zeitschrift für Heilpädagogik*, *66*, 133–149.

De Vries, C. (2013). *DIFMaB Diagnostisches Inventar zur Förderung Mathematischer Basiskompetenzen*. Dortmund: verlag modernes lernen.

Wahl, D. (2005). Die besondere Bedeutung einer früh („in advance“) im Lernprozess vermittelten Experten-Struktur („organizer“). *Lernumgebungen erfolgreich gestalten. Vom trägen Wissen zum kompetenten Handeln* (S. 139–154). Bad Heilbrunn: Klinkhardt.

Walter, J. (2002). Einer flog übers Kuckucksnest“ oder welche Interventionsformen erbringen im sonderpädagogischen Feld welche Effekte? Ergebnisse ausgewählter us-amerikanischer Meta- und Mega-Analysen. *Zeitschrift für Heilpädagogik*, *11*, 442–450.

Walter, J. (2008a). Curriculumbasiertes Messen (CBM) als lernprozessbegleitende Diagnostik: Erste deutschsprachige Ergebnisse zur Validität, Reliabilität und Veränderungssensibilität eines robusten Indikators zur Lernfortschrittsmessung beim Lesen. *Heilpädagogische Forschung*, *34* (2), 62–79.

Walter, J. (2008b). Adaptiver Unterricht erneut betrachtet: Über die Notwendigkeit systematischer formativer Evaluation von Lehr- und Lernprozessen und die daraus resultierende Diagnostik und Neudefinition von Lernstörungen nach dem RTI-Paradigma. *Zeitschrift für Heilpädagogik*, *59*, 202–215.

Walter, J. (2009). *Lernfortschrittsdiagnostik Lesen (LDL). Ein curriculumbasiertes Verfahren*. Göttingen: Hogrefe.

Walter, J. (2014a). Lernfortschrittsdiagnostik Lesen (LDL) und Verlaufsdiagnostik sinnerfassenden Lesens (VSL): Zwei Verfahren als Instrumente der formativ orientierten Lesediagnostik. In M. Hasselhorn, W. Schneider & U. Trautwein (Hrsg.), *Lernverlaufsdiagnostik (Bd. 12)* (S. 165–201). Göttingen: Hogrefe.

Walter, J. (2014b). *Verlaufsdiagnostik sinnerfassenden Lesens (VSL)*. Göttingen: Hogrefe.

Walter, S. (2016). Matheabenteuer - Ein pragmatischer Forschungsansatz zur Unterstützung der Inklusion im Erstrechenunterricht. *Zeitschrift für Heilpädagogik*, *67*, 180–191.

Walther, G., Geiser, H., Langeheine, R. & Lobemeier, K. (2003). Mathematische Kompetenzen am Ende der vierten Jahrgangsstufe. In W. Bos, E. Lankes, M. Prenzel, K. Schwippert, G. Walther & R. Valtin (Hrsg.), *Erste Ergebnisse aus IGLU. Schülerleistungen am Ende der vierten Jahrgangsstufe im internationalen Vergleich* (S. 189–226). Münster: Waxmann.

Warnke, A. (2005). Umschriebene Entwicklungsstörungen. *Kinder- und Jugendpsychiatrie. Eine praktische Einführung* (S. 142–155). Stuttgart: Thieme.

Watkins, C.L. (1996). Follow Through: Why Didn’t We? *Effective School Practices*, *15* (1), 57–66. Verfügbar unter: http://darkwing.uoregon.edu/~adiep/ft/watkins.htm

Weber, A. & Stefanek, J. (1998). Überblick über die Längsschnittstudie LOGIK. In F.E. Weinert (Hrsg.), *Entwicklung im Kindesalter* (S. 39–52). Weinheim und Basel: Beltz.

Weinert, F.E. (2001). Vergleichende Leistungsmessung in Schulen - eine umstrittene Selbstverständlichkeit. In F.E. Weinert (Hrsg.), *Leistungsmessungen in Schulen* (S. 17–31). Weinheim und Basel: Beltz Verlag.

Weiß, R. (2006). *Culture Fair Intelligence Test. Grundintelligenztest Skala 2 Revision*. Göttingen: Hogrefe.

Weissbach, B. (1986). Sekundarstufenschock in Gesamtschulen. Ursachen, Erscheinungsformen ... u. was Schule dagegen tun kann. *Westermanns pädagogische Beiträge*, *1*, 21–25.

Weißhaupt, S. & Peucker, S. (2009). Entwicklung arithmetischen Vorwissens. In A. Fritz, G. Ricken & S. Schmidt (Hrsg.), *Handbuch Rechenschwäche* (S. 52–76). Weinheim und Basel: Beltz Verlag.

Weltgesundheitsorganisation. (2013). *Internationale statistische Klassifikation der Krankheiten und verwandter Gesundheitsprobleme (International Statistical Classification of Diseases and Related Health Problems)*. Verfügbar unter: http://www.icd-code.de/

Wember, F. (2007). Direkter Unterricht. In U. Heimlich & F. Wember (Hrsg.), *Didaktik des Unterrichts im Förderschwerpunkt Lernen* (S. 163–176). Stuttgart: Kohlhammer.

Wendt, H., Bos, W., Selter, C. & Köller, O. (2012). TIMSS 2011: Wichtige Ergebnisse im Überblick. *TIMSS 2011. Mathematische und naturwissenschaftliche Kompetenz von Grundschulkindern in Deutschland im internationalen Vergleich*. (S. 13–26). Münster: Waxmann.

Wexler, J., Reed, D.K., Pyle, N., Mitchell, M. & Barton, E.E. (2015). A Synthesis of Peer-Mediated Academic Interventions for Secondary Struggling Learners. *Journal of Learning Disabilities*, *48* (5), 451–470.

Whitaker, T. (2009). *Was gute Lehrer anders machen: 14 Dinge, auf die es wirklich ankommt*. Weinheim: Beltz.

Whitehurst, G. (2002). Evidence-Based Education. Verfügbar unter: http://ies.ed.gov/director/pdf/2002_10.pdf

Wiechmann, J. (2016). Direkte Instruktion. In J. Wiechmann & S. Wildhirt (Hrsg.), *12 Unterrichtsmethoden* (S. 40–51). Weinheim: Beltz.

Wilbert, J. & Linnemann, M. (2011). Kriterien zur Analyse eines Tests zur Lernverlaufsdiagnostik. *Empirische Sonderpädagogik*, *3*, 225–242.

Winter, K. (2010). *Entwicklung von Item-Distraktoren mit diagnostischem Potential zur individuellen Defizit- und Fehleranalyse. Didaktische Überlegungen, empirische Untersuchungen und konzeptionelle Entwicklung für ein internetbasiertes Mathematik-Self-Assessment*. Westfälische Wilhelms-Universität Münster.

Wittmann, E.C. & Müller, G.N. (1990). *Handbuch produkiver Rechenübungen*. Stuttgart: Klett.

Wittmann, G. (2007a). Von Fehleranalysen zur Fehlerkultur. *Beiträge zum Mathematikunterricht*, 175–178.

Wittmann, G. (2007b). Fehleranalysen in der Bruchrechnung. *Beiträge zum Mathematikunterricht*, 173–174.

Wunram, F. (2014). *Überprüfung der Validität eines selbst entwickelten curriculumbasierten Messverfahrens zu ausgewählten mathematischen Kompetenzen in der 4. Klasse*.

Unveröffentlichte Masterarbeit. Oldenburg: Carl von Ossietzky Universität Oldenburg.

Xin, F. (1996a). *The Effects of Computer-Assisted Cooperative Learning in Mathematics in Integrated Classrooms for Students with and without Disabilities. Final Report. Study One*. New Jersey: Rowan College, Special Education Department. Verfügbar unter: http://files.eric.ed.gov/fulltext/ED412696.pdf

Xin, F. (1996b). *The Effects of Computer-Assisted Cooperative Learning in Mathematics in Integrated Classrooms for Students with and Without Disabilities. Final Report. Study two*. New Jersey: Rowan College, Special Education Department. Verfügbar unter: http://files.eric.ed.gov/fulltext/ED412696.pdf

Xin, Y.P., Jitendra, A.K. & Deatline-Buchman, A. (2005). Effects of Mathematical Word Problem–Solving Instruction on Middle School Students with Learning Problems. *The Journal of Special Education*, *3* (39), 181–192.

Zielinski, W. (1980). *Lernschwierigkeiten. Verursachungsbedingungen, Diagnose, Behandlungsansätze*. Stuttgart: Kohlhammer.

Zielinski, W. (1998). *Lernschwierigkeiten*. Stuttgart: Kohlhammer.

Zierer, K. (2014). Wie sich Effektstärken in der Erziehungswissenschaft interpretieren lassen. *JfAD*, *115–122*.

Zimmermann, K.R. (2014). *Rechenschwierigkeiten erkennen und bewältigen*. Weinheim und Basel: Beltz.

Zöfel, P. (2003). *Statistik für Psychologen im Klartext*. München: Pearson.

Zucker, S. (2004). *Scientifically Based Research: NCLB and Assessment. Policy Report*. San Antonio: Pearson Education. Verfügbar unter: http://images.pearsonassessments.com/images/tmrs/tmrs_rg/ScientificallyBasedResearch.pdf

13 Anhang

Anhang 1 Implementationsfragebogen der Interventionsleiterinnen und -leiter ... 328
Anhang 2 Ergebnisse des Implementationsfragebogen der Interventionsleiterinnen und -leiter auf Itemebene (deskriptiv) ... 333
Anhang 3 Implementationsfragebögen der Schülerinnen und Schüler ... 336
Anhang 4 Ergebnisse des Implementationsfragebogen der Schülerinnen und Schüler auf Itemebene (deskriptiv) ... 339
Anhang 5 Überprüfung auf Normalverteilung im Hypothesenblock I und II ... 341
Anhang 6 Überprüfung auf Normalverteilung im Hypothesenblock III ... 343
Anhang 7 Überprüfung auf Normalverteilung im Hypothesenblock IV ... 344

ANHANG 1 IMPLEMENTATIONSFRAGEBOGEN DER INTERVENTIONSLEITERINNEN UND -LEITER

Fragebogen zur Bewertung der Maßnahme zur Förderung der mathematischen Basiskompetenzen

Sehr geehrte/r Teilnehmer/in,

der nachfolgende Fragebogen soll Aufschluss über die Umsetzung der Maßnahme zur Förderung der mathematischen Basiskompetenzen geben. Das Ziel ist die Einschätzung der Realisierbarkeit der Maßnahme.

Bitte entscheiden Sie sich bei jeder Frage für eine Antwort. Wählen Sie bei Unsicherheiten die Antwort, die Ihrer Einschätzung am ehesten entspricht.

Ihre Daten werden anonym erhoben, vertraulich behandelt und ausschließlich zur Evaluation der Maßnahme verarbeitet. Anhand der veröffentlichten Ergebnisse sind keine Rückschlüsse auf Einzelpersonen möglich.

(Die im Folgenden verwendete Abkürzung SuS steht gleichermaßen für SchülerInnen und Schüler.)

		stimmt genau	stimmt ziemlich	stimmt kaum	stimmt gar nicht
	Ich habe die Olympischen Rechenspiele gerne durchgeführt				
AK_{ZI}	zu Beginn	☐	☐	☐	☐
AK_{ZI}	nach der Hälfte	☐	☐	☐	☐
AK_{ZI}	am Ende	☐	☐	☐	☐
AK_{ZI}	Die Intervention hat meine Erwartungen insgesamt erfüllt.	☐	☐	☐	☐
Ü	Ich würde die Olympischen Rechenspiele ein weiteres Mal durchführen.	☐	☐	☐	☐
Ü	Ich würde die Olympischen Rechenspiele weiterempfehlen.	☐	☐	☐	☐
AK_{ZI}	Die Olympischen Rechenspiele haben mir Anregungen für eine zukünftige Förderung der mathematischen Basiskompetenzen gegeben.	☐	☐	☐	☐
AN_{ZG}	Die Olympischen Rechenspiele sind für die Zielgruppe (SuS an Ober- oder Gesamtschulen) angemessen.	☐	☐	☐	☐
AN_{ZG}	Die Olympischen Rechenspiele motivieren SuS zum Rechnen.	☐	☐	☐	☐
AN_{ZG}	Die Olympischen Rechenspiele machen den SuS Freude beim Rechnen.	☐	☐	☐	☐
AN_{PI}	Die Inhalte der Olympischen Rechenspiele sind zur Förderung des Stellenwertsystems und der Grundrechenarten angemessen.	☐	☐	☐	☐
AN_{ZP}	Die Zeitplanung ist insgesamt stimmig.	☐	☐	☐	☐

AN_{ZP}	Die Zeitplanung in den Stunden ist zu kurz. Bei ‚Stimmt genau/stimmt ziemlich' – in welchen Stunden? _____ _____	☐	☐	☐	☐
AN_{ZP}	Die Zeitplanung in den Stunden ist zu lang. Bei „Stimmt genau/stimmt ziemlich' – in welchen Stunden? _____ _____	☐	☐	☐	☐
AN_{ZG}	Die Olympischen Rechenspiele fördern besonders die mathematischen Basiskompetenzen der leistungsstarken SuS.	☐	☐	☐	☐
AN_{ZG}	Die Olympischen Rechenspielen fördern besonders die mathematischen Basiskompetenzen der leistungsschwachen SuS.	☐	☐	☐	☐
AN_{ZG}	Die Olympischen Rechenspielen fördern besonders die mathematischen Basiskompetenzen von SuS mit diagnostiziertem Unterstützungsbedarf im Bereich Lernen.	☐	☐	☐	☐
AN_{ZP}	Die Stundenplanung ermöglicht es, einzelnen SuS gezielt Hilfe)stellung zu geben.	☐	☐	☐	☐
AN_{ZG}	Die Materialien sind für leistungsstarke SuS lernförderlich differenziert.	☐	☐	☐	☐
AN_{ZG}	Die Materialien sind für leistungsschwache SuS lernförderlich differenziert.	☐	☐	☐	☐
AN_{PME}	Die eingesetzten Methoden der Olympischen Rechenspiele sind zur Förderung des Stellenwertsystems und der Grundrechenarten angemessen. Bei "Stimmt kaum/gar nicht?" - warum nicht? _____ _____	☐	☐	☐	☐
AN_{AO}	Der Advance Organizer verdeutlicht den SuS das Thema der Stunde.	☐	☐	☐	☐
AN_{AO}	Der Advance Organizer gibt den SuS einen Rückblick und eine Vorschau auf die Unterrichtsthemen.	☐	☐	☐	☐
AN_{AB}	Die Arbeitsblätter helfen den SuS beim Lösen von Aufgaben.	☐	☐	☐	☐

AN_{DI}	Die Strategieerklärungen an der Tafel helfen den SuS beim Lösen von Aufgaben.	☐	☐	☐	☐
AN_{DI}	Das Nachsprechen mit der gesamten Klasse hilft den SuS beim Verstehen der Rechenstrategien.	☐	☐	☐	☐
AN_{SP}	Die Spiele führen zu einem positiven Stundenabschluss.	☐	☐	☐	☐
AN_{PM}	Der Einsatz von heterogenen Arbeitsgruppen hilft den SuS beim Lösen von Aufgaben.	☐	☐	☐	☐
AN_{PM}	Das Üben mit einem festen Partner hilft den SuS beim Lösen von Aufgaben.	☐	☐	☐	☐
AN_{PM}	Das Üben im Trainingslager hilft den SuS beim Lösen von Aufgaben.	☐	☐	☐	☐
AN_{ZG}	Die Olympischen Rechenspiele führen zu einem lernförderlichen Klima. Bei “Stimmt kaum/gar nicht?” - warum nicht? _____ _____	☐	☐	☐	☐
AN_{CBM}	Das Diagramm im Stadionplan macht den SuS die individuelle Entwicklung ihres Lernfortschritts sichtbar.	☐	☐	☐	☐
AN_{CBM}	Die SuS möchten sich in den Rechensprints verbessern.	☐	☐	☐	☐
AN_{VF}	Die vorgegebenen Regeln sind angemessen.	☐	☐	☐	☐
AN_{VF}	Die roten und gelben Karten sind als negative Verstärker angemessen.	☐	☐	☐	☐
AN_{VF}	Die SuS halten sich an die Regeln, um keine rote oder gelbe Karte zu bekommen.	☐	☐	☐	☐
AN_{VF}	Nach dem Einsatz des negativen Verstärkers (rote/gelbe Karte) halten die SuS die Regeln ein.	☐	☐	☐	☐
AN_{VF}	Die Stempel sind als positive Verstärker angemessen.	☐	☐	☐	☐
AN_{VF}	Die SuS halten sich an die Regeln, um einen Stempel zu bekommen.	☐	☐	☐	☐

M_{PR}	Zur erfolgreichen Umsetzung der Intervention waren meine Fähigkeiten ausreichend. Bei "Stimmt kaum/gar nicht?" - welche fehlten ihnen? _____ _____	☐	☐	☐	☐
M_{SR}	Meine Schule verfügte über die notwendigen Ressourcen (Materialien, Ausstattung) zu einer erfolgreichen Umsetzung. Bei "Stimmt kaum/ gar nicht?" - welche fehlten Ihrer Meinung nach?	☐	☐	☐	☐
M_{SR}	Die betreuende Lehrkraft der Schule stand der Umsetzung der Olympischen Rechenspiele positiv gegenüber.	☐	☐	☐	☐
M_{SR}	Die Olympischen Rechenspiele sind in den Unterrichtsalltag implementierbar.	☐	☐	☐	☐
M_{SR}	Die Olympischen Rechenspiele lassen sich durch eine einzelne Lehrperson durchführen.	☐	☐	☐	☐
M_{SR}	Der Zeitaufwand ist in der Vorbereitung ökonomisch und realistisch.	☐	☐	☐	☐
M_{SR}	Der Materialaufwand ist in der Vorbereitung ökonomisch und realistisch.	☐	☐	☐	☐
M_{M}	Das Durchführungsmanual ist einfach zu handhaben.	☐	☐	☐	☐
M_{M}	Die Materialien sind eindeutig gekennzeichnet.	☐	☐	☐	☐
M_{M}	Das Materialmanual ist einfach zu handhaben.	☐	☐	☐	☐
M_{M}	Die Formulierungen und Aufgabenstellungen im Manual sind für den Durchführenden deutlich und verständlich.	☐	☐	☐	☐
M_{M}	Die erforderlichen Vorbereitungen für die Arbeitsphasen und deren Übergänge werden durch das Manual erkennbar.	☐	☐	☐	☐
W	Ich habe den Advance Organizer nach Vorgabe des Manuals umgesetzt.	☐	☐	☐	☐
W	Ich habe die Formulierungen der direkten Instruktionen übernommen.	☐	☐	☐	☐
W	Ich habe die bestehenden Materialien verwendet.	☐	☐	☐	☐

W	Ich habe die bestehenden Materialien mit eigenen Materialien ergänzt. Bei „Stimmt genau/ziemlich?“ - welche Stunden haben Sie womit ergänzt? _____ _____	☐	☐	☐	☐
W	Ich habe die vorgegebene Zeitstruktur der einzelnen Unterrichtsstunden eingehalten.	☐	☐	☐	☐
W	Ich habe die vorgegebene Zeitstruktur der Unterrichtseinheiten eingehalten.	☐	☐	☐	☐

AK_{Note}	Bitte bewerten Sie die Olympischen Rechenspiele abschließend mit einer Schulnote.				
1 (sehr gut)	2 (gut)	3 (befriedigend)	4 (ausreichend)	5 (mangelhaft)	6 (ungenügend)
☐	☐	☐	☐	☐	☐

Welche Materialien haben sich bei der Durchführung als besonders hilfreich erwiesen? Warum?

_____

_____

Welche Schwierigkeiten zeigten sich bei der inhaltlichen und konzeptionellen Durchführung der Olympischen Rechenspiele? Warum?

_____

_____

Welche weiteren Anregungen bzw. Verbesserungsvorschläge haben Sie?

_____

_____

Vielen Dank für Ihre Unterstützung!

Anmerkung: Die Darstellung des Fragebogens kann bezüglich des Layouts vom Original abweichen.

ANHANG 2 ERGEBNISSE DES IMPLEMENTATIONSFRAGEBOGEN DER INTERVENTIONSLEITERINNEN UND -LEITER AUF ITEMEBENE (DESKRIPTIV)

1 = stimmt genau, 2 = stimmt ziemlich, 3 = stimmt kaum, 4 = stimmt gar nicht

Item/Aussage		M	SD
	Ich habe die Olympischen Rechenspiele gerne durchgeführt	–	–
AK_{ZI}	zu Beginn	1.85	0.899
AK_{ZI}	nach der Hälfte	2.23	0.599
AK_{ZI}	am Ende	1.92	0.494
AK_{ZI}	Die Intervention hat meine Erwartungen insgesamt erfüllt.	2.50	0.522
AK_{ZI}	Die Olympischen Rechenspiele haben mir Anregungen für eine zukünftige Förderung der mathematischen Basiskompetenzen gegeben.	1.85	0.689
Ü	Ich würde die Olympischen Rechenspiele ein weiteres Mal durchführen.	2.08	0.862
Ü	Ich würde die Olympischen Rechenspiele weiterempfehlen.	2.23	0.725
AN_{ZG}	Die Olympischen Rechenspiele sind für die Zielgruppe (SuS an Ober- oder Gesamtschulen) angemessen.	1.85	0.555
AN_{ZG}	Die Olympischen Rechenspiele motivieren SuS zum Rechnen.	2.15	0.689
AN_{ZG}	Die Olympischen Rechenspiele machen den SuS Freude beim Rechnen.	2.38	0.768
AN_{PI}	Die Inhalte der Olympischen Rechenspiele sind zur Förderung des Stellenwertsystems und der Grundrechenarten angemessen.	1.69	0.480
AN_{ZP}	Die Zeitplanung ist insgesamt stimmig.	3.62	0.506
AN_{ZP}	Die Zeitplanung in den Stunden ist zu kurz.*	3.23	1.166
AN_{ZP}	Die Zeitplanung in den Stunden ist zu lang.*	1.62	1.044
AN_{ZG}	Die Olympischen Rechenspiele fördern besonders die mathematischen Basiskompetenzen der leistungsstarken SuS.	2.08	0.494
AN_{ZG}	Die Olympischen Rechenspielen fördern besonders die mathematischen Basiskompetenzen der leistungsschwachen SuS.	2.85	0.555
AN_{ZG}	Die Olympischen Rechenspielen fördern besonders die mathematischen Basiskompetenzen von SuS mit diagnostiziertem Unterstützungsbedarf im Bereich Lernen.	3.58	0.669
AN_{ZP}	Die Stundenplanung ermöglicht es, einzelnen SuS gezielt Hilfestellung zu geben.	2.91	0.831
AN_{ZG}	Die Materialien sind für leistungsstarke SuS lernförderlich differenziert.	2.15	0.555
AN_{ZG}	Die Materialien sind für leistungsschwache SuS lernförderlich differenziert.	2.77	0.725
AN_{PME}	Die eingesetzten Methoden der Olympischen Rechenspiele sind zur Förderung des Stellenwertsystems und der Grundrechenarten angemessen.	2.08	0.641

AN_{AO}	Der Advance Organizer verdeutlicht den SuS das Thema der Stunde.	1.46	0.519
AN_{AO}	Der Advance Organizer gibt den SuS einen Rückblick und eine Vorschau auf die Unterrichtsthemen.	1.46	0.519
AN_{AB}	Die Arbeitsblätter helfen den SuS beim Lösen von Aufgaben.	2.08	0.669
AN_{DI}	Die Strategieerklärungen an der Tafel helfen den SuS beim Lösen von Aufgaben.	1.92	0.494
AN_{DI}	Das Nachsprechen mit der gesamten Klasse hilft den SuS beim Verstehen der Rechenstrategien.	2.62	0.870
AN_{SP}	Die Spiele führen zu einem positiven Stundenabschluss.	1.69	0.630
AN_{PM}	Der Einsatz von heterogenen Arbeitsgruppen hilft den SuS beim Lösen von Aufgaben.	2.69	0.751
AN_{PM}	Das Üben mit einem festen Partner hilft den SuS beim Lösen von Aufgaben.	2.50	0.674
AN_{PM}	Das Üben im Trainingslager hilft den SuS beim Lösen von Aufgaben.	3.25	0.754
AN_{ZG}	Die Olympischen Rechenspiele führen zu einem lernförderlichen Klima.	2.46	0.519
AN_{CBM}	Das Diagramm im Stadionplan macht den SuS die individuelle Entwicklung ihres Lernfortschritts sichtbar.	1.54	0.660
AN_{CBM}	Die SuS möchten sich in den Rechensprints verbessern.	1.85	0.689
AN_{VF}	Die vorgegebenen Regeln sind angemessen.	1.77	0.832
AN_{VF}	Die roten und gelben Karten sind als negative Verstärker angemessen.	2.00	0.816
AN_{VF}	Die SuS halten sich an die Regeln, um keine rote oder gelbe Karte zu bekommen.	2.50	0.522
AN_{VF}	Nach dem Einsatz des negativen Verstärkers (rote/gelbe Karte) halten die SuS die Regeln ein.	2.50	0.674
AN_{VF}	Die Stempel sind als positive Verstärker angemessen.	1.23	0.439
AN_{VF}	Die SuS halten sich an die Regeln, um einen Stempel zu bekommen.	1.77	0.832
M_{PR}	Zur erfolgreichen Umsetzung der Intervention waren meine Fähigkeiten ausreichend.	1.85	0.555
M_{SR}	Meine Schule verfügte über die notwendigen Ressourcen (Materialien, Ausstattung) zu einer erfolgreichen Umsetzung.	2.08	0.862
M_{SR}	Die betreuende Lehrkraft der Schule stand der Umsetzung der Olympischen Rechenspiele positiv gegenüber.	2.77	0.832
M_{SR}	Die Olympischen Rechenspiele sind in den Unterrichtsalltag implementierbar.	2.27	0.786

M_{SR}	Die Olympischen Rechenspiele lassen sich durch eine einzelne Lehrperson durchführen.	2.23	0.725
M_{SR}	Der Zeitaufwand ist in der Vorbereitung ökonomisch und realistisch.	1.85	0.689
M_{SR}	Der Materialaufwand ist in der Vorbereitung ökonomisch und realistisch.	2.46	1.05
M_{M}	Das Durchführungsmanual ist einfach zu handhaben.	1.46	0.519
M_{M}	Die Materialien sind eindeutig gekennzeichnet.	1.46	0.519
M_{M}	Das Materialmanual ist einfach zu handhaben.	1.46	0.519
M_{M}	Die Formulierungen und Aufgabenstellungen im Manual sind für den Durchführenden deutlich und verständlich.	1.85	0.376
M_{M}	Die erforderlichen Vorbereitungen für die Arbeitsphasen und deren Übergänge werden durch das Manual erkennbar.	1.85	0.689
W	Ich habe den Advance Organizer nach Vorgabe des Manuals umgesetzt.	1.38	0.506
W	Ich habe die Formulierungen der direkten Instruktionen übernommen.	2.23	0.832
W	Ich habe die bestehenden Materialien verwendet.	1.31	0.480
W	Ich habe die vorgegebene Zeitstruktur der einzelnen Unterrichtsstunden eingehalten.	2.62	0.961
W	Ich habe die vorgegebene Zeitstruktur der Unterrichtseinheiten eingehalten.	2.15	0.801
AK_{Note}	Bitte bewerten Sie die Olympischen Rechenspiele mit einer Schulnote.	2.69	0.947

* Anmerkung: hier gilt in der Auswertung: 1 = stimmt gar nicht, 2 = stimmt kaum, 3 = stimmt ziemlich, 4 = stimmt genau

ANHANG 3 IMPLEMENTATIONSFRAGEBÖGEN DER SCHÜLERINNEN UND SCHÜLER

Schülerfragebogen zu den Olympischen Rechenspielen

Anleitung

Auf den folgenden Seiten stehen einige Sätze. Damit wollen wir herausfinden, wie die Olympischen Rechenspiele den Schülerinnen und Schülern gefallen haben.

Lies die Sätze der Reihe nach durch und kreuze an, ob der Satz für dich stimmt.

Markiere die Kästchen (❑), indem du sie ausfüllst (■). Mache bitte keine einfachen Kreuze oder Häkchen!

Überlege bei den einzelnen Sätzen nicht zu lange und antworte so, wie es dir sofort als richtig erscheint.

		stimmt genau	stimmt ziemlich	stimmt kaum	stimmt gar nicht
AK_{LE}	Nach den Olympischen Rechenspielen fällt mir das Lesen der Stellenwerttafel leichter. ***Beispiel:*** H Z E 	☐	☐	☐	☐
AK_{LE}	Nach den Olympischen Rechenspielen fällt mir das Plusrechnen leichter. ***Beispiel:*** 345 + 212 =?	☐	☐	☐	☐
AK_{LE}	Nach den Olympischen Rechenspielen fällt mir das Minusrechnen leichter. ***Beispiel:*** 612 – 323 = ?	☐	☐	☐	☐
AK_{LE}	Nach den Olympischen Rechenspielen fällt mir das Malrechnen leichter. ***Beispiel:*** 23 · 34 = ?	☐	☐	☐	☐
AK_{LE}	Nach den Olympischen Rechenspielen fällt mir das Geteiltrechnen leichter. ***Beispiel:*** 241 : 3 = ?	☐	☐	☐	☐
		stimmt genau	**stimmt ziemlich**	**stimmt kaum**	**stimmt gar nicht**
AN_{AB}	Ich konnte die Aufgaben auf den Arbeitsblättern lösen.	☐	☐	☐	☐
AN_{AB}	Ich habe alle Aufgaben auf den Arbeitsblättern geschafft.	☐	☐	☐	☐
AN_{AB}	Die Rechenbeispiele auf den Arbeitsblättern haben mir beim Lösen der Aufgaben geholfen.	☐	☐	☐	☐
AN_{AB}	Ich habe die Aufgabenstellung gelesen, bevor ich mit dem Lösen der Aufgaben angefangen bin.	☐	☐	☐	☐
AN_{AB}	Ich wusste, wie ich die Aufgaben auf den Arbeitsblättern lösen konnte.	☐	☐	☐	☐

		stimmt genau	stimmt ziemlich	stimmt kaum	stimmt gar nicht
AN_{AO}	Durch das Plakat am Unterrichtsbeginn kannte ich das Thema der Stunde. *Plakat:* 	☐	☐	☐	☐
AN_{AO}	Durch das Plakat am Unterrichtsbeginn wusste ich, welche Themen wir schon in den Olympischen Rechenspielen hatten und welche noch kommen.	☐	☐	☐	☐
AN_{DI}	Die Erklärungen der Rechenwege an der Tafel haben mir beim Lösen von Aufgaben geholfen.	☐	☐	☐	☐
AN_{DI}	Das Nachsprechen der Rechenwege mit der gesamten Klasse hat mir geholfen, die Rechenwege zu verstehen.	☐	☐	☐	☐
AN_{AB}	Das Üben auf den Arbeitsblättern hat mir geholfen, Aufgaben zu lösen.	☐	☐	☐	☐
AN_{PM}	Das Üben mit meinen Mitschülern im Trainingslager hat mir beim Lösen von Aufgaben geholfen.	☐	☐	☐	☐
AN_{PM}	Das Üben mit meinem festen Partner hat mir beim Lösen von Aufgaben geholfen.	☐	☐	☐	☐
AN_{CBM}	Ich habe auf dem Stadionplan erkannt, ob ich mich in den Rechensprints verbessert oder verschlechtert habe.	☐	☐	☐	☐
AN_{CBM}	Ich wollte mich in den Rechensprints verbessern.	☐	☐	☐	☐
AK_{ZF}	Der Unterrichtsbeginn mit dem Plakat hat mir gefallen. *Plakat:* 	☐	☐	☐	☐
AK_{ZF}	Die Erklärungen des Lehrers an der Tafel haben mir gefallen.	☐	☐	☐	☐
AK_{ZF}	Das Nachsprechen der Rechenwege mit der gesamten Klasse hat mir gefallen.	☐	☐	☐	☐
AK_{ZF}	Das Üben mit den Arbeitsblättern hat mir gefallen.	☐	☐	☐	☐
AK_{ZF}	Das Üben mit meinen Mitschülern im Trainingslager hat mir gefallen.	☐	☐	☐	☐
AK_{ZF}	Das Üben mit meinem festen Partner hat mir gefallen.	☐	☐	☐	☐
AK_{ZF}	Die Spiele haben mir gefallen.	☐	☐	☐	☐
AK_{ZF}	Die Rechensprints haben mir gefallen.	☐	☐	☐	☐

AK_{ZF}	Der Stempel und der Stadionplan haben mir gefallen.	☐	☐	☐	☐
		stimmt genau	**stimmt ziemlich**	**stimmt kaum**	**stimmt gar nicht**
AN_{VF}	Ich habe mich an die Regeln gehalten, um keine rote oder gelbe Karte zu bekommen.	☐	☐	☐	☐
AN_{VF}	Ich habe mich an die Regeln gehalten, um einen Stempel zu bekommen.	☐	☐	☐	☐
AN_{VF}	Ich habe mal eine rote oder gelbe Karte bekommen.	☐	☐	☐	☐
AN_{VF}	Wenn „stimmt genau“: Ich habe mich nach dem Austeilen einer roten oder gelben Karte an die Regeln gehalten.	☐	☐	☐	☐
		stimmt genau	**stimmt ziemlich**	**stimmt kaum**	**stimmt gar nicht**
AK_{ZI}	Die Olympischen Rechenspiele haben mir gefallen.	☐	☐	☐	☐
AK_{ZI}	Durch die Olympischen Rechenspiele hatte ich Freude am Rechnen.	☐	☐	☐	☐

AK_{Note}	Ich gebe den Olympischen Rechenspielen die Note:				
1 (sehr gut)	2 (gut)	3 (befriedigend)	4 (ausreichend)	5 (mangelhaft)	6 (ungenügend)
☐	☐	☐	☐	☐	☐

Was hat dir an den Olympischen Rechenspielen besonders Spaß gemacht?

_____

_

Was würdest du an den Olympischen Rechenspielen ändern?

_____

_

Vielen Dank für deine Mitarbeit!

Anmerkung: Die Darstellung des Fragebogens kann bezüglich des Layouts vom Original abweichen.

ANHANG 4 ERGEBNISSE DES IMPLEMENTATIONSFRAGEBOGEN DER SCHÜLERINNEN UND SCHÜLER AUF ITEMEBENE (DESKRIPTIV)

1 = stimmt genau, 2 = stimmt ziemlich, 3 = stimmt kaum, 4 = stimmt gar nicht

		M	SD
AK_{LE}	Nach den Olympischen Rechenspielen fällt mir das Lesen der Stellenwerttafel leichter.	1.92	1.126
AK_{LE}	Nach den Olympischen Rechenspielen fällt mir das Plus-rechnen leichter.	2.08	1.123
AK_{LE}	Nach den Olympischen Rechenspielen fällt mir das Minus-rechnen leichter.	2.17	1.024
AK_{LE}	Nach den Olympischen Rechenspielen fällt mir das Mal-rechnen leichter.	2.26	1.033
AK_{LE}	Nach den Olympischen Rechenspielen fällt mir das Geteilt-rechnen leichter.	2.41	1.087
AN_{AB}	Ich konnte die Aufgaben auf den Arbeitsblättern lösen.	1.69	0.717
AN_{AB}	Ich habe alle Aufgaben auf den Arbeitsblättern geschafft.	2.38	0.857
AN_{AB}	Die Rechenbeispiele auf den Arbeitsblättern haben mir beim Lösen der Aufgaben geholfen.	2.14	1.068
AN_{AB}	Ich habe die Aufgabenstellung gelesen, bevor ich mit dem Lösen der Aufgaben angefangen bin.	1.69	0.884
AN_{AB}	Ich wusste, wie ich die Aufgaben auf den Arbeitsblättern lösen konnte.	1.76	0.793
AN_{AO}	Durch das Plakat am Unterrichtsbeginn kannte ich das Thema der Stunde.	1.85	1.024
AN_{AO}	Durch das Plakat am Unterrichtsbeginn wusste ich, welche Themen wir schon in den Olympischen Rechenspielen hatten und welche noch kommen.	1.80	0.983
AN_{DI}	Die Erklärungen der Rechenwege an der Tafel haben mir beim Lösen von Aufgaben geholfen.	1.96	0.943
AN_{DI}	Das Nachsprechen der Rechenwege mit der gesamten Klasse hat mir geholfen, die Rechenwege zu verstehen.	2.33	1.016
AN_{AB}	Das Üben auf den Arbeitsblättern hat mir geholfen, Aufgaben zu lösen.	1.97	0.936
AN_{PM}	Das Üben mit meinen Mitschülern im Trainingslager hat mir beim Lösen von Aufgaben geholfen.	2.48	1.048
AN_{PM}	Das Üben mit meinem festen Partner hat mir beim Lösen von Aufgaben geholfen.	2.27	1.056
AN_{CBM}	Ich habe auf dem Stadionplan erkannt, ob ich mich in den Rechensprints verbessert oder verschlechtert habe.	1.78	0.976
AN_{CBM}	Ich wollte mich in den Rechensprints verbessern.	1.74	0.982
AK_{ZF}	Der Unterrichtsbeginn mit dem Plakat hat mir gefallen.	2.00	0.986
AK_{ZF}	Die Erklärungen des Lehrers an der Tafel haben mir gefallen.	2.00	0.952
AK_{ZF}	Das Nachsprechen der Rechenwege mit der gesamten Klasse hat mir gefallen.	2.31	1.036
AK_{ZF}	Das Üben mit den Arbeitsblättern hat mir gefallen.	2.05	1.019
AK_{ZF}	Das Üben mit meinen Mitschülern im Trainingslager hat mir gefallen.	2.44	1.093

AK_{ZF}	Das Üben mit meinem festen Partner hat mir gefallen.	2.17	1.057
AK_{ZF}	Die Spiele haben mir gefallen.	1.76	0.985
AK_{ZF}	Die Rechensprints haben mir gefallen.	2.16	1.102
AK_{ZF}	Der Stempel und der Stadionplan haben mir gefallen.	1.73	1.006
AN_{VF}	Ich habe mich an die Regeln gehalten, um keine rote oder gelbe Karte zu bekommen.	1.54	0.816
AN_{VF}	Ich habe mich an die Regeln gehalten, um einen Stempel zu bekommen.	1.56	0.830
AN_{VF}	Ich habe mal eine rote oder gelbe Karte bekommen.	1.75	0.432
AN_{VF}	Wenn „stimmt genau“: Ich habe mich nach dem Austeilen einer roten oder gelben Karte an die Regeln gehalten.	2.08	1.164
AK_{ZI}	Die Olympischen Rechenspiele haben mir gefallen.	2.03	1.026
AK_{ZI}	Durch die Olympischen Rechenspiele hatte ich Freude am Rechnen.	2.25	1.041
AK_{Note}	Bitte bewerte die Olympischen Rechenspiele abschließend mit einer Schulnote.	2.11	0.707

ANHANG 5 ÜBERPRÜFUNG AUF NORMALVERTEILUNG IM HYPOTHESENBLOCK I UND II

			Kolmogorov-Smirnov			Shapiro-Wilk		
		n	df	p	normalverteilt	df	p	normalverteilt
Gesamtgruppe	EG Prä	251	251	0.200	ja	251	0.515	ja
	EG Post	239	239	0.200	ja	239	0.596	ja
	EG FU-I	226	226	0.200	ja	226	0.608	ja
	EG FU-II	228	228	0.200	ja	228	0.103	ja
	KG Prä	359	359	0.200	ja	359	0.288	ja
	KG Post	356	356	0.200	ja	356	0.548	ja
	KG FU-I	335	335	0.200	ja	335	0.266	ja
	KG FU-II	328	328	0.045	nein	328	0.242	ja
Geringes Vorwissen	EG Prä	75	75	0.002	nein	75	0.000	nein
	EG Post	68	68	0.200	ja	68	0.258	0.200
	EG FU-I	65	65	0.200	ja	65	0.667	0.200
	EG FU-II	60	60	0.200	ja	60	0.429	0.200
	KG Prä	116	116	0.019	nein	116	0.000	nein
	KG Post	110	110	0.200	ja	110	0.403	ja
	KG FU-I	106	106	0.042	nein	106	0.000	nein
	KG FU-II	100	100	0.200	nein	100	0.004	nein
Geringe kognitive Fähigkeiten	EG Prä	52	52	0.200	ja	52	0.426	ja
	EG Post	49	49	0.200	ja	49	0.377	ja
	EG FU-I	45	45	0.073	ja	45	0.061	ja
	EG FU-II	45	45	0.200	ja	45	0.396	ja
	KG Prä	63	63	0.200	ja	63	0.335	ja
	KG Post	60	60	0.200	ja	60	0.783	ja
	KG FU-I	56	56	0.200	ja	56	0.940	ja
	KG FU-II	56	56	0.200	ja	56	0.940	ja

			Kolmogorov-Smirnov			Shapiro-Wilk		
		n	df	p	normalverteilt	df	p	normalverteilt
Geringe Lesekompetenzen	EG Prä	154	154	0.200	ja	154	0.417	ja
	EG Post	142	142	0.200	ja	142	0.777	ja
	EG FU-I	137	137	0.200	ja	137	0.594	ja
	EG FU-II	134	134	0.200	ja	134	0.378	ja
	KG Prä	197	197	0.200	ja	197	0.557	ja
	KG Post	189	189	0.200	ja	189	0.493	ja
	KG FU-I	170	170	0.200	ja	170	0.384	ja
	KG FU-II	169	169	0.200	ja	169	0.103	ja
Risikobehaftetes Sozialverhalten	EG Prä	18	18	0.200	ja	18	0.267	ja
	EG Post	18	18	0.200	ja	18	0.416	ja
	EG FU-I	17	17	0.200	ja	17	0.419	ja
	EG FU-II	17	17	0.073	ja	17	0.420	ja
	KG Prä	20	20	0.200	ja	20	0.400	ja
	KG Post	19	19	0.200	ja	19	0.109	ja
	KG FU-I	17	17	0.200	ja	17	0.379	ja
	KG FU-II	13	13	0.049	nein	13	0.092	ja
Risikobehaftetes Lernverhalten	EG Prä	22	22	0.019	nein	22	0.094	ja
	EG Post	21	21	0.200	ja	21	0.589	ja
	EG FU-I	22	22	0.200	ja	22	0.170	ja
	EG FU-II	20	20	0.200	ja	20	0.273	ja
	KG Prä	23	23	0.200	ja	23	0.473	ja
	KG Post	23	23	0.200	ja	23	0.127	ja
	KG FU-I	23	23	0.200	ja	23	0.115	ja
	KG FU-II	20	20	0.183	ja	20	0.283	ja
Unterstützungsbedarf Lernen	EG Prä	7	7	0.200	ja	7	0.314	Ja
	EG Post	7	7	0.200	ja	7	0.544	Ja
	EG FU-I	7	7	0.200	ja	7	0.666	Ja
	EG FU-II	7	7	0.200	ja	7	0.443	Ja
	KG Prä	12	12	0.200	ja	12	0.761	ja
	KG Post	12	12	0.200	ja	12	0.331	ja
	KG FU-I	12	12	0.200	ja	12	0.233	ja
	KG FU-II	12	12	0.090	ja	12	0.162	ja

ANHANG 6 ÜBERPRÜFUNG AUF NORMALVERTEILUNG IM HYPOTHESENBLOCK III

		Kolmogorov-Smirnov			Shapiro-Wilk		
	n	df	p	normalverteilt	df	p	normalverteilt
Lernfreude	243	243	0.200	ja	243	0.493	ja
Durchschn. Standardabweichung der Klasse in der Prä-Erhebung	243*	243	0.000	nein	243	0.000	nein

Anmerkung: * Jedem Schüler/jeder Schülerin wurde die durchschnittliche Standardabweichung im HRT 1–4 (Prä) der eigenen Klasse zugeordnet

ANHANG 7 ÜBERPRÜFUNG AUF NORMALVERTEILUNG IM HYPOTHESENBLOCK IV

			Kolmogorov-Smirnov			**Shapiro-Wilk**		
		n	df	p	normalverteilt	df	p	normalverteilt
Schülerinnen und Schüler	Akzeptanz	250	250	0.000	nein	250	0.000	nein
	Angemessenheit	250	250	0.007	nein	250	0.000	nein
Interventionsleiterinnen und -leiter	Akzeptanz	13	13	0.058	ja	13	0.181	ja
	Übernahme	13	13	0.006	nein	13	0.009	nein
	Angemessenheit	13	13	0.200	ja	13	0.681	ja
	Machbarkeit	13	13	0.005	nein	13	0.925	ja
	Wiedergabetreue	13	13	0.200	ja	13	0.565	ja

ERKLÄRUNG

Hiermit versichere ich, dass ich die vorliegende Dissertation mit dem Titel *„Die Förderung mathematischer Basiskompetenzen zu Beginn der Sekundarstufe I – Evaluation und Implementation eines Trainingsprogramms zur Förderung der mathematischen Basiskompetenzen im inklusiven Setting zu Beginn der Sekundarstufe I“* selbstständig verfasst und deren Inhalt nicht schon für eine Diplom-, Bachelor- oder Masterarbeit verwendet habe. Die benutzten Hilfsmittel habe ich vollständig angegeben.

Hiermit versichere ich, dass die Leitlinien guter wissenschaftlicher Praxis an der Carl von Ossietzky Universität Oldenburg befolgt worden sind.

Carolina Käter