Unterrichtsintegrierte Förderung des Operationsverständnisses

Eine Interventionsstudie im 2. Schuljahr

Dissertation

zur Erlangung der Würde eines Doktors der Philosophie

vorgelegt dem

Institut für Bildungswissenschaften der Universität Basel

von

Jon Florin
aus
Val Müstair GR

Zürich, 2021

Druckerei: Kopierwerk, Zürich

Originaldokument gespeichert auf dem Dokumentenserver der Universität Basel
edoc.unibas.ch

Dieses Werk ist lizenziert unter einer Creative Commons Namensnennung 4.0 International Lizenz.

Genehmigt vom Institut für Bildungswissenschaften der Universität Basel, auf Antrag von Prof. Dr. Christine Streit, Prof. Dr. Jérémy Blanc und Prof. Dr. Michael Gaidoschik.

Basel, den 25. Mai 2021

Die Direktorin Prof. Dr. Elena Makarova

Zusammenfassung

Die Fähigkeit von Kindern, Verknüpfungen zwischen mathematischen Operationen und Sachsituationen, Handlungen oder Bildern vorzunehmen («den Darstellungswechsel vollziehen zu können»), wird als wesentliches Merkmal zur Einschätzung ihres Operationsverständnisses herangezogen. Diese Fähigkeit der Kinder wurde in der vorliegenden Arbeit untersucht.

Im Rahmen einer Interventionsstudie wurde mit Kindern in der 2. Primarschulklasse untersucht, ob sich eine Förderung des Operationsverständnisses auf die allgemeinen Rechenfähigkeiten der Kinder auswirkt. Die Interventionen erfolgten unterrichtsintegriert im Rahmen eines kooperativ-kommunikativen Settings. Es wurde auch der Frage nachgegangen, ob sich die Wirkung der Interventionen bei den rechenstarken und bei den rechenschwachen Kindern unterscheidet. Gleichzeitig wurde ein neuartiges Erhebungsinstrument zur Erfassung des Operationsverständnisses für die Multiplikation erprobt und eingesetzt. Die vorliegende Arbeit umfasst auch eine Analyse und Beurteilung dieses neuen Erhebungsinstruments.

Im Schuljahr 2019/20 konnten die für die Beantwortung der Forschungsfragen notwendigen Längsschnittuntersuchungen in insgesamt 46 Schulklassen des Forschungsprojekts MALKA (www.malka1und2.ch) durchgeführt werden. Für die Datenauswertungen wurde nebst der klassischen Testtheorie vor allem das eindimensionale Rasch-Modell aus der Item Response Theorie eingesetzt. Die Analysen wurden mit Hilfe von Mehrebenen-Modellen unter Berücksichtigung der hierarchischen Struktur der Daten (Kinder in Schulklassen) vorgenommen.

Es wurde beobachtet, dass die rechenschwachen Kinder von den Interventionen profitierten. Die rechenstarken Kinder hingegen erzielten im Unterricht in der Kontrollgruppe, in der die Interventionen nicht eingesetzt wurden, einen grösseren Lernfortschritt als in der Interventionsgruppe. Eine positive Wirkung der Interventionen beim Vergleich ganzer Klassen konnte nicht nachgewiesen werden.

Die Förderung des Operationsverständnisses im kooperativ-kommunikativen Setting mit wissenschaftlicher Erfassung der Auswirkungen dieser Unterrichtsform auf den Lernzuwachs der Kinder geschah erstmalig.

Das Schuljahr 2019/20 wurde durch eine unerwartete sechswöchige Schulschliessung aufgrund der Covid-19-Pandemie beeinflusst. Durch die Schulschliessung wurde der Einsatz der Unterrichts-Interventionen unterbrochen. Die Ergebnisse dieser Arbeit sind aufgrund dieses Vorfalls mit entsprechender Vorsicht zu behandeln.

Danksagung

Diese Arbeit war in ein mehrjähriges und umfangreiches Projekt eingebettet, zu dem viele Personen beigetragen haben, denen ich mich in verschiedenster Weise zu Dank verpflichtet sehe. Es wird mir kaum gelingen, mit diesem Text eine vollständige Danksagung zusammen zu stellen, bei der ich niemanden vergesse. Die folgenden Ausführungen können zumindest als eine Bemühung angesehen werden, derart Unmögliches zu erreichen.

Zuerst und ganz besonders danken möchte ich Prof. Dr. Christine Streit, für die Anstellung als Wissenschaftlicher Mitarbeiter und damit auch für die Ermöglichung dieser Projektarbeit sowie natürlich für die langjährige, regelmässige, optimistische und unermüdliche Begleitung und Unterstützung während der Entstehung dieser Arbeit. Ohne ihr Mitdenken und Mitwirken wäre diese Arbeit nicht entstanden.

Diese Arbeit wäre aber auch nicht möglich gewesen ohne das Führungsgremium und ohne das gesamte Team des Forschungsprojekts MALKA. Meine Arbeit wurde des Weiteren unterstützt durch viele studentische Hilfskräfte der Universität Zürich sowie durch viele Lehrpersonen, die sich in unterschiedlicher Form – bei Erprobungen des Materials oder durch die aktive Teilnahme mit ihren Klassen − im Forschungsprojekt beteiligt haben. Nicht zuletzt gilt mein Dank auch einer grossen Zahl von Kindern und Eltern sowie erziehungsberechtigten Personen, die alle durch ihre Teilnahme bzw. durch ihr Einverständnis dieses Forschungsprojekt ermöglicht haben.

Ebenfalls danken möchte ich den Zweit- und Drittgutachtern dieser Arbeit für ihren Einsatz, den Kolleginnen und Kollegen am Institut für Erziehungswissenschaften der Universität Zürich für das gemeinsame Arbeiten, den Kolleginnen und Kollegen der Mathematikdidaktik am Institut für Kindergarten und Unterstufe der PH FHNW in Muttenz, den weiteren Kolleginnen und Kollegen aus der Mathematikdidaktik der PH FHNW, dem Team des Instituts für Bildungswissenschaften der Universität Basel, der einzigen studentischen Hilfskraft der FHNW in der Person von Chantal Strasser, den Kolleginnen und Kollegen, die als Promovierende an der Universität Basel in den letzten Jahren mit mir im selben Boot sassen, den Methodenberatungen der verschiedenen beteiligten Hochschulen, den Expertinnen und Experten für die Rückmeldungen zu den Fördereinheiten, verschiedenen Referentinnen und Referenten, von denen ich im Rahmen der Promotionsausbildung viel Wichtiges und insgesamt einiges mehr lernen konnte, als ich mir langfristig werde merken können, swissuniversities für die Finanzierung und Ermöglichung dieser Forschungsarbeit, dem Internet als unerschöpflicher Quelle von Wissen und Weisheiten aller Art sowie allen, die nach wie vor Bücher und Hefte in Papierform drucken, an denen ich trotz online-Lehre und Digitalisierung immer noch viel Freude finde.

Auch meinem privaten Umfeld, das mich während der Entstehung dieser Arbeit ganz wesentlich unterstützt hat, gilt ein ganz herzlicher Dank für die viele Ermutigung, Unterhaltung, für Ablenkung durch interessante Stellungen auf dem Schachbrett oder durch Sonstiges, für gelegentliche Verköstigung, für in seiner Auswirkung nicht zu unterschätzendes Geplauder über meine berufliche Tätigkeit und für gutes Zureden jeglicher Art.

Inhaltsverzeichnis

1 EINLEITUNG ... 1

2 OPERATIONSVERSTÄNDNIS ALS THEORETISCHES KONSTRUKT ... 5

2.1 Operation, Verständnis und Operationsverständnis ... 5

2.1.1 Die Operation ... 5

2.1.2 Das Verständnis ... 8

2.1.3 Das Operationsverständnis ... 14

2.2 Darstellungswechsel, Grundvorstellungen und Strukturierungsfähigkeiten ... 19

2.2.1 Der Darstellungswechsel ... 19

2.2.2 Der Darstellungswechsel als Indiz für Operationsverständnis ... 21

2.2.3 Der Darstellungswechsel zwischen dem Term und anderen Darstellungen ... 26

2.2.4 Das Konzept der Grundvorstellungen ... 29

2.2.5 Grundvorstellungen und Operationsverständnis ... 33

2.2.6 Strukturierungsfähigkeiten und Operationsverständnis ... 38

2.3 Das Konstrukt Operationsverständnis ... 40

2.3.1 Die vier einzelnen Grundrechenarten und deren vereinigte Betrachtung ... 40

2.3.2 «Operation Sense» − die andersartige Betrachtung im englischsprachigen Raum ... 42

2.3.3 Das Konstrukt Operationsverständnis – eine Zusammenfassung ... 45

3 OPERATIONSVERSTÄNDNIS – BEDEUTUNG, BEFUNDE UND FÖRDERUNG ... 47

3.1 Operationsverständnis in seiner Bedeutung für das Rechnenlernen ... 47

3.1.1 Voraussetzungen für Operationsverständnis ... 47

3.1.2 Operationsverständnis und Rechenleistung ... 50

3.1.3 Rechenschwäche und Operationsverständnis als Folge vielfältiger Einflüsse ... 53

3.1.4 Bedeutung und Besonderheiten der Grundrechenart «Multiplikation» ... 56

3.2 Befunde zum «Operationsverständnis» ... 60

3.2.1 Untersuchungen unter Verwendung von Interviews ... 60

3.2.2 Interventionsstudien und quantitative Untersuchungen ... 62

3.3 Förderung von Operationsverständnis in unterschiedlichen Settings ... 64

3.3.1 Behandlung im regulären Unterricht ... 64

3.3.2 Spezifische Förderung des Operationsverständnisses ... 65

3.3.3 Sprachsensibilität der Förderung und Bedeutung der Sprache in der Mathematik ... 67

3.3.4 Einsatz von Material und Anschauungsmitteln für Förderzwecke ... 68

4 EMPIRISCHER TEIL 1: FORSCHUNGSPROJEKT UND METHODISCHES VORGEHEN 71

4.1 Einbettung in ein umfassendes Forschungsprojekt 71
4.1.1 Beschreibung des Forschungsprojekts 71
4.1.2 Gesamtstichprobe 74
4.1.3 Erhebungsinstrumente 75
4.1.4 Durchführung der Erhebungen 78

4.2 Darstellung der Untersuchung und Vorgehen 81
4.2.1 Forschungsfragen 81
4.2.2 Design 82
4.2.3 Stichprobe 83
4.2.4 Eingesetzte Instrumente 87
4.2.5 Intervention 88
4.2.6 Hypothesen 92

4.3 Fördereinheiten 93
4.3.1 Grundlagen der Fördereinheiten 93
4.3.2 Das kooperative Setting als Unterrichtsform 95
4.3.3 Entwicklung der Fördereinheiten 99
4.3.4 Aufbau und Umsetzung des Förderpakets 100
4.3.5 Aufbau einer einzelnen Fördereinheit 104

4.4 Das Erhebungsinstrument für das Multiplikationsverständnis 108
4.4.1 Vorarbeiten für einen Test des Multiplikationsverständnisses 108
4.4.2 Entwicklung und Erprobungen des Testinstruments 109
4.4.3 Beurteilung und Diskussion der Testgütekriterien 117
4.4.4 Einschränkungen des Geltungsbereichs 121
4.4.5 Vor- und Nachtest «Multiplikationsverständnis» 124

4.5 Auswertungsmethoden 127
4.5.1 Klassische Testtheorie 127
4.5.2 Item Response Theorie (IRT) 128
4.5.3 Überprüfung von Differential Item Functioning 133
4.5.4 Modellprüfung und Kriterien zur Modellbewertung 138
4.5.5 Mehrebenenanalyse 139
4.5.6 Auswertungsmethoden für die Längsschnittanalyse 143

5 EMPIRISCHER TEIL 2: ANALYSE DES ERHEBUNGSINSTRUMENTS FÜR DAS MULTIPLIKATIONSVERSTÄNDNIS ... 152

5.1 Datenaufbereitung und deskriptive Statistiken (Vor- und Nachtest) ... 152
5.1.1 Korrektur und Umwandlung von Sondercodes, Umgang mit fehlenden Daten ... 152
5.1.2 Testscore und deskriptive Statistiken ... 154

5.2 Analyse mittels klassischer Testtheorie (Vor- und Nachtest) ... 158
5.2.1 Item-Schwierigkeitsindex ... 158
5.2.2 Trennschärfen und interne Konsistenz ... 160
5.2.3 Inter-Item-Korrelationen ... 162

5.3 Faktoranalysen ... 163
5.3.1 Faktoranalyse für den Vortest ... 163
5.3.2 Faktoranalyse für den Nachtest ... 168

5.4 Analyse mittels IRT (Vortest) ... 172
5.4.1 Parameter des Rasch-Modells ... 172
5.4.2 DIF-Analysen für die einzelnen Items ... 175

5.5 Analyse mittels IRT (Nachtest) ... 179
5.5.1 Parameter des Rasch-Modells ... 179
5.5.2 DIF-Analysen für die einzelnen Items ... 181

5.6 Zusammenhänge mit anderen an der Stichprobe erfassten Grössen (Vortest) ... 183
5.6.1 Korrelationen ... 183
5.6.2 Einfluss der Einteilung in die Schulklassen ... 184
5.6.3 Einfluss des Lehrmittels ... 185

5.7 Itemauswahl, Konstruktion einer Testskala ... 186
5.7.1 Konstruktion einer Testskala (Vortest) ... 186
5.7.2 Konstruktion einer Testskala (Nachtest) ... 187

6 EMPIRISCHER TEIL 3: ERGEBNISSE DER LÄNGSSCHNITTANALYSEN ... 188

6.1 Längsschnittanalyse mit dem Instrument BASIS-MATH (BM) ... 188
6.1.1 Lernfortschritt aller Kinder (BM) ... 188
6.1.2 Lernfortschritt von Kindern mit unterschiedlicher Rechenleistung (BM) ... 190
6.1.3 Variationen des gewählten Mehrebenen-Modells ... 198

6.2 Längsschnittanalyse für das Operationsverständnis (gesamte Skala) ... 210
6.2.1 Lernfortschritt aller Kinder (OV) ... 210
6.2.2 Lernfortschritt von Kindern mit unterschiedlicher Rechenleistung (OV) ... 211

6.3 Längsschnittanalyse für das Operationsverständnis (Brückenitems) ... 216
6.3.1 DIF-Analyse der 12 Brückenitems im Längsschnitt ... 216
6.3.2 Lernfortschritt aller Kinder (Brückenitems) ... 218

7 DISKUSSION UND AUSBLICK ... 220

7.1 Diskussion ... 220
7.1.1 Erhebungsinstrument für das Multiplikationsverständnis ... 220
7.1.2 Längsschnittanalysen ... 221
7.1.3 Weitere Beobachtungen ... 226

7.2 Ausblick ... 227

8 ANHÄNGE ... 230

8.1 Organisatorisches im Projekt MALKA ... 230
8.1.1 Strukturdaten ... 230
8.1.2 Testzeitpunkte ... 231

8.2 Testmaterial ... 232
8.2.1 Vortest Operationsverständnis (im Projekt verwendete Version) ... 232
8.2.2 Codiermanual des Vortests Operationsverständnis ... 235
8.2.3 Nachtest Operationsverständnis (im Projekt verwendete Version) ... 238
8.2.4 Codiermanual des Nachtests Operationsverständnis ... 241

8.3 Interventionsmaterial ... 245
8.3.1 Fördereinheit FE06 – Ablauf des Unterrichts ... 245
8.3.2 Fördereinheit FE06 – Arbeitsunterlagen ... 249

8.4 Auswertungen (weitere Daten) ... 251
8.4.1 Inter-Item-Korrelationsmatrizen ... 251
8.4.2 Kommunalitäten der Faktoranalyse ... 254
8.4.3 Screeplots für die Faktoranalyse ... 256
8.4.4 Graphische DIF-Analyse ... 257

9 ABBILDUNGS-, TABELLEN- UND LITERATURVERZEICHNIS ... 259

9.1 Abbildungsverzeichnis ... 259

9.2 Tabellenverzeichnis ... 261

9.3 Literaturverzeichnis ... 264

1 Einleitung

Operationsverständnis

Die meisten Kinder können sich zu den vier arithmetischen Grundoperationen ohne grössere Schwierigkeiten passende Handlungen, Situationen oder Bilder vorstellen. Für manche Kinder aber sind Rechnungen oftmals nur bedeutungslose Manipulationen mit Ziffern (Radatz, 1991; Ruwisch, 2001; Gleissberg, 2018). Diese Situation ist problematisch, weil die gedankliche Verknüpfung der Rechenoperationen mit passenden Sachsituationen, Handlungen oder Bildern für den Aufbau und die Entwicklung des Verständnisses für die Grundoperationen wesentlich ist (Schütte, 2008; Prediger, 2009). Die Fähigkeit der Kinder, solche gedanklichen Verknüpfungen vorzunehmen, wird in der mathematikdidaktischen Literatur als Indiz und als Gradmesser für Operationsverständnis bezeichnet (Lesh et al., 1987; Huinker, 1993; Bönig, 1995; Gerster & Schultz, 2004; Schäfer, 2005).

Aus Problemen beim Operationsverständnis können sich bei betroffenen Kindern Rechenschwächen entwickeln, die mit gravierenden Folgen für das spätere Lernen von Mathematik verbunden sein können. Darüber hinaus werden Rechenschwächen verschiedentlich auch genannt als mitverantwortliche Ursachen für Schwierigkeiten im späteren Leben, so z.B. beim selbstständigen Umgang mit den persönlichen Finanzen oder auch bei der Berufswahl (Geary, 1994; Moser Opitz, 2005, S. 113 ff.; Akinwunmi & Deutscher, 2014; Gaidoschik, 2016a, S. 15; Rittle-Johnson, 2017). Dementsprechend wird einer frühzeitigen (schulischen) Förderung und Unterstützung beim Aufbau der gedanklichen Verknüpfungen zwischen den Grundoperationen und dazu passenden Sachsituationen, Bildern und Handlungen eine hohe Bedeutung beigemessen.

Struktur dieser Arbeit

In dieser Arbeit werden zuerst in Kapitel 2 die theoretischen Grundlagen des in der Mathematikdidaktik verwendeten Konstrukts Operationsverständnis dargelegt und besprochen. Das Kapitel beginnt mit der Besprechung der Begriffe «Operation», «Verständnis» und «Operationsverständnis». In Kapitel 2.2 werden sodann der Begriff «Darstellungswechsel» sowie das Konzept der Grundvorstellungen und das Konstrukt der Strukturierungsfähigkeit erläutert. Es wird beschrieben, wie mit diesen Begriffen in der Mathematikdidaktik gearbeitet wird und wie sie im Zusammenhang mit dem Operationsverständnis verwendet werden. In Kapitel 2.3 wird beschrieben, wie – unter Verwendung der früher diskutierten Begriffe – die vier Grundrechenarten unter einem Konstrukt «Operationsverständnis» vereint betrachtet werden. Dieses im deutschsprachigen Raum verwendete Konstrukt wird zum Abschluss des Kapitels noch mit dem im englischsprachigen Raum verwendeten Begriff «Operation Sense» verglichen.

In Kapitel 3 wird die Befundlage zum Operationsverständnis dargelegt. Der Begriff wird in seiner Bedeutung beim Rechnenlernen beleuchtet. Der Blick wird auf praktische Schwierig-

keiten im Unterricht und in der Förderarbeit gerichtet. Es werden verschiedene Formen des Unterrichts und der Förderung besprochen.

Im empirischen Teil dieser Arbeit wird zuerst eine Interventionsstudie beschrieben, die in den Schuljahren 2018/19 und 2019/20 im Rahmen eines Forschungsprojekts verschiedener Hochschulen in der deutschsprachigen Schweiz durchgeführt wurde. Die Beschreibung des Gesamtprojekts findet sich in Kapitel 4.1. In Kapitel 4.2 wird der Teil des Projekts beschrieben, der für diese Arbeit besonders relevant ist.

Im Rahmen dieser Arbeit wurden als Interventionen im Mathematikunterricht des zweiten Primarschuljahrs Unterrichtsmaterialien zur Förderung des Operationsverständnisses eingesetzt und deren Wirkung auf den Lernzuwachs der Kinder untersucht. Mit einem neuen Erhebungsinstrument wurde der Lernstand der Kinder zu verschiedenen Zeitpunkten erfasst. Diese Arbeit umfasst in den Kapiteln 4.3 und 4.4 einen Bericht über die Entwicklung der entsprechenden Materialien, deren Ausgestaltung und deren Einsatz im Unterricht bei verschiedenen Schweizer Primarschulklassen. Eine vorbereitende Beschreibung der später eingesetzten Auswertungsmethoden findet sich in Kapitel 4.5.

In den Kapiteln 5 und 6 werden die Ergebnisse der Untersuchungen dargelegt. Das für diese Arbeit eigens entwickelte neue Erhebungsinstrument für das Multiplikationsverständnis (siehe Kapitel 5) wird anhand der Daten aus dem Vor- und dem Nachtest mit verschiedenen analytischen Mitteln eingehend überprüft. Es erfolgt auch eine Diskussion der Gütekriterien für das neue Erhebungsinstrument.

Dieses Erhebungsinstrument wurde zusammen mit einem standardisierten Mathematik-Test für die Längsschnittanalyse im zweiten Schuljahr eingesetzt. Die Ergebnisse dieser Untersuchungen finden sich in Kapitel 6.

Die Arbeit endet mit der Diskussion der Ergebnisse in Kapitel 7.1 und mit einem Ausblick auf mögliche weiterführende Forschungsarbeiten in Kapitel 7.2.

Sprachgebrauch

Mit «diese Arbeit» wird die Arbeit in Zusammenhang mit dem in diesem Dokument beschriebenen Promotionsprojekt bezeichnet. Der Wortlaut «diese Studie» bezieht sich auf die Studie, die dieser Arbeit zugrunde liegt. Die Tätigkeiten, welche in diese Arbeit einfliessen, waren in einem umfangreicheren Forschungsprojekt eingebettet, welches zu weiteren Studien und Promotionsarbeiten geführt hat. Dieses Forschungsprojekt wird in dieser Arbeit auch das «Projekt» oder das «Projekt MALKA» genannt.

Vorbemerkung zur Covid-19-Situation in der Schweiz im Frühjahr 2020

Das Hauptaugenmerk dieser Arbeit liegt auf dem zweiten Beobachtungsjahr, d.h. auf dem Schuljahr 2019/20. Dieses Schuljahr wurde durch die Ausbreitung des Coronavirus SARS-CoV-2 stark tangiert. Durch dieses Coronavirus wurde eine weltweite Pandemie («Covid-19-

Pandemie») ausgelöst, von der auch die Schweiz sowie die Schulen in der Schweiz in hohem Masse betroffen waren. Um die Ausbreitung des Virus in der Schweiz zu verhindern bzw. zu verlangsamen, wurden im Frühjahr 2020 sämtliche Schulen in der Schweiz für acht Wochen (vom 16. März bis zum 11. Mai 2020, inklusive zwei Wochen Frühlingsferien) geschlossen. Diese Schulschliessung betrifft auch die vorliegende Arbeit und hatte für die geplante Studie verschiedene Konsequenzen. Zum einen war die Umsetzung der Interventionen bei Schulschliessung am 16. März 2020 noch nicht abgeschlossen. Bei verschiedenen Klassen war die Umsetzung zwar schon recht weit fortgeschritten. Einige Lehrpersonen hatten ihre Unterrichtstätigkeit aber auch so eingeteilt, dass in den Frühlingsmonaten ein grosser Teil der Interventionen hätte durchgeführt werden sollen. Zudem hat die Schulschliessung die Jahresplanung der Lehrpersonen beeinflusst. Der Stoff, der im Frühling hätte unterrichtet werden sollen, musste mit sehr kurzer Vorbereitungszeit anders (für das «home-schooling») aufbereitet oder aber auf einen späteren Zeitpunkt vertagt werden. Auch der Unterricht nach Wiedereröffnung der Schulen ab dem 11. Mai 2020 war immer noch stark von der unerwarteten Schulschliessung beeinflusst. Schulische Lücken mussten in verschiedenen Fächern bearbeitet, die Gesamtsituation mit den Kindern aufgearbeitet und verschiedenste Schutzmassnahmen beachtet und mit den Kindern eingeübt werden. Durch die Schulschliessung ergaben sich neuartige Verschiedenheiten zwischen Schulklassen, welche nicht vorhersehbar waren und deren Effekte die Messung der Wirkung der eingesetzten Interventionsmaterialien erschweren. Nicht nur die Kinder, sondern auch die Eltern und die Lehrpersonen gingen unterschiedlich mit der Schulschliessung um. Plötzlich spielten auch die Eltern beim Unterricht für die Kinder eine wichtige Rolle. Das «home-schooling» war vor allem dann erfolgreich, wenn die Eltern das Kind optimal unterstützten. Nebst unterschiedlicher Wirksamkeit des «home-schooling» kamen auch nach der Wiedereröffnung der Schulen am 11. Mai 2020 weitere Differenzen zwischen Klassen und Kantonen zum Tragen. In einigen Kantonen wurde z.B. während vier Wochen vorerst in Halbklassen unterrichtet, was für die Kinder eine tiefere Wochenstundenzahl zur Folge hatte. Andere Kantone entschieden sich für die sofortige Wiederaufnahme des Schulbetriebs in ganzen Klassen. In einigen Kantonen wurde zudem für den Abschluss im Frühjahr 2020 auf Zeugnisnoten verzichtet, in anderen wurden trotz Schulschliessung Zeugnisse mit Noten ausgestellt. Die Zeit vom 11. Mai 2020 bis zu den Sommerferien 2020 war zudem geprägt von verschiedenen Unsicherheiten und unterschiedlichen Schutzmassnahmen und Schutzkonzepten in Zusammenhang mit der Covid-19-Situation in der Schweiz.

Trotz diesen unerwarteten Unterschiedlichkeiten zwischen den Klassen und zwischen der Beschulung der einzelnen Kinder konnte diese Arbeit regulär abgeschlossen werden. Die Corona-bedingte Situation führte zwar dazu, dass der Umfang der Testungen geringer ausfiel als ursprünglich geplant. Die für diese Arbeit entscheidenden Erhebungen konnten mit 46 der zu Schuljahrsbeginn verfügbaren 68 Klassen durchgeführt werden. Die anderen Klassen wurden aus unterschiedlichen Gründen nicht getestet. Bei einigen Klassen war der Aufwand für die Lehrperson zu gross oder aber die Schulleitung hatte entschieden, das jeweilige kantonale Schutzkonzept streng auszulegen und den Zugang zum Schulareal für sämtliche externen Personen − und somit auch für die Testleitung des Forschungsprojekts − zu untersagen. In einigen Klassen musste die Lehrperson ersetzt werden, weil die ursprüngliche Lehrperson der Covid-19-«Risikogruppe» angehörte. Einige weitere Klassen wurden nicht getestet, weil sie

die Interventionen erst im Frühjahr hätten umsetzen wollen und die Schulschliessung ihre Planung im Hinblick auf die Interventionen zu stark beeinträchtigt hatte.

Im Zeitraum vom 2. Juni bis zum 10. Juli 2020 konnten dann aber – unter Einsatz eines eigens ausgearbeiteten «Covid-19 Schutzkonzepts» für das Forschungsprojekt – ein Grossteil der geplanten Erhebungen durchgeführt werden. Die Anzahl getesteter Kinder und Klassen erlaubte es, die ursprünglich im Rahmen dieser Arbeit geplanten Berechnungen und Längsschnittanalysen durchzuführen.

Die Ergebnisse dieser Auswertung und dieser Arbeit sind aber immer auch im Lichte der achtwöchigen Schulschliessung zu betrachten. In dieser Arbeit wird deshalb an einigen Stellen auf die Bedeutung der Schulschliessung aufgrund der «Covid-19-Pandemie» für diese Untersuchungen hingewiesen. Es ist nicht zu erwarten, dass sämtliche Ergebnisse unter «regulären» schulischen Bedingungen repliziert werden können. Zu erwarten wäre beispielsweise, dass die im Frühjahr 2020 durchgeführten Erhebungen in einem anderen Schuljahr bei Kindern der 2. Primarschulklasse, die kurz zuvor keine derartige Schulschliessung erlebt haben, bessere Ergebnisse liefern würden. Von den Anpassungen, die im Forschungsprojekt aufgrund der Covid-19-Situation vorgenommen werden mussten, wird in dieser Arbeit an den jeweiligen Stellen laufend berichtet. Selbstverständlich wird diese Thematik auch in der Diskussion der Ergebnisse in Kapitel 7.1 wieder aufgenommen.

2 Operationsverständnis als theoretisches Konstrukt

In Kapitel 2 werden zuerst die Begriffe erläutert, die in der Mathematikdidaktik für die Beschreibung des Konstrukts Operationsverständnis verwendet werden. Die beiden Kapitel 2.1 und 2.2 umfassen solche Begriffsbeschreibungen und Zusammenfassungen von anderen Konzepten und Konstrukten aus der Mathematikdidaktik. In Kapitel 2.3 werden die eingeführten Begriffe dann verwendet, um das Konstrukt Operationsverständnis so zu umschreiben, wie es im späteren Verlauf dieser Arbeit verwendet werden wird.

2.1 Operation, Verständnis und Operationsverständnis

In diesem Abschnitt werden die Begriffe «Operation» und «Verständnis» aus Sicht verschiedener Wissenschaften beleuchtet und beschrieben. Der Begriff der Operation wird dabei aus Sicht der Mathematik und aus Sicht der Mathematikdidaktik betrachtet. Im Anschluss daran wird der Begriff des Verständnisses besprochen und dargelegt. Dieser Begriff wird in der Literatur verschiedener Fachdisziplinen auf sehr unterschiedliche Art verwendet und umschrieben. In diesem Kapitel werden zudem einige der Betrachtungsweisen beschrieben, welche mit dem Lernen von Mathematik in Zusammenhang stehen. Es wird dabei auch erläutert, welche dieser Betrachtungsweisen in dieser Arbeit weiter Verwendung finden wird.

Die beiden Begriffe können sprachlich zum «Operationsverständnis» verbunden werden. Dabei handelt es sich um einen Begriff, der vorwiegend in der Mathematikdidaktik verwendet wird. Seine Bedeutung darzulegen, ist eines der Anliegen dieser Arbeit. In diesem Kapitel erfolgt eine erste Begriffsklärung, die in den Kapiteln 2.2 und 2.3 vertieft und geschärft wird.

2.1.1 Die Operation

Der Begriff der Operation wird in den beiden Wissenschaften Mathematik und Mathematikdidaktik unterschiedlich verwendet. In diesem Kapitel werden die beiden Begriffsverwendungen in ihrer Unterschiedlichkeit und in ihrer Gemeinsamkeit beleuchtet.

Der Begriff der Operation in der Mathematik

In der Mathematik beschreibt der Begriff der *n-stelligen Verknüpfung* eine Funktion, die n Elementen $a_1, a_2, \ldots, a_n$ aus n Mengen $A_1, A_2, \ldots, A_n$ ein Element b aus einer Menge B zuordnet.

Eine *zweistellige* (oder *binäre*) *Verknüpfung* beschreibt den Spezialfall $n = 2$, also eine Funktion, welche zwei Elementen a_1 und a_2 aus zwei Mengen A_1 und A_2 ein Element b aus einer Menge B zuordnet.

Sind alle Mengen dieselben, also im Falle einer zweistelligen Verknüpfung die drei Mengen A_1, A_2 und B identisch (Menge $M = A_1 = A_2 = B$), so wird die Verknüpfung eine *innere Verknüpfung* oder auch eine *Operation* genannt.

Eine *zweistellige Operation* ist also eine Funktion oder Verknüpfungsvorschrift f, die zwei Elementen x und y aus einer Menge M ein Element $f(x, y)$ aus derselben Menge M zuordnet (Bronštejn & Semendjaev, 2006, S. 308; Scheid, 1981, S. 480; Stingl, 2004, S. 27). Anstelle der Funktionsschreibweise $f(x, y)$ kann auch ein Operationszeichen, z.B. $x \circ y$ verwendet werden.

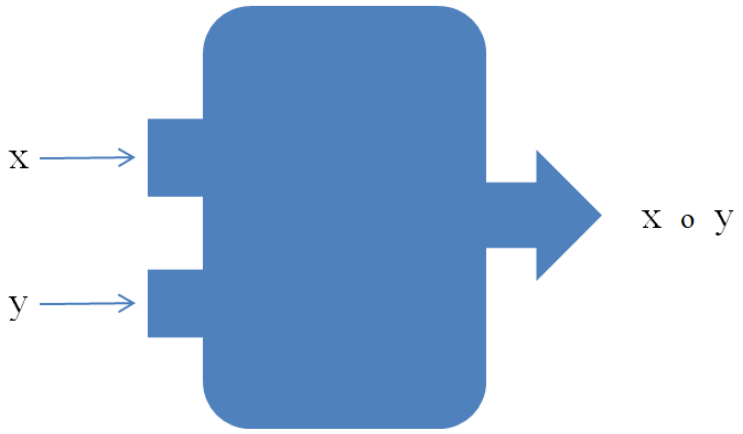


Abbildung 1: Veranschaulichung einer zweistelligen Operation

Beispiele von Operationen in der Menge $\mathbb{N} = \{1, 2, 3, 4, \ldots\}$ der natürlichen Zahlen sind z.B.

$f(a, b) = a + b$, die Summe der beiden Zahlen a und b

$f(a, b) = a \cdot b$, das Produkt der beiden Zahlen a und b

$f(a, b) = \mathrm{ggT}(a, b)$, der grösste gemeinsame Teiler der beiden Zahlen a und b

$f(a, b) = \mathrm{kgV}(a, b)$, das kleinste gemeinsame Vielfache der beiden Zahlen a und b

$f(a, b) = \max(a, b)$, das Maximum der beiden Zahlen a und b

$f(a, b) = a^2 + b^2$

usw.

Keine Operationen in der Menge der natürlichen Zahlen sind die Funktionen:

$f(a, b) = a - b$ und

$f(a, b) = \frac{a}{b}$,

da bei diesen Funktionen der ursprüngliche Zahlenraum der natürlichen Zahlen (durch geeignete Wahl der Werte von a und b) verlassen werden kann (Scheid, 1981, S. 480).

Durch geeignete Zahlenraumerweiterungen werden aber auch diese beiden Funktionen zu Operationen, so ist $f(a, b) = a - b$ eine Operation auf der Menge der ganzen Zahlen $\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$ und $f(a, b) = \frac{a}{b}$ wird zur Operation in der Menge $\mathbb{Q}\backslash\{0\}$ der

rationalen Zahlen ohne die Null und natürlich auch in den Mengen $\mathbb{R}\backslash\{0\}$ und $\mathbb{C}\backslash\{0\}$ der reellen bzw. der komplexen Zahlen ohne die 0.

Durch entsprechende Zahlenraumerweiterungen werden also auch die Subtraktion und die Division zu inneren Verknüpfungen und damit zu Operationen im Sinne der Mathematik.

Beachtenswert an der Definition einer Operation in der Mathematik ist die Tatsache, dass sie sich nicht auf Mengen M von Zahlen beschränkt, sondern auch alle anderen denkbaren Mengen umfasst. Die Definition ist z.B. auch anwendbar auf Matrizen (anstelle von Zahlen), sie könnte die Vereinigung oder die Durchschnittbildung von Teilmengen einer Menge (als Elemente der Potenzmenge einer beliebigen Ausgangsmenge) beschreiben (Bronštejn & Semendjaev, 2006) oder auch auf geometrische Vorgänge wie beispielsweise die Addition von Vektoren in einem beliebigen Vektorraum angewendet werden (Scheid, 1981). In diesem Fall bildet die Menge M den Vektorraum und die zweistellige Operation entspricht der Addition der Vektoren. Auch die Bildung des Vektorprodukts zweier Vektoren aus einem Vektorraum M stellt eine zweistellige Operation im obigen Sinne dar, nicht aber die Berechnung des Skalarprodukts, da dessen Ergebnis zu einem skalaren Wert und damit nicht zu einem Element aus der ursprünglichen Menge M (dem Vektorraum) führt. Das Skalarprodukt stellt aus diesem Grunde eine zweistellige Verknüpfung, nicht aber eine innere Verknüpfung (Operation) dar.

Der Begriff der Operation in der Mathematikdidaktik

In der Mathematikdidaktik ist es – insbesondere in der Literatur, welche sich ausschliesslich oder vorwiegend auf die Primarschulstufe bezieht − üblich, den Begriff der Operation zu verwenden, um damit die vier Grundrechenarten (die vier *Grundoperationen*) zu beschreiben. Der Begriff wird auf diese Weise verwendet, da sich die vier Grundrechenarten bei entsprechender Zahlenraumerweiterung zu Operationen «entwickeln». Die Begriffsverwendung unterscheidet sich aber wie bereits angedeutet von derjenigen aus der Mathematik. Diese Unterschiedlichkeit zeigt sich zum Beispiel daran, dass der Begriff der Operation − genauso wie der damit verbundene Begriff des Operationsverständnisses − auch und sogar vorwiegend für Untersuchungen in den ersten drei Primarschuljahren verwendet wird. In diesen Jahren bewegen sich die Kinder gedanklich noch ausschliesslich in der Menge der natürlichen Zahlen.

In dieser Arbeit wird der Begriff der Operation im Sinne der Mathematikdidaktik verwendet und wird – sofern nicht genauer spezifiziert – die vier Grundrechenarten umfassen.

Die Begriffe «Operation», «Grundoperation» und «Grundrechenart» werden in dieser Arbeit im Folgenden als Synonyme verwendet.

Alle Operationen werden in dieser Arbeit im Zahlenraum der natürlichen Zahlen untersucht. Dieser Zahlenraum bleibt bei den Untersuchungen in dieser Arbeit auch eingeschränkt auf den Hunderterraum, d.h. auf den Zahlenraum der natürlichen Zahlen von 1 bis 100. Es wurde bereits davon berichtet, dass sich Operationsverständnis, das sich in kleineren Zahlenräumen zeigt, nicht notwendigerweise auf grössere Zahlenräume übertragen lässt (Schäfer, 2005; Freesemann, 2014, S. 44). Diese Aussagen wurden im Zusammenhang mit den Operationen

Addition und Subtraktion gemacht. Ob diese Beobachtung auch für die Multiplikation gemacht werden kann, bleibt dabei offen.

Die Addition und die Subtraktion werden die Grundrechenarten oder die Operationen *erster Stufe* genannt. Die Multiplikation und die Division bilden die Grundrechenarten bzw. die Operationen *zweiter Stufe* (Oehl, 1962, S. 58).

2.1.2 Das Verständnis

Bevor in dieser Arbeit der Begriff Operationsverständnis beschrieben wird, werden verschiedene Ansätze zur Beschreibung der Begriffe «Verständnis» und «Verstehen» im Zusammenhang mit dem Erlernen von Mathematik beleuchtet.

Die möglichen Bedeutungen der Worte «Verstehen» und «Verständnis» sind a priori ausgesprochen vielfältig. Sie können in Bezug auf verschiedene Gegenstandsbereiche z.B. auf die folgende Art voneinander abgegrenzt werden:

«Man kann:

(1) Menschen, Handlungen, Situationen verstehen, etwa die Motive, Ziele usw. der Beteiligten;
(2) Äusserungen medial bzw. formal verstehen (Lautstärke, Fremdsprache);
(3) Äusserungen inhaltlich verstehen (bei Mitteilungen, Texten, Wendungen, Wörtern, Symbolen, …);
(4) Schliesslich einen Sachverhalt verstehen (Sachverstand)»

(Bender, 1991, S. 53 f.)

In dieser Arbeit wird mit «Verstehen» in der obigen Einteilung in vier Kategorien die Bedeutung (4) gemeint. Auch im Begriff «Operationsverständnis» ist die Bedeutung (4) gemeint, gegebenenfalls ergänzt durch Aspekte der Bedeutung (3), falls z.B. Symbole (Zahlen oder Operationszeichen) den Kindern noch unklar oder «unverstanden» sein sollten.

Vielfalt der Ansätze zur Beschreibung von «Verstehen»

Die Ansätze zur Schärfung und klaren Erfassung der Begrifflichkeiten unterscheiden sich zwischen verschiedenen Wissenschaften. Während in der Psychologie Verständnis unabhängig vom Lerngegenstand beschrieben wird, stehen in den verschiedenen Fachdidaktiken die Prozesse des Lernens und des Verstehens meistens in klarem Bezug zum jeweiligen Fach. Innerhalb des Fachs wiederum sind sie meist auch bezogen auf einen jeweiligen Lerngegenstand. Beispielsweise wird in Drollinger-Vetter (2011) das Verstehen mit Bezug auf Aebli (2011) als «inhaltsabhängig» angesehen. Es wird dabei postuliert, das Verstehen sei «in den Begriffen der Sache» zu beschreiben. Auch Weigand (2015) schlägt für das Verstehen eine «domänenspezifische» Untersuchung vor. Zentrale Begriffe für das Verständnis eines einzel-

nen Gebiets der Mathematik unterscheiden sich also auch innerhalb der Mathematik von Gebiet zu Gebiet.

Lernprozesse bzw. Verstehens-Prozesse finden also nicht nur in fachspezifisch unterschiedlicher Art und Weise statt, sondern unterscheiden sich auch innerhalb eines Fachs von Gebiet zu Gebiet. Damit ist es notwendig, Verstehens-Prozesse sehr eingehend und bezogen auf das jeweilige Themenfeld zu untersuchen. Übergreifende und allgemeine Beschreibungen eines Verstehens-Prozesses werden nach Ansicht der Fachdidaktiker der Sache und dem Inhalt häufig nicht gerecht. Die Mathematikdidaktik wie auch die Fachdidaktiken anderer Fächer begründen u.a. mit dieser Beobachtung ihren Anspruch, als eigenständige Wissenschaften zu gelten (Bigalke, 1974).

Innerhalb der Mathematikdidaktik illustriert das Konzept der Grundvorstellungen die Tatsache, dass verschiedene Lerngegenstände sehr unterschiedliche innere Vorstellungen des Lernenden benötigen. Nur so können die Lerngegenstände in angemessener Weise «verstanden» werden. Angewandt auf die Grundrechenart *Multiplikation* kann dies bedeuten, dass ein Kind zeitlich-sukzessive Beispiele («Peter geht dreimal in die Küche und holt dort jedes Mal vier Teller») der Multiplikation bzw. einer passenden Malaufgabe (einem Multiplikationsterm) zuordnen kann, während diese Zuordnung bei kombinatorischen Beispielen («Peter hat drei Pullover und vier Hosen. Auf wie viele Arten kann er sich anziehen?») noch nicht gelingt und dieses zweite Beispiel damit im Sinne des Operationsverständnisses unverstanden bleibt.

Es wird in der Literatur zudem darauf hingewiesen, dass selbst innerhalb der Mathematikdidaktik die Begriffe «Verstehen» und «Verständnis» sehr unterschiedlich gebraucht werden und mit verschiedener Bedeutung versehen sind (Weigand, 2015). Versuche, die Begriffsverwendung zu vereinheitlichen, gelingen meist nicht, weil die Inhalte zu vielschichtig und zu unterschiedlich seien (Bönig, 1995).

Häufig kommt es aber in der mathematikdidaktischen Literatur vor, dass der Begriff des «Verstehens» in Zusammenhang gebracht wird mit dem Begriff von Vorstellungen, die dann zum Teil «innere Vorstellungen», «Modellvorstellungen» oder eben auch Grundvorstellungen genannt werden. Weigand (2015) hält zum Beispiel fest, dass zum «Verstehen» mehr gehöre als die reine Kenntnis einer Definition. Für das Verständnis eines Begriffs seien auch Vorstellungen zu den Eigenschaften und den Merkmalen des Begriffs notwendig («Begriffsinhalt»). Der Lernende sollte weiter einen Überblick gewinnen über die Gesamtheit aller Objekte, die mit dem entsprechenden Begriff bezeichnet werden («Begriffsumfang»). Zudem sollte der Lernende in der Lage sein, die Beziehungen des Begriffs zu anderen Begriffen zu erfassen («Begriffsnetz»). Als Letztes wird die Anwendung und der angemessene Umgang mit dem Begriff genannt («Begriffsanwendungen»). In all diesen Aspekten ist der Lernende gefordert, wenn er ein umfassendes Begriffsverständnis aufbauen möchte. Aufgabe der Lehrperson ist es jeweils, die lernende Person beim Aufbau eines solchen Verständnisses zu unterstützen. Der Autor beschreibt weiter die Notwendigkeit, dass der Lernende für den Aufbau von Verständnis «mentale Objekte» aufbaut, und verweist auf den Begriff der Grundvorstellungen (Weigand, 2015, S. 260 ff.).

Ansätze aus der Psychologie mit Bezug zur Mathematikdidaktik

Zuweilen werden auch Ansätze aus der Psychologie innerhalb der Fachdidaktiken angewandt, um den Prozess des Verstehens eines Sachverhalts zu beschreiben, so z.B. in Drollinger-Vetter (2011). Der Einbezug solcher Ansätze ist innerhalb der Mathematikdidaktik aber umstritten und wird in diesem Fachgebiet rege diskutiert (Lorenz, 2017; Obersteiner et al., 2018; Griesel et al., 2019).

Sowohl in der Psychologie als auch in der Mathematikdidaktik kann der Begriff des «Verständnisses» bzw. der Vorgang des «Verstehens» unterschiedlich gefasst und beschrieben werden. Einige solche Ansätze rund um den Begriff des «Verstehens» − alle zumindest mit Bezug zur Mathematikdidaktik – werden im Folgenden kurz beleuchtet.

Viele Autoren und Didaktiker stimmen darin überein, dass der Prozess des Lernens, welcher auch den Aufbau von Verständnis umfasst, ein konstruktivistischer Prozess ist. Flanders (2014) legt mit Bezug auf Hiebert & Carpenter (1992) zusammengefasst dar, der Konstruktivismus gehe davon aus, dass die Lernenden ihr eigenes Verständnis eigenständig aufbauen und somit konstruieren. Die lernende Person baut sich dabei selbst ein Netzwerk von gedanklichen Verknüpfungen auf. Stern (2002) postuliert, dass neue kognitive Kompetenzen das Ergebnis einer aktiven geistigen Konstruktion sind. Sie entstehen nicht durch eine passive Übertragung vom Lehrenden zum Lernenden. Verstehen wird angesehen als Aufbau und Verbindungsausbau in einem gedanklichen Netzwerk und ein Sachverhalt («fact») gilt als verstanden, wenn er Teil eines solchen gedanklichen Netzwerks ist. Das Ausmass des Verstehens hängt ab von der Dichte und vom Ausbau dieses gedanklichen Netzwerks. Hiebert & Carpenter (1992) verweisen darauf, dass diese Auffassung von Verständnis eine lange mathematikdidaktische Tradition habe, die sich durch das ganze 20. Jahrhundert hindurch ziehe. Sie unterstützen den Gebrauch von Materialien im schulischen Unterricht als Mittel zum Zweck der Förderung des Verständnisses für Mathematik. Sie weisen aber auch auf deren Gefahren hin und begründen die Bedeutung des Darstellungswechsels für das Verständnis. Es wird zudem hingewiesen auf die Bedeutung des Darstellungswechsels besonders im Zusammenhang mit mathematischen Symbolen wie Zahlen oder Rechentermen. Für das Verständnis der Schulkinder für solche mathematische Symbole sind Verbindungen in einem gedanklichen Netzwerk und der Wechsel zu anderen Darstellungsarten notwendig (Hiebert & Carpenter, 1992, S. 66 ff.).

Die Entwicklung von Verständnis wird als bedeutungsvoll diskutiert, denn dieses ist mit vielerlei vorteilhaften Effekten verbunden. Die lernende Person erinnert sich leichter an Dinge, die sie verstanden hat. Ihr Gedächtnis wird entlastet, wenn sie einen Sachverhalt versteht und diesen nicht auswendig lernen muss. Auch fallen Transferleistungen wie die Anwendung des Gelernten dem Lernenden leichter, wenn er etwas verstanden hat und dieses neue Wissen in ein entsprechendes Netzwerk eingebettet ist.

Verstehen wird also beschrieben als ein Prozess, bei dem gedankliche Verbindungen hergestellt werden. Dies kann entweder zwischen bereits bekannten Informationsteilen oder auch zwischen Bekanntem und neuer Information stattfinden. Dieser Verbindungsaufbau und die

Herstellung von Verknüpfungen u.a. auch zwischen inneren Bildern und Darstellungen führen zum Begriff des Darstellungswechsels.

Verständnis entwickelt sich somit, wenn ein gedankliches Netzwerk dichter, besser vernetzt und verbunden wird und die bestehenden Verbindungen gestärkt werden (Flanders, 2014). Diese konstruktivistischen Gedanken gehen zurück bis zu Piaget, der formuliert hat, dass «nur die aktive Auseinandersetzung eines Individuums mit Gegebenheiten seiner Umwelt» dazu führt, dass Bedeutung konstruiert wird und Lernen stattfindet (Gerlach, 2007, S. 104).

In manchen Arbeiten wurde dabei mit starkem Bezug zu Piaget der Bezug zur Handlung des Lernenden besonders hervorgehoben. Dörfler beispielsweise betont den Zusammenhang zwischen mathematischen Begriffen und dem Handeln. Der Autor sieht die praktischen Handlungen als «Ursprung und als Mittel zur Herausbildung mathematischer Begriffe und Operationen» und beschreibt in seiner Arbeit das «Lernen von (grundlegender) Mathematik in und durch Handlungen». Er arbeitet dabei mit den Begriffen «Frames» und «kognitiven Schemata» und sieht Wissenserwerb als individuelle Konstruktion, die aber auch von aussen durch Material und durch die Lehrpersonen gesteuert wird. Der sprachlichen Ebene schreibt der Autor – genauso wie auch die jüngere Literatur − bei solchen Prozessen des Lernens und der Entwicklung von Verstehen besonders hohe Bedeutung zu (Dörfler, 1988).

Auch in der Kognitionspsychologie werden Lernprozesse bzw. Prozesse der Entwicklung und des Aufbaus von Verständnis unter dem Aspekt des «Strukturaufbaus» beschrieben. Diese Betrachtungsweise wird auch in mathematikdidaktischen Zusammenhängen angewandt und beleuchtet. Drollinger-Vetter (2011) beschreibt «Verstehen» mit Bezug auf Aebli (1994) als «Aufbau von beweglichen und transparenten kognitiven Strukturen». Verstehens-Prozesse können als Prozesse der Begriffsbildung verstanden werden. Die Begriffsbildung wiederum sei ein «Prozess des Aufbaus und der Objektivierung von Strukturen». Das Verstehen einer Sache wird in Zusammenhang mit einem Netz von Beziehungen, die der Sache innewohnen, angesehen. Dieses Netz «transparent» zu sehen und sich darin geistig bewegen zu können, ist laut den Autoren die eigentliche Bedeutung von «Verstehen».

Ein weiterer Ansatz aus der Kognitionspsychologie, Verstehens- und Lernprozesse zu beschreiben, verwendet die Einteilung von Wissen in konzeptuelles und prozedurales Wissen. Diese Einteilung wurde schon in Hiebert & Carpenter (1992) und in früheren Werken ausgeführt, aber auch in neuerer Zeit vermehrt wieder aufgenommen und diskutiert.

Konzeptuelles Wissen umschreibt eine Einsicht, eine Kenntnis der Bedeutung der verwendeten Begriffe und Verständnis für die Verfahren und Begriffe. Im Gegensatz dazu beschreibt prozedurales Wissen die Fähigkeit, Abläufe («Prozeduren») auszuführen. Ergänzt werden diese beiden Begriffe durch den Aspekt der prozeduralen Flexibilität. Dieser Begriff bezeichnet die erweiterte Fähigkeit, Abläufe auszuführen und zu entscheiden, welcher dieser Abläufe der geeignetste bzw. effizienteste in der jeweiligen Situation ist.

Anhand von Beispielen aus der Mathematik wird dargelegt, wie konzeptuelles und prozedurales Wissen aufeinander aufbauen und einander gegenseitig ergänzen und zu prozeduraler Flexibilität des Lernenden führen. Prozedurale Flexibilität wird u.a. auch durch Vergleichsprozesse gefördert. Verstehen wird in diesem Ansatz beschrieben als Ergebnis eines geeignet

verzahnten Aufbaus von konzeptuellem und prozeduralem Wissen beim Lernenden (A. J. Baroody et al., 2007; Rittle-Johnson & Star, 2007, 2011; Durkin et al., 2017; Rittle-Johnson, 2017).

Fortschritte in den beiden Aspekten des konzeptuellen und prozeduralen Wissens sind eng verzahnt und finden in iterativer Weise und in enger Wechselwirkung miteinander statt. Fortschritte in einem der beiden Aspekte fördern und beschleunigen Fortschritte im anderen. Bezogen auf den Lernprozess und vor allem auch auf die Ausgestaltung von Unterricht ist dies eine wichtige Erkenntnis mit praktischen Konsequenzen. Unterricht, der vorwiegend oder sogar ausschliesslich einen der beiden Aspekte beachtet, bleibt mangelhaft. Rittle-Johnson (2017) beschreibt Lerntechniken, welche mehr als nur einen dieser Aspekte von Wissen bedienen. Zu diesen Lerntechniken gehört insbesondere auch die Vergleichsarbeit, z.B. zwischen verschiedenen Lösungswegen (Durkin et al., 2017; Rittle-Johnson & Star, 2011).

Der Blick der Semiotik auf den Verstehens-Begriff

In der semiotischen Schule werden Prozesse des Verstehens mittels Zeichen oder Signalen beschrieben. Diese Schule geht zurück auf die griechischen Gelehrten Platon und Aristoteles. Vertreter dieser Schule in der neueren Zeit sind Charles Sanders Peirce und Ferdinand de Saussure (Dörfler, 2019). Beschrieben wird in dieser Betrachtungsweise der Dreiklang zwischen dem Bezeichneten (dem Gegenstand), dem Zeichen (dem gesprochenen Wort) und der Vorstellung, die sich dabei «in der Seele» ergibt und entwickelt. De Saussure unterscheidet zwischen dem Bezeichneten («signifié») und dem Bezeichnenden («signifiant»). Die Mathematikdidaktik bezieht sich aber stärker auf Peirce, der im Gegensatz zu De Saussure stärker mit einer Dreiteilung in Objekt, Representamen und Interpretant arbeitet. Aufgegriffen wird die semiotische Betrachtungsweise auch bei der Beschreibung von Darstellungswechseln, der Einteilung in verschiedene Register und den Wechseln zwischen oder innerhalb von solchen Registern (Duval, 2006). Wechsel zwischen Registern werden dabei «conversions», Wechsel der Betrachtungsweise bzw. der Darstellungsart innerhalb eines Registers werden «treatments» genannt. Die Bandbreite der Anwendungen dieser Theorien in der Mathematikdidaktik ist gross und ihre Reichweite erstreckt sich auch auf Themen, die ausserhalb des Primarschulunterrichts behandelt werden. Entsprechend werden beispielsweise auch Verstehens-Prozesse in der Algebra oder auch die komplexen Zahlen bzw. abstraktere Objekte der höheren Algebra unter semiotischer Perspektive beleuchtet. Immer wieder werden dabei multiple Darstellungen und der Wechsel zwischen solchen Darstellungen diskutiert. An der Fähigkeit, zwischen Darstellungen zu wechseln, zeigt sich letztendlich das Verstehen der Lernenden (Dörfler, 2015; Duval, 2006). Diese Fähigkeit zum Repräsentationswechsel und die Gewandtheit bei diesem Wechsel («fluency») wird u.a. auch in Flanders (2014) als Zeichen von Verständnis formuliert.

Semiotiker sehen im Lernen von Mathematik nicht nur in der elementaren Arithmetik, sondern auf jedem Niveau «Zeichenspiele» oder «Zeichenpraxis» (Rosch, 2015; Dörfler, 2019). Das Zeichenspiel der Arithmetik könnte zwar ebenso zur reinen «Unterhaltung» nach den geltenden Regeln erfolgen, sollte aber − wenn möglich − für den Lernenden auch einen Sinn

ergeben. Diesen Sinn erhält es, wenn den Zeichen und den Regeln Bedeutung zukommt (Dörfler, 2019). Angewandt auf die Arithmetik der vier Grundrechenarten kann dies so formuliert werden, dass das Zahlenspiel Sinn bekommt, wenn es vom rechnenden Kind mit Operationsverständnis betrieben wird. Ansonsten bleibt es beim Zahlenspiel bzw. bei der in der Einleitung erwähnten und in der Literatur beschriebenen «bedeutungslosen Manipulation mit Ziffern». Thomas (2008) hält mit Bezug auf Peirce fest, dass ein Symbol ein Zeichen sei, das durch dessen Gebrauch erst eine Bedeutung erhält. Diese Bedeutung gilt es zuerst zu erschliessen; eine Verwendung von Zeichen, ohne deren Bedeutung zu kennen, bleibt ein sinnloser Vorgang. Rechnen ohne Operationsverständnis ist *ein* Beispiel eines solchen Vorgangs, der wenig Sinn macht.

Zeichen können unterschiedlich verwendet und mit Sinn gefüllt werden. Laakmann (2013) beschreibt, wie Begriffe und Zeichen entweder individuell oder auch konventionell mit Bedeutung ausgefüllt sein können.

Semiotiker präsentieren mit ihrer Sicht auf die Mathematik eine alternative Betrachtungsweise, deren Fokus auf den Zeichentätigkeiten liegt. Gemäss ihrer eigenen Sicht wird die Mathematik dadurch «humanisiert». Sie wird durch diese Betrachtungsweise zu einer sozialeren und beobachtbaren Tätigkeit (Dörfler, 2019).

«Kommognitionale» Betrachtung des Verstehens-Begriffs

Eine weitere alternative Sicht innerhalb der Mathematikdidaktik liefert die «kommognitionale» Erforschung von Lernprozessen. In dieser Schule werden die beiden Vorgänge des Denkens («thinking», «cognition») und der Kommunikation («communication») in enger Verbindung zueinander gesehen. Die beiden Vorgänge werden als Manifestationen desselben Prozesses beschrieben. Die Wortkombination in englischer Sprache führt zum neuen Begriff «commognition», welcher dieser Sichtweise ihren Namen gibt (Sfard, 2012a, 2012b).

Beschrieben werden in dieser Betrachtungsweise Routinen («routines»), welche sowohl in «practical» und «discursive» als auch in «process-oriented» und «product-oriented» eingeteilt werden können (Lavie et al., 2018).

Denken wird als eine Art Selbstgespräch betrachtet. Mathematik zu betreiben wird als Gesprächsführung bzw. als Diskurs angesehen. Auf diesen Grundvoraussetzungen aufbauend wird dann der Lernprozess bzw. das «Verstehen» von Mathematik beschrieben (Sfard, 2012b).

Die Untersuchungen münden in der Erkenntnis, dass Lernprozesse als Prozesse der «Routinisierung» betrachtet werden können und damit eigentlich keine Notwendigkeit mehr besteht, den Begriff des «Verstehens» überhaupt zu verwenden (Lavie et al., 2018, S. 20 f.).

Diese Forschungsrichtung soll aufzeigen, wie weitreichend und unterschiedlich die Erfassung des Begriffs «Verstehens» auch in der Mathematikdidaktik ist. In dieser Schule wird der Begriff des Verstehens, welchen andere Forscher zu beschreiben und zu fassen versuchen, aufgrund der getätigten Überlegung als unnötig betrachtet.

2.1.3 Das Operationsverständnis

In diesem Kapitel erfolgt eine Beschreibung des Begriffs Operationsverständnis unter Verwendung der bisher besprochenen Begriffe. Das Kapitel liefert damit auch eine Annäherung an das Konstrukt Operationsverständnis, das in Kapitel 2.3 beschrieben wird.

Grundverständnis

In der mathematikdidaktischen Literatur wird der Begriff Operationsverständnis verwendet, um ein Grundverständnis der Kinder für die vier arithmetischen Grundrechenarten zu bezeichnen. Der Begriff der Operation wird dabei verwendet für die in Kapitel 2.1.1 dargelegten Operationen aus der Mathematikdidaktik. Es geht in dieser Arbeit also um das Verständnis der Kinder für die vier Grundrechenarten, also um die vier Grundoperationen der Addition, Subtraktion, Multiplikation und Division.

Die vier Grundrechenarten werden bei dieser Begriffsverwendung entweder gemeinsam betrachtet oder aber getrennt untersucht. Bei der getrennten Untersuchung ist dann jeweils vom Operationsverständnis für die jeweilige Grundrechenart, z.B. vom «Operationsverständnis der Multiplikation» oder auch vom «Multiplikationsverständnis» die Rede. Dabei wird meist selbstredend davon ausgegangen, dass das Konstrukt Operationsverständnis in vier einzelne sinnvolle Konstrukte für die vier Grundrechenarten zerfällt. Von dieser Annahme wird auch in der vorliegenden Arbeit ausgegangen. Für den Zweck dieser Arbeit wird das Konstrukt Operationsverständnis so verwendet, wie es in der mathematikdidaktischen Literatur im Moment gebräuchlich ist, d.h. als Oberbegriff bzw. als Sammelbegriff für das Operationsverständnis sämtlicher vier einzelner Grundrechenarten.

Zuweilen wird das Multiplikationsverständnis auch mit anderen Begriffen oder mit einer Umschreibung wie z.B. dem «umfassenden multiplikativen Verständnis» in der Literatur behandelt (Ruwisch, 2001). In dieser Arbeit werden solche Formulierungen ohne weitere Rücksicht auf allfällige sprachliche Nuancen als Synonyme verwendet.

Die vier Grundrechenarten werden im schulischen Unterricht sowohl in der Schweiz als auch in Deutschland in den ersten drei Primarschuljahren eingeführt und behandelt. Sie gehören für alle Kinder – insbesondere auch für jene mit besonderem Förderbedarf − zu den zentralen Lernzielen der ersten Schuljahre (Moser Opitz & Schmassmann, 2007). Verschaffel et al. (2007) bezeichnen die Grundrechenarten («operations») als dominanten Teil der mathematischen Grundausbildung («a dominant part of elementary mathematics education»).

Dabei geht es bei der Einführung in die Grundrechenart typischerweise um ein erstes grundlegendes Verständnis der Kinder für die jeweilige Grundrechenart. Das Einüben (bzw. das Auswendiglernen oder Abrufen) von Rechenaufgaben und deren Ergebnissen (z.B. das Einüben des Einmaleins für die Multiplikation) folgen im Unterricht zu späterem Zeitpunkt. Dieses Einüben erfolgt erst, nachdem das inhaltliche Denken ermöglicht und das Grundverständnis für die Bedeutung der Grundrechenart gefestigt bzw. die notwendigen Grundvorstellungen aufgebaut worden sind (Smith & Smith, 2006; Prediger, 2009; Gaidoschik, 2016a, 2016b).

Der in dieser Arbeit verwendete Begriff «Operationsverständnis» beschreibt dieses frühe Grundverständnis für die Bedeutung der jeweiligen Grundrechenart. Er beschreibt hingegen nicht die Fähigkeit, Rechnungen zu lösen und beispielsweise Ergebnisse von Multiplikationsaufgaben aus dem Gedächtnis abzurufen.

Die Bedeutung von Operationsverständnis ändert sich in höheren Klassen nicht. Wissenschaftliche Untersuchungen des Operationsverständnisses zu späterem Zeitpunkt (in späteren Primarschuljahren oder in der Sekundarstufe) betreffen daher häufig Mängel im Grundverständnis, das in den ersten drei Schuljahren hätte aufgebaut werden sollen. Die inhaltliche Bedeutung des Begriffs «Operationsverständnis» ist in diesen Arbeiten dieselbe wie bei den Untersuchungen in den frühen Primarschuljahren (Moser Opitz, 2005; Ehlert et al., 2013).

In der Literatur wird auch darauf hingewiesen, dass es genau das tiefliegende Verständnis für die eigentlichen wesentlichen Inhalte der Mathematik ist (beispielsweise das Operationsverständnis), welches die spätere Mathematikleistung voraussagt. Weniger aussagekräftige Prädiktoren sind die prozeduralen Fertigkeiten oder der schnelle Abruf von auswendig gelernten Ergebnissen zu einzelnen Multiplikationsaufgaben (Moser Opitz, 2013; Smith & Smith, 2006). Der zu starke Fokus des Unterrichts auf prozedurale Aspekte im Gegensatz zu konzeptuellem Verständnis wird in der Literatur auch als Erklärung für Schwierigkeiten der Kinder im Umgang mit der Multiplikation herangezogen (Hurst & Hurrel, 2016). Die Autoren stellen fest, dass die Leichtigkeit, mit der die Kinder Ergebnisse von Multiplikationsaufgaben abrufen, keinerlei Indiz für ihr Multiplikationsverständnis sei.

In umfangreicheren empirischen Untersuchungen werden hingegen häufig Rechenleistungen und prozedurale Aspekte geprüft. Als Erklärung hierfür dient sicherlich die Tatsache, dass sich das Erfassen des Grundverständnisses sehr viel schwieriger gestaltet. In Stanat et al. (2017) werden bei den verschiedenen Kompetenzstufen, die für das Fach Mathematik erfasst werden, schon auf tiefer Stufe mehrheitlich Rechenleistungen wie z.B. das Lösen von Einmaleins-Rechnungen genannt. Es bleibt also unklar, ob und in welchem Umfang bei dieser Untersuchung Operationsverständnis gemessen wurde. In vielen weiteren Studien verhält sich dies nicht anders.

Der Begriff «Operationsverständnis» – so wie er in der Mathematikdidaktik verwendet wird – bezeichnet also ein grundlegendes Verständnis für die zentralen und wichtigsten Aspekte und Bedeutungen der jeweiligen Grundrechenart. Er wird nicht dazu verwendet, Nuancen in Fertigkeiten mit den Grundrechenarten zu beschreiben, wenn dieses «grundlegende» Verständnis einmal gegeben ist. So ist die Unterscheidung zwischen Personen, die in der späteren Schulzeit oder auch im Erwachsenenalter unterschiedlich flink oder auch unterschiedlich geschickt die Grundrechenarten anwenden, keine Frage des Operationsverständnisses im Sinne der Mathematikdidaktik.

Verwendung von «Operationsverständnis» in der Mathematikdidaktik

Verschiedene Konstrukte aus der Psychologie (z.B. «Extrovertiertheit» oder «Intelligenz») erfassen Begriffe, deren Verwendungen im Alltag mit vielen Nuancen versehen sind. Für die Erfassung solcher Konstrukte im Rahmen von Erhebungsinstrumenten mittels Tests oder Fragebogen sind jeweils Operationalisierungen notwendig, die mit Vereinfachungen des zu erfassenden Begriffs verbunden sind (Bortz & Döring, 2016, S. 221 ff.).

Im Gegensatz zu den obigen Beispielen aus der Psychologie handelt es sich beim Operationsverständnis aber um einen Begriff, der bereits inhaltlich «eng» gefasst ist.

Operationsverständnis beschreibt die grundlegenden Aspekte des Verstehens, die bei einem Kind in der Primarschule für die Grundrechenarten aufgebaut werden sollten. Operationsverständnis beschreibt nicht einen raffinierteren oder «verständigeren» Umgang mit der Operation zu späterem Zeitpunkt. Eine erwachsene Person, welche beispielsweise eine Multiplikationsaufgabe wie $16 \cdot 14$ flinker löst (z.B. durch Abrufen des Ergebnisses von 15^2 und eine darauffolgende Differenzbildung $225 - 1$) als eine andere Person, mag verständiger, effizienter, flexibler oder geschickter mit dieser Grundrechenart umgehen. Sie zeichnet sich im umgangssprachlichen Sinne des Wortes also auch durch ein höheres Verständnis für die Operation aus. Sie zeigt aber – so die Begriffsverwendung des Konstrukts in der Mathematikdidaktik – gegenüber der anderen Person kein höheres oder stärker ausgeprägtes Operationsverständnis.

In dieser Hinsicht unterscheidet sich Operationsverständnis vom englischen Begriff «operation sense», denn die denkbare verbleibende Weite des Begriffs kommt in der englischen Sprache viel deutlicher zum Ausdruck. In manchen Arbeiten aus dem englischsprachigen Raum wird sogar ausdrücklich auf diese Weite des Konstrukts hingewiesen (Slavit, 1999, S. 258) oder die Beschreibung des Konstrukts wird mit Formulierungen wie «operation sense enables students to apply and use operations with meaning and flexibility» (Huinker, 1993, S. 80) bewusst weit gefasst.

In beiden Sprachräumen findet aber das in der Mathematik weit verbreitete Konzept des Wechsels zwischen Darstellungen Anwendung. In der Fachliteratur besteht ein Konsens dahingehend, dass der Darstellungswechsel einen zentralen, wesentlichen Aspekt des Konstrukts Operationsverständnis darstellt.

Etwas weniger Einheit besteht hinsichtlich der Frage, welche weiteren Aspekte das Konstrukt umfassen soll. In der Literatur sind einerseits «Definitionen» zu finden, die besagen, das Operationsverständnis sei gleichzusetzen mit der Fähigkeit zum Darstellungswechsel, und damit auch suggerieren, die Definition des Konstrukts würde bzw. sollte sich auf diese Fähigkeit beschränken (Gerster & Schultz, 2004; Freesemann, 2014, S. 111). In anderen Arbeiten wird etwas vorsichtiger formuliert, der Darstellungswechsel sei ein Indiz oder Indikator oder ein zentraler Aspekt von Operationsverständnis (Moser Opitz, 2005, S. 123; Kuhnke, 2013, S. 29). Es wird festgehalten, der Darstellungswechsel stehe in enger Beziehung zum Operationsverständnis, und Schwierigkeiten beim Operationsverständnis würden sich beim Hin- und Herwechseln zwischen unterschiedlichen Repräsentationsformen äussern bzw. seien durch

Mängel bei den Fähigkeiten des Hin- und Herwechselns (des «Mathematisierens») bedingt (Moser Opitz, 2009, S. 37, 2013, S. 205; Freesemann & Breucker, 2014, S. 8).

In der deutschsprachigen Literatur wird ausser auf den Darstellungswechsel auch auf die Bedeutung der Grundvorstellungen hingewiesen. Dies bedeutet, dass der Darstellungswechsel den Kindern nicht nur für einzelne Aufgabentypen, sondern auch für möglichst viele Aufgabentypen gelingen sollte.

Zudem wird auch im deutschsprachigen Raum Flexibilität im Umgang mit den Grundrechenarten als Indiz für ein angemessenes Verständnis angesehen. Diesem Aspekt der Formulierung von Huinker (1993) wird in der Mathematikdidaktik ebenfalls Bedeutung beigemessen. Es geht dabei aber meist nicht um Flexibilität im Sinne des geschickten Rechnens, sondern eher um gedankliche Flexibilität bei den Übersetzungsprozessen zwischen den einzelnen Darstellungen. Der Begriff «Operationsverständnis» wird aber nach wie vor lediglich verwendet, um ein Grundverständnis für die jeweilige Grundrechenart zu beschreiben. Um dies am Beispiel der Multiplikation zu präzisieren: Gemeint ist also ein Erkennen von multiplikativen Strukturen in Alltagsbildern, in Alltagssituationen oder in Punktebildern, nicht aber Raffinesse oder Effizienz beim Anwenden des Einmaleins oder beim Lösen von Malrechnungen in höheren Zahlenräumen. Flexibilität spielt hier insofern mit, als ein flexibler Umgang z.B. mit Punktebildern eine Übersetzung in einen Term erleichtern kann. Ein derartiges Erkennen von Strukturen (z.B. am Hunderterpunktefeld) und das Arbeiten mit Verwandtschaften zwischen einzelnen Multiplikationsaufgaben können zum Operationsverständnis beitragen.

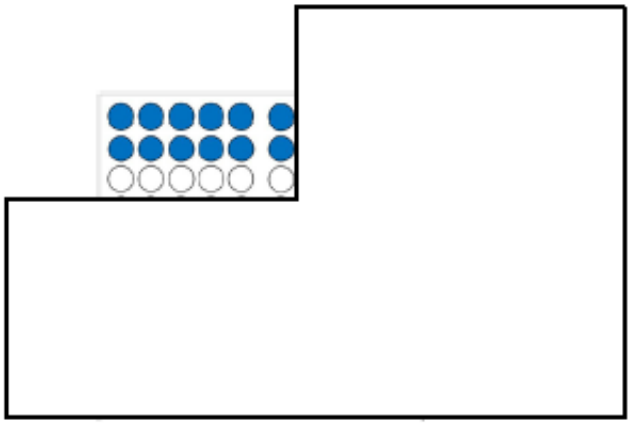


Abbildung 2: Vergrösserung eines 2 · 6 Punktefelds durch Hinzunahme neuer Punkte; aus Florin et al. (2019)

Ein Erkennen eines 3 · 6 Punktefelds am Hunderterfeld mit Hilfe des Wissens, dass dieses aus dem 2 · 6 Punktefeld durch Hinzunahme einer Zeile hervorgeht, kann unter dem Gesichtspunkt des Operationsverständnisses betrachtet werden. Entsprechendes Arbeiten der Kinder mit Strukturen kann der Förderung von Operationsverständnis dienen.

Für den Zweck dieser Arbeit wird Operationsverständnis somit verstanden als die Fähigkeit, zwischen der mathematischen Symbolik und realen Darstellungen (in bildlicher oder textba-

sierter Form) hin und her zu wechseln (also den Darstellungswechsel zwischen Mathematik und Realität zu vollziehen). Einbezogen wird die Notwendigkeit, die verschiedenen Grundvorstellungen der einzelnen Operationen zu bedienen. Zudem wird für ein «ausreichendes» Operationsverständnis eine angemessene Flexibilität im Umgang mit Strukturmerkmalen bzw. der Verwandtschaft verschiedener mathematischer Terme verlangt.

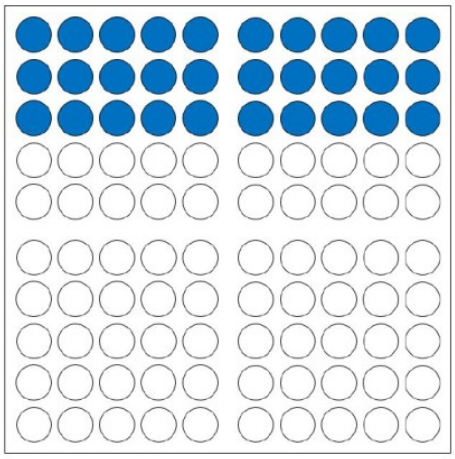


Abbildung 3: Punktefeld zu den Malaufgaben 3 · 10 sowie 2 · 15; ; aus Florin et al. (2019)

Hierzu zählt auch die Fähigkeit zum Darstellungswechsel innerhalb der mathematischen Symbolik (der «intramodale Transfer»), z.B. beim Erkennen der Verwandtschaft der beiden Terme 3 · 10 und 2 · 15 oder zum Beispiel auch der Terme 3 · 5 und 5 + 5 + 5.

2.2 Darstellungswechsel, Grundvorstellungen und Strukturierungsfähigkeiten

Operationsverständnis wird immer wieder in Zusammenhang gebracht mit den drei Begriffen «Darstellungswechsel», «Grundvorstellungen» und «Strukturierungsfähigkeiten». In diesem Kapitel werden die drei Begriffe einzeln und in ihrer Beziehung zueinander erörtert. Die Begriffe werden dann in Kapitel 2.3 verwendet, um das Konstrukt Operationsverständnis zu beschreiben, das im weiteren Verlauf dieser Arbeit behandelt wird.

2.2.1 Der Darstellungswechsel

Darstellungswechsel werden bei der Beschreibung von Lernprozessen in vielen Gebieten verwendet. Sie werden nicht nur in der Mathematik, sondern auch in andern Schulfächern herangezogen, um Lernsituationen zu beschreiben und zu analysieren. Sie werden auch verwendet um Lernsituationen zu schaffen.

In Leisen (2005) wird der Wechsel von Darstellungsformen sogar als Unterrichtsprinzip für sämtliche Fächer vorgeschlagen.

Unterschieden wird zwischen gegenständlichen, bildlichen, sprachlichen, symbolischen und mathematischen Darstellungen. Die letztgenannten mathematischen Darstellungen werden auch auf Struktur- oder Flussdiagramme, Tabellen und Mindmaps sowie auf Formeln und Gesetze aus der Physik, Chemie oder Biologie angewendet. Dabei wird der Wechsel zwischen Darstellungsformen als «der didaktische Schlüssel zum fachlichen Verstehen» bezeichnet. Die Verwendung und der Nutzen des Darstellungswechsels erstrecken sich dabei nicht nur über die naturwissenschaftlichen Fächer und deren Inhalte, sondern umfassen auch sprachliche Aspekte. So wird beispielsweise der Darstellungswechsel in Form eines Wechsels der verwendeten Sprache innerhalb des zweisprachigen Unterrichts als gewinnbringend und verständnisfördernd angesehen und entsprechend diskutiert (Leisen, 2005; Prediger et al., 2011).

Brandt (2019) weist darauf hin, dass Wechsel zwischen unterschiedlichen Darstellungsformen als lernförderlich angesehen werden und aus diesem Grunde in Deutschland auch in den Bildungsstandards verankert wurden.

Darstellungen und Darstellungswechsel im Schulfach Mathematik

Der Wechsel der Darstellungsart sowie die Verwendung vielfältiger Darstellungsarten werden in den verschiedensten Gebieten der Mathematik beschrieben und mit Nutzen für didaktische Zwecke verbunden. Die Fähigkeit, verschiedene Darstellungen nutzen zu können, wird als Zeichen flexibel einsetzbaren Begriffswissens und als ein Bestandteil mathematischer Kompetenz angesehen. Diese Fähigkeit ist Bestandteil von Lehrplänen und Bildungsstandards und wird zu den wichtigen Ideen (den «big ideas») der Mathematikdidaktik gezählt (Lesh et al., 1987; Duval, 2006; Kuntze, 2013, S. 270; Kuhnke, 2013, S. 3). Durch das Üben von Darstellungswechseln im frühen Primarschulunterricht lernen die Kinder also nicht nur ein bestimmtes Thema dieser Schulstufe, sondern üben sich auch in einem Vorgehen, das ihnen in vielen

späteren Gebieten der Mathematik und anderer Schulfächer wieder begegnen wird. Sie erarbeiten sich also auch eine fächerübergreifende Kompetenz.

Die Klassierung der einzelnen Darstellungen, zwischen denen gewechselt werden soll, kann dabei – in Abhängigkeit des Themengebiets, in welchem gearbeitet wird − sehr unterschiedlich ausfallen. Zudem unterscheidet sich die Verwendung der Begriffe nicht nur zwischen Sprachräumen, sondern auch innerhalb desselben Sprachraums. So wird z.B. der englische Begriff «representation» sowohl für den deutschen Begriff «Vorstellung» als auch für «Darstellung» verwendet, was einen Vergleich der Literatur aus unterschiedlichen Sprachräumen erschwert.

Die Verwendung verschiedener Darstellungsarten wird in der englischen Literatur typischerweise mit der Abkürzung MER (use of «multiple external representations») umschrieben (Acevedo Nistal et al., 2009; Ainsworth et al., 2002). In einer semiotisch verwurzelten Definition von «representation» wird dieser Begriff in englischer Sprache definiert als «etwas, das für etwas Anderes steht» (Duval, 2006, S. 103). Unterschieden wird zwischen Wechseln innerhalb gleichartiger Darstellungsarten («Registern») und Wechseln zwischen andersartigen Darstellungsarten. Erstere werden «treatments» oder «intramodale Transfers» genannt, letztere bezeichnet man als «conversions» oder «intermodale Transfers». Betont wird häufig auch die Wichtigkeit der Rückübersetzung («reversibility») und der Flexibilität im Umgang mit solchen Übersetzungen zwischen Darstellungsarten (Flanders, 2014).

Unter den verwendeten Darstellungsarten wird zwischen alltäglichen oder handlungsbezogenen, bildlichen, symbolischen und sprachlichen Darstellungen unterschieden (Prediger et al., 2011). Die jeweiligen Modelle können unterschiedlich konstruiert werden. So kann beispielsweise die sprachliche Darstellung den symbolischen Darstellungen zugeordnet werden (wobei die Sprache als eigene Symbolik betrachtet wird) oder auch den alltäglichen Darstellungen (wenn z.B. Sachsituationen oder Rechengeschichten in sprachlicher Form wiedergegeben werden). Die sprachliche Darstellungsart kann auch als eigenständige Darstellungsform betrachtet werden. Bei Untersuchungen zum zweisprachigen Unterricht spricht man sogar von einer «duplizierten» sprachlichen Darstellungsart (Leisen, 2005; Prediger et al., 2011).

Die Anwendungsgebiete innerhalb des Fachs Mathematik, in denen Darstellungswechsel auftreten bzw. untersucht werden, sind vielfältig und erstrecken sich über die verschiedensten Stufen des Unterrichts von der Vorschule bis zur Hochschule (Zimmermann & Bescherer, 2013). Typische Übersetzungsprozesse sind solche zwischen Zahlen oder Operationen einerseits und bildlichen Darstellungen dieser Zahlen und Operationen, beispielsweise in Form von Punktebildern, andererseits (Ayub et al., 2013; Ladel, 2011a; Padberg & Büchter, 2015). Ebenso wird zwischen Darstellungsarten gewechselt bei der Betrachtung von Funktionen, Wertetabellen und Graphen, die zu diesen Funktionen passen (Ainsworth et al., 2002; Rider, 2007; Laakmann, 2013). Übersetzungsprozesse kommen aber auch zur Anwendung innerhalb der Geometrie bzw. der Vektorgeometrie, der Analysis oder der Stochastik (Duval, 2006; Rider, 2007; Laakmann, 2013; Till & Sproesser, 2013; Flanders, 2014)

Ainsworth et al. (2002) beschreiben als Folge und Nutzen des Einübens von Darstellungswechseln ein robusteres und flexibleres Wissen. Darstellungswechsel werden allgemein als

Bestandteil sowie als Ausdruck von Verständnis angesehen (Thomas, 2008; Spiegel & Selter, 2013).

Zum Teil wird auch die Begrifflichkeit diskutiert und beispielsweise die Verwendung des Begriffs «Darstellungsvernetzung» anstelle von «Darstellungswechsel» vorgeschlagen (Prediger et al., 2011).

Brandt (2019) weist auch darauf hin, dass die Aktivitäten «Darstellen» und «Kommunizieren» in enger wechselseitiger Beziehung zueinander stehen und sich als Begrifflichkeiten kaum scharf voneinander abgrenzen lassen.

Vereinzelt und vorwiegend in jüngerer Zeit wird auch darauf verwiesen, dass Computerunterstützung für das Einüben von Darstellungswechseln und die damit verbundenen Lernprozesse ein verheissungsvolles Mittel sein könnte (Ladel, 2011b; Laakmann, 2013; Ladel, 2018).

2.2.2 Der Darstellungswechsel als Indiz für Operationsverständnis

Für Kinder auf Primarschulstufe ist die Fähigkeit zum sinnvollen und flexiblen Wechsel zwischen verschiedenen Darstellungsarten der arithmetischen Grundoperationen ein besonders häufig genanntes Indiz bzw. ein Indikator für Operationsverständnis (Huinker, 1993; Bönig, 1995; Gerster & Schultz, 2004; Schäfer, 2005; Moser Opitz, 2005; Scherer & Moser Opitz, 2010; Royar, 2013; Kuhnke, 2013; Freesemann & Breucker, 2014; Akinwunmi & Deutscher, 2014; Gaidoschik, 2016b).

Wendet man die Sprache der Semiotik aus Duval (2006) auf das Operationsverständnis an, so wäre das zuweilen beobachtete sinnlose Manipulieren mit Ziffern und Zahlen, welches die Kinder ohne Verständnis betreiben, ein Arbeiten in einem einzelnen «Register». Es fehlt die Verknüpfung mit anderen Registern und die Tätigkeit der Kinder bleibt ein Vorgehen ohne tieferes Verständnis für die Sache. Die Register bleiben isoliert bzw. «compartmentalized». Verständnis ist nur fragmentarisch möglich. Notwendig für Verständnis in der Mathematik ist aber die Koordination der verschiedenen Register (Duval, 2006, S. 124 ff.). Auf das Operationsverständnis angewandt ist dies also die Fähigkeit der Kinder, zwischen Registern bzw. zwischen Darstellungsarten zu wechseln bzw. diese in einen sinnvollen Zusammenhang miteinander zu bringen.

Die Fähigkeit zum Darstellungswechsel bzw. zur Übersetzung zwischen verschiedenen Darstellungen der Grundrechenarten wird damit zum Erkennungsmerkmal und zum Gradmesser für Operationsverständnis.

Diese Fähigkeit erlaubt nun eine Operationalisierung des Konstrukts Operationsverständnis so wie bereits skizziert (Bortz & Döring, 2016, S. 228). Damit verbunden ist die Möglichkeit einer Erfassung des Konstrukts sowohl in Tests und quantitativen Studien als auch in diagnostischen Instrumenten zwecks Einschätzung des Lernstands eines Kindes. Solche Einschätzungen des Lernstands dienen primär dem Zweck, eine allfällige Förderung optimal zu gestalten und auf die Bedürfnisse des Kindes auszurichten.

Notwendig ist aber vorerst eine Präzisierung im Hinblick auf die verschiedenen Darstellungsarten bzw. Darstellungen, zwischen denen das Kind in der Lage sein sollte, flexibel zu übersetzen.

Typischerweise kommt für das Operationsverständnis eine Klassierung in drei verschiedene Darstellungsarten zur Anwendung. Eine dieser Darstellungsarten umfasst dabei meist konkrete Sachsituationen oder Handlungen, eine zweite die bildlichen («ikonischen») Darstellungen, und die dritte Darstellungsart umfasst die mathematisch-symbolische Schreibweise in Form von Termen für die entsprechende Operation.

Diese Dreiteilung hat ihren Ursprung und ihre Wurzeln im EIS-Modell (EIS für «enaktiv», «ikonisch» und «symbolisch») der Entwicklungsstadien von Piaget. Sowohl die Bedeutung der Darstellungsarten als auch der genaue Inhalt haben sich dabei aber gewandelt. Die Darstellungsarten sind nämlich in keiner Weise in einer zeitlichen Entwicklung zu sehen. Bei der Beschreibung von Operationsverständnis stehen die verschiedenen Darstellungsarten lediglich für unterschiedliche Kontexte, in denen die Grundrechenarten in Erscheinung treten. Die symbolische Darstellung stellt also gegenüber anderen Darstellungsarten keinerlei Höherentwicklung dar, sondern ist lediglich eine andere Darstellungsart für den jeweiligen Kontext. Sie stellt eine «Sprache» dar, die man verwenden kann und in die («zu der hin») man übersetzen kann oder die man interpretieren kann («aus der man zurück in den ursprünglichen Kontext übersetzen kann»).

Stellvertretend für die Modelle vieler verschiedener Autoren sei folgendes Diagramm der drei Darstellungsarten mit je einem Beispiel und insgesamt sechs verschiedenen Übersetzungsrichtungen aus Gerster & Schultz (2004) angeführt:

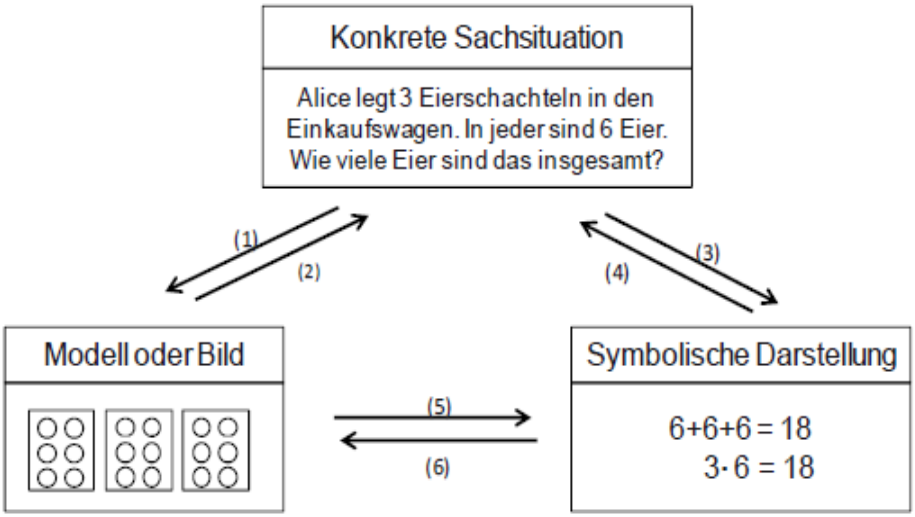


Abbildung 4: Modell mit drei Darstellungsarten und sechs Übersetzungsrichtungen (Gerster & Schultz, 2004)

Die einzelnen Übersetzungsrichtungen werden in der Literatur entweder wie im obigen Diagramm nummeriert oder auch mit passenden Begriffen namentlich bezeichnet. So nennt beispielsweise Baireuther (1999) die Übersetzungsrichtung (1) «Fixieren», die Übersetzungsrichtung (2) «Dynamisieren», (3) «Formalisieren», (4) «Konkretisieren», (5) «Symbolisieren» und (6) «Visualisieren».

Auch in Gleissberg (2018) werden Examensarbeiten zitiert, in denen das Verständnis der Kinder in Zusammenhang gebracht wird mit der Fähigkeit bzw. der Unfähigkeit, solche Übersetzungsprozesse zu leisten.

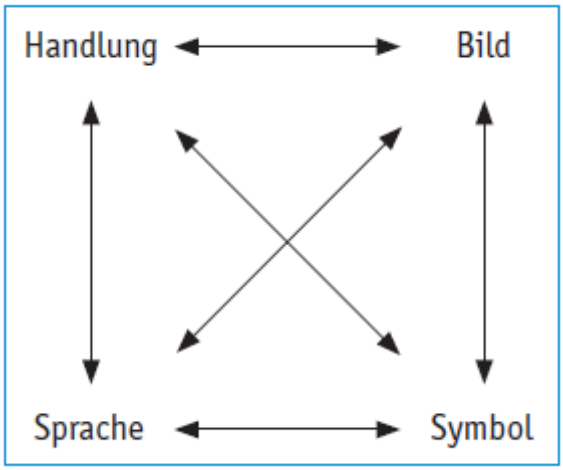


Abbildung 5: Beispiel eines Modells mit vier Darstellungsarten (Freesemann & Breucker, 2014)

In der Literatur finden sich auch vielerlei Hinweise auf alternative Modelle, z.B. mit vier oder mehr Darstellungsarten (Lesh et al., 1987; Bönig, 1995; Leisen, 2005; Prediger et al., 2011; Freesemann & Breucker, 2014; Kaufmann & Wessolowski, 2017).

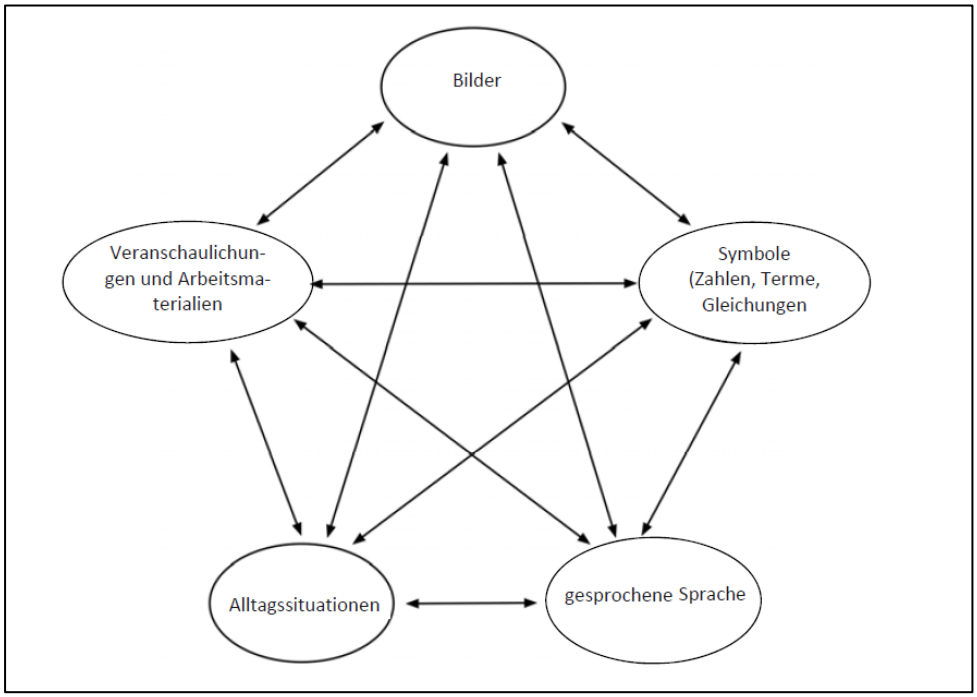


Abbildung 6: Modell mit fünf Darstellungsarten und 20 Übersetzungsrichtungen (Lesh, Post & Behr, 1987)

Ein Modell mit fünf Darstellungsarten beispielsweise entsteht dann, wenn man statische Bilder von manipulierbaren Modellen unterscheidet und ausserdem reale Begebenheiten («real scripts») von gesprochener Sprache («spoken language») trennt und diese gegenüber der ge-

schriebenen symbolischen Sprache («written symbols») abgrenzt (Ayub et al., 2013; Lesh et al., 1987). Aber auch bereits die rein sprachliche Umformulierung der Aufgabe in eine zweite Sprache – möglicherweise die Erstsprache der lernenden Kinder – kann ein Diagramm mit drei Darstellungsarten um eine vierte Darstellungsart ergänzen. Eine ähnliche Ergänzung um eine neue Darstellungsart stellt die Zweiteilung der konkreten Sachsituation in eine «alltagssprachliche Situationsbeschreibung» und eine «fachsprachliche Situationsbeschreibung» dar. Diese Trennung kann zu einer neuen Darstellungsart führen wie z.B. in Prediger et al. (2011) bzw. in Prediger & Wessel (2013) unter Verwendung des Begriffs der «Darstellungsebenen».

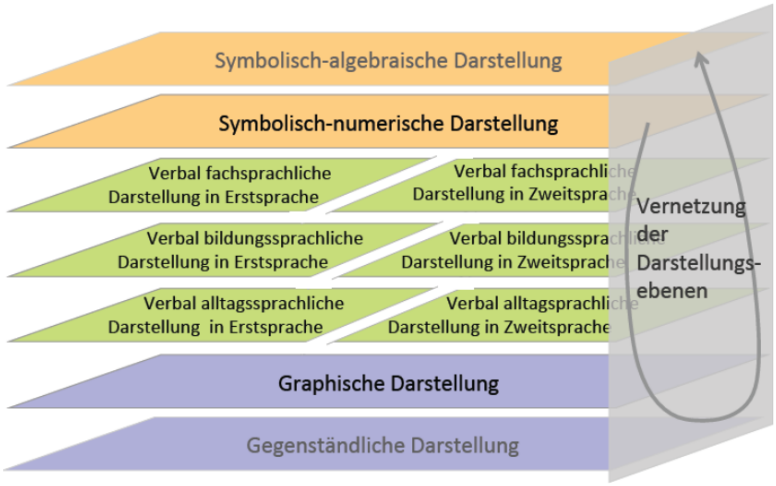


Abbildung 7: Beispiel eines Modells mit verschiedenen Darstellungsebenen (Prediger et al., 2011)

Ein älteres und auf Bruner (1964) zurück gehendes Modell mit vier Darstellungsarten wird in Gleissberg & Eichler (2019) für eine Untersuchung zur Multiplikation verwendet. Die Autoren berichten in ihrer Arbeit von erheblichen Schwierigkeiten von Kindern beim Vollzug des Darstellungswechsels. Sie analysieren Schulbücher im Hinblick auf das Vorkommen des Darstellungswechsels und auf die vollständige Erfassung aller denkbaren Übersetzungsrichtungen. Die Autoren kommen dabei zum Schluss, dass die Schulbücher («programs») von Lehrpersonen durch zusätzliches Material ergänzt werden müssen.

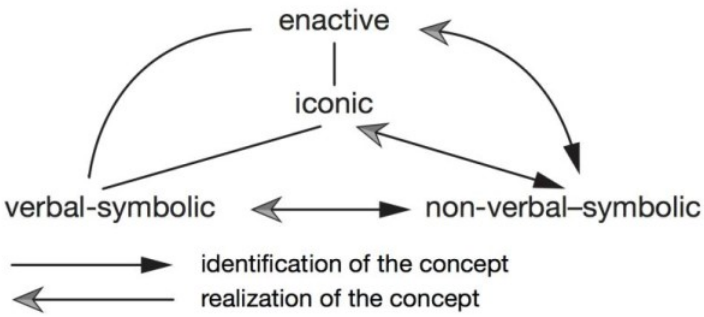


Abbildung 8: Modell von Bruner (1964) mit «Levels of representation»

Zusätzlich zum Übergang zwischen den verschiedenen Darstellungsarten (dem «intermodalen Transfer») wird in der Literatur auch der Übergang zwischen Darstellungen derselben Art (der sogenannte «intramodale Transfer») beschrieben (Bönig, 1995; Kuhnke, 2013). Dieser kommt zur Anwendung, wenn beispielsweise ein Alltagsbild in ein anderes, aber zur selben Grundoperation passendes Bild (z.B. ein Punktefeld) übertragen wird. Ob ein Transfer nun als «intermodaler» oder als «intramodaler» Transfer zu betrachten ist, hängt natürlich wiederum vom gewählten Modell ab. Fasst man z.B. die sprachliche Behandlung einer Sachsituation in der Erst- und in der Zweitsprache in derselben Darstellungsart zusammen, so stellt die rein sprachliche Übersetzungsleistung einen intramodalen Transfer dar. Betrachtet man die beiden sprachlichen Formulierungen aber als unterschiedliche Darstellungsarten, so wird aus der sprachlichen Übertragung ein intermodaler Transfer. Eine Übersetzungsleistung von einem symbolischen Ausdruck (einem Term wie $3 \cdot 6$) in einen anderen symbolischen Ausdruck (z.B. den Term $6 + 6 + 6$) wird in den meisten Modellen als intramodaler Transfer innerhalb der symbolischen Darstellungsart betrachtet. Solche Übersetzungsleistungen zwischen Termen gehören ebenfalls zu den Darstellungswechseln, welche für die Beschreibung des Operationsverständnisses herangezogen werden.

Vom «Indiz» für Operationsverständnis zur «Definition» von Operationsverständnis

Wie dargelegt wird die Fähigkeit zum Darstellungswechsel häufig als ein Indiz für Operationsverständnis genannt. Diese Fähigkeit ist aber vorerst nicht mit Operationsverständnis gleichzusetzen, auch wenn sich im Darstellungswechsel Verständnis oder gegebenenfalls ein Mangel an Verständnis zeigen kann. Er dient dazu, angemessenes oder mangelhaftes Operationsverständnis sichtbar zu machen. Damit dient die Fähigkeit zum Darstellungswechsel der Operationalisierung dieses Konstrukts. Operationsverständnis sollte diese Fähigkeit zum Darstellungswechsel umfassen, genauso wie Verständnis – auch angewandt auf andere Themenbereiche der Mathematik – den Umgang mit und den Wechsel zwischen verschiedenen Darstellungsarten einschliessen sollte. In Thomas (2008, S. 3) wird dies entsprechend formuliert: «Understanding of a mathematical concept should include the ability to relate corresponding elements of different representations of a concept". Verständnis wird aber beschrieben als ein weitreichenderer Begriff als die Fähigkeit, einen oder mehrere Darstellungswechsel korrekt vollziehen zu können. Operationsverständnis umfasst weitere Kompetenzen und Nuancen. Operationsverständnis sollte die Fähigkeit zum Darstellungswechsel enthalten («include»), ist aber nicht identisch mit dieser Fähigkeit und daher nicht mit ihr gleichzusetzen. Eine Gleichsetzung erfolgt höchstens in Form einer Operationalisierung des Konstrukts zum Zweck, diese Fähigkeit z.B. in Erhebungsinstrumenten, Tests oder allenfalls in Prüfungen im Schulunterricht zu erfassen.

Bezogen auf den Unterricht in den ersten drei Primarschuljahren und bezogen auf die grundlegenden Lernziele, die im Hinblick auf die vier Grundrechenarten mit den Kindern in diesen Jahren erreicht werden sollen, ist die Fähigkeit der Kinder zum Darstellungswechsel mehr als nur *ein Indiz* für Operationsverständnis. Die Fähigkeit zum Darstellungswechsel wird − unter Hinzunahme des Anspruchs, diesen auf eine «flexible» Art und Weise vollziehen zu können –

einheitlich als zentrales Kriterium und Erkennungsmerkmal für das Operationsverständnis der Kinder genannt.

Insgesamt betrachtet können die hohe Bedeutung des Darstellungswechsels und die Tatsache, dass alternative Erklärungsmodelle oder andere Indizien für Operationsverständnis selten in der Literatur erwähnt werden, so gedeutet werden, dass dem Darstellungswechsel für dieses Konstrukt zentrale Bedeutung zukommt. Zumindest mit Blick auf die frühen Primarschuljahre und mit gleichzeitigem Fokus auf die rechenschwachen Kinder, mit denen erst einmal das grundlegende Verstehen der einzelnen Operationen sichergestellt werden muss, mag diese Betrachtungsweise berechtigt sein. Mit diesem Fokus lässt es sich auch erklären, dass in manchen Forschungsarbeiten das Operationsverständnis definiert wird als die Fähigkeit der Kinder, den Darstellungswechsel vollziehen zu können (Gerster & Schultz, 2004; Freesemann, 2014).

Der Darstellungswechsel kann also − mit Blick auf die frühen Schuljahre und die rechenschwachen Kinder − nicht nur als *ein*, sondern als *das* wesentliche und zentrale Indiz für Operationsverständnis genannt werden.

In dieser Arbeit wird Operationsverständnis nun operationalisiert als die Fähigkeit, den Darstellungswechsel korrekt und auf sinnvolle Art und Weise vollziehen zu können. Dies geschieht unter Berücksichtigung der verschiedenen inhaltlichen Grundvorstellungen für die einzelnen vier Grundrechenarten. Auch wird das Augenmerk auf die Flexibilität der Kinder im Umgang mit den Grundrechenarten und insbesondere auf die Strukturierungsfähigkeiten der Kinder gelegt.

Unter den verschiedenen Darstellungswechseln wird in dieser Arbeit denjenigen eine besondere Bedeutung beigemessen, bei denen die mathematisch-symbolische Darstellungsart, also die Schreibweise als Term, eine Rolle spielt. Übersetzungsleistungen, bei denen Terme eine Rolle spielen, die also entweder «zum Term hin» oder «vom Term weg» führen, werden als zentral angesehen.

2.2.3 Der Darstellungswechsel zwischen dem Term und anderen Darstellungen

Für das Operationsverständnis werden in der Literatur diejenigen Übersetzungsrichtungen als besonders relevant beschrieben, in denen die symbolische Darstellung entweder als Ausgangs- oder als Endpunkt vorkommt. Die Symbolsprache der Mathematik – als Term oder als «Rechnung» für die Kinder − bildet also bei den wesentlichen Übersetzungsrichtungen entweder den Ausgangs- oder den Endpunkt.

Anderen Darstellungsarten wird häufig unterstützende Wirkung für das Gelingen solcher Übersetzungsprozesse zugebilligt (Lesh et al., 1987). Sie sind aber von weniger zentraler Bedeutung als diejenigen, die entweder «zum Term hin» oder «vom Term weg» führen. Auch der bildlichen Darstellungsart wird bei Übersetzungsprozessen zum Teil eine Vermittlerrolle zugeschrieben (Gerster & Schultz, 2004; Kaur, 2019). Diese Vermittlerrolle kann man sich besonders gut vorstellen bei Aufgabentypen, bei denen von einer Rechengeschichte (passend

zu einer räumlich-simultanen Grundvorstellung) in die Symbolsprache der Mathematik und somit «zu einem Term hin» übersetzt werden soll. Ein Bild kann bei einem solchen Übersetzungsprozess den Sachverhalt so darstellen, dass die Übersetzung dem Kind leichter fällt. Söbbeke (2005) weist in diesem Zusammenhang auch auf gedächtnispsychologische Untersuchungen hin, welche gezeigt haben, dass Menschen sich besser an Bilder als an Worte erinnern können.

Mustersituation, grafische Darstellungen und passende Grundvorstellungen können den Darstellungswechsel bzw. den Übergang von einer realen Sachsituation zur formalen symbolischen Darstellung erleichtern (Prediger, 2009, S. 220 ff.). Dies kann auch via rein gedankliche Vermittlung erfolgen, also auch dann, wenn die Mustersituation nicht explizit zu Papier gebracht wird.

Die entscheidenden Übersetzungsleistungen sind aber jeweils diejenigen, welche die symbolische mathematische Sprache betreffen und/oder eine sinnvolle Verständigung über deren Formalismus erfordern (Barmby et al., 2009, S. 220; Royar, 2013, S. 120 f.).

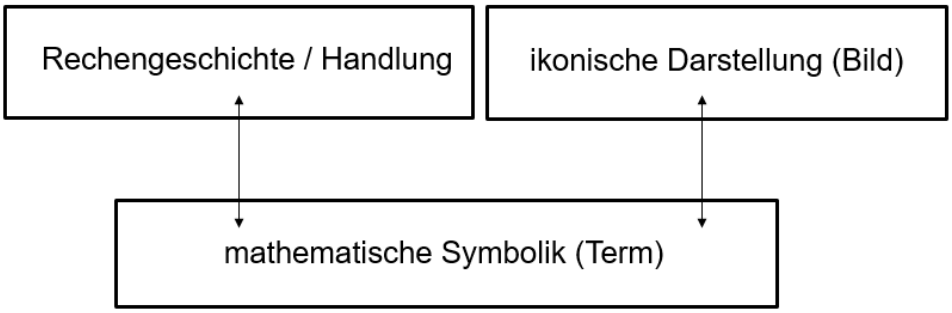


Abbildung 9: Für das Operationsverständnis wesentliche Darstellungswechsel

Operationsverständnis wird somit beschrieben als die Fähigkeit, zwischen realer Situation und formaler mathematischer Schreibweise zu übersetzen (Huinker, 1993; J. Walter et al., 2001, S. 145 f.; Freesemann & Breucker, 2014, S. 9; Moser Opitz & Schmassmann, 2007, S. 266). Die Übersetzungsleistung zwischen realen Situationen und mathematischer Sprache kann dabei – in Anlehnung an den Modellierungskreislauf der Mathematikdidaktik − mit den Begriffen des «Mathematisierens» und des «Interpretierens» bezeichnet werden (Prediger, 2009, 2010). Die Übersetzungsleistung des «Interpretierens» wird zuweilen auch «Entmathematisieren» genannt (Freesemann, 2014).

Möglich ist es auch, beide Übersetzungsrichtungen mit demselben Begriff zu versehen. Bei diesem Sprachgebrauch beschreibt die «Mathematisierungsfähigkeit» den Darstellungswechsel bzw. den Übersetzungsprozess in beide Richtungen, nämlich «zum Term hin» und «vom Term weg». Beschrieben werden dabei die Tätigkeiten «aus realen Gegebenheiten den mathematischen Gehalt herauszuschälen und mit mathematischen Methoden zu bearbeiten *und umgekehrt*». Diese Fähigkeit wird mit Bezug auf die Grundrechenarten explizit in Zusam-

menhang mit Operationsverständnis gestellt. Sie «stellt einen zentralen Aspekt des Operationsverständnisses» dar (Moser Opitz, 2005, S. 123).

Schulz et al. (2019) umschreiben den Darstellungswechsel «von und zum Term hin» mit der Fähigkeit, Situationen (verbal beschriebene Situationen, Handlungen, Bilder oder Textaufgaben) mit mathematisch-symbolischer Schreibweise in Beziehung zu bringen. In Royar (2013) liegt der Schwerpunkt auf dem Übersetzungsprozess «vom Formalismus» zur Sprache mit der Betonung, dass das Kind die eigene Interpretation des Formalismus allgemein verständlich kommunizieren können muss. Dies steht in Einklang mit der Forderung nach Lehrplänen, die berücksichtigen sollten, dass Kommunikation und Begründungen zentral sind für das Verständnis von Mathematik («being able to reason is essential to understanding mathematics»; NCTM nach (Barmby et al., 2009)).

In der hier vorliegenden Arbeit wird für das Konstrukt Operationsverständnis mit der Fähigkeit der Kinder zu Übersetzungsleistungen in diese beiden Richtungen gearbeitet. Das Konstrukt Operationsverständnis wird also mit den Darstellungswechseln zur formalen mathematisch-symbolischen Schreibweise hin («zum Term hin») bzw. von dieser Schreibweise ausgehend in die reale Welt der Handlungen, Geschichten und Bilder («vom Term weg») operationalisiert.

Die verschiedenen Grundvorstellungen der einzelnen Rechenoperationen sollen sich bei der Behandlung dieser Operationen mit den Kindern natürlich niederschlagen. Operationsverständnis kann erst dann *ausreichend* sein, wenn den Kindern die Übersetzungsprozesse nicht nur für einzelne, sondern für die meisten bzw. die in ihrer Stufe üblicherweise behandelten Grundvorstellungen gelingen. Auch wenn die kombinatorische Grundvorstellung bei den Kindern auf der 2. Primarschulstufe noch ausgeklammert bleibt, sind für diese Grundrechenart die anderen Grundvorstellungen zu bedienen und für die beiden Operationen erster Stufe sämtliche in der Literatur genannten Grundvorstellungen bzw. die verschiedenen möglichen Sachsituationen mit den Kindern zu behandeln und zu üben (Gerster & Schultz, 2004).

Verschiedene Aufgabentypen sollen von den Kindern behandelt werden können; deren Fähigkeit zu Übersetzungsleistungen darf sich also z.B. bei der Multiplikation nicht auf rein räumlich-simultane Situationen beschränken. Der Umgang der Kinder mit Termen soll flexibel bleiben; die Kinder sollen also auch Unterschiede zwischen Bildern erkennen, Unterschiede zwischen Termen in Bilder übersetzen und Terme ineinander umwandeln und auf diese Art z.B. eine Malrechnung in eine wiederholte Additionsrechnung übersetzen können. Die Arbeit an strukturiertem Material wie z.B. Punktefeldern ist dafür eine der wesentlichen Vorgehensweisen bei der Förderung der Kinder (Häsel-Weide, 2016).

2.2.4 Das Konzept der Grundvorstellungen

In diesem Kapitel wird zuerst auf die Verwendung des Begriffs «Grundvorstellungen» eingegangen. Es folgt eine kurze Darlegung der geschichtlichen Entwicklung des Begriffs und seiner heutigen Verwendung. Das Kapitel schliesst mit einer Diskussion der Grenzen seiner Nützlichkeit.

Begriffsverwendung und -bedeutung

Der Begriff der «Grundvorstellungen» wird in der deutschsprachigen Mathematikdidaktik seit vielen Jahrzehnten und in Zusammenhang mit verschiedenen Themengebieten der Mathematik verwendet und diskutiert. Er wurde dabei häufig in exemplarischer und pragmatischer Art und Weise und von verschiedenen Autoren durchaus unterschiedlich gebraucht. Eine Übersicht und Zusammenfassung der Begriffsverwendung bis 1995 liefert vom Hofe (1995). Der Autor beleuchtet dabei als Ausgangspunkt und Ziel seiner Untersuchung genau jenes Problemfeld, das sich besonders bei Untersuchungen zum Operationsverständnis der Kinder eröffnet. Dabei handelt es sich um die Tatsache, dass die Kinder mit Begriffen und Symbolen aus der Mathematik zwar umzugehen lernen, dass sie diese aber nicht immer mit adäquaten Deutungen füllen. Damit wird das Rechnen der Kinder zu einem «Manipulieren» von «inhaltslosen Zeichen über auswendiggelernte Regeln». Die Kinder lernen also, sich in der Zahlenwelt der Mathematik zu bewegen und den Unterrichtsalltag im Fach Mathematik zu bewältigen, ohne den eigentlichen Inhalt der verwendeten Symbole in angemessener Weise zu verstehen. Dieses Phänomen wird in der didaktischen Literatur verschiedentlich beobachtet und beschrieben (vom Hofe, 1995, S. 10; Ruwisch, 2001; Lorenz, 2013). Radatz (1991) beschreibt es, indem er das mathematische Formelsystem als ein «bei vielen Schülern dominantes und autonomes» System darstellt. Dieses werde «von vielen Schülern nicht als eine mathematische Abstraktion oder eine mathematische Generalisation verstanden, sondern wie ein unabhängiges und selbständiges System». Mit solchen Vorstellungen und dem Auswendiglernen von «Lösungsregeln» könne ein Kind dann kurzfristig erfolgreich den Anforderungen des Mathematikunterrichts genügen (Radatz, 1991, S. 85 f.).

Gemäss diesen Autoren muss ein Kind für einen verständigen Umgang mit den Grundrechenarten eine angemessene und näher zu spezifizierende Vorstellung entwickeln, damit der Umgang mit diesen Grundrechenarten und das Rechnen mit den Zahlen einen Sinn erhält.

Mit dem Begriff der Grundvorstellungen wird nun diejenige Vorstellung für ein bestimmtes Themengebiet der Mathematik oder einen bestimmten mathematischen Begriff beschrieben, die beim Kind oder beim Lernenden vorhanden sein *soll*. Erst wenn solche Vorstellungen aufgebaut wurden, wird das Agieren des Kindes oder des Lernenden mit den Symbolen und Zeichen in der Mathematik bedeutungsvoll. Es geschieht dann nicht mehr inhaltlos und abgekoppelt von Sinn und Bedeutung, sondern erhält einen Bezug zur Realität.

Es handelt sich bei den Grundvorstellungen also um eine normative Begrifflichkeit, die nicht beschreibt, was beobachtet wird, sondern was sein *soll*. Die beobachteten Vorstellungen von Kindern oder Lernenden werden «Kindervorstellungen», «Schülervorstellungen» oder auch

«individuelle Vorstellungen» genannt. Sie können mit den erwünschen Grundvorstellungen ganz oder zumindest teilweise übereinstimmen. Sie können aber auch von diesen abweichen; oft sogar in solchem Mass, dass sie mit dem Begriff «Fehlvorstellungen» bezeichnet werden müssen.

Damit ergibt sich auch der Nutzen des Grundvorstellungskonzepts. Das Konzept kann erstens herangezogen werden, um beobachtete Vorstellungen der Kinder zu beschreiben (deskriptiver Aspekt), zweitens, um die Differenz zwischen den normativen (erwünschten) Grundvorstellungen und den beobachteten (individuellen) Vorstellungen zu erfassen, und drittens, um auf der konstruktiven Ebene den Unterricht bzw. den Lern- und Lehrprozess so zu gestalten, dass die individuellen Vorstellungen und die Grundvorstellungen sich einander annähern (Griesel et al., 2019; vom Hofe, 1995).

Das Konzept der Grundvorstellungen wird auch als Brückenbau zwischen mathematikdidaktischen und psychologischen Betrachtungen des Lernprozesses von Kindern angesehen. Einzelne Autoren betrachten das Grundvorstellungskonzept als «stoffdidaktische Kategorie zwischen Mathematik und Psychologie», welche unter anderem auch zwischen Mathematikdidaktik und Kognitionspsychologie vermitteln bzw. dazu beitragen soll, Spannungen zwischen diesen beiden Wissenschaften zu überbrücken (Griesel et al., 2019, S. 124 ff.; Lorenz, 2017).

Geschichtliche Entwicklung des Begriffs

In seiner Übersichtsarbeit beschreibt vom Hofe (1995) die geschichtliche Entwicklung des Begriffs der Grundvorstellungen. Er beginnt dabei mit Arbeiten des Mathematikers Leonhard Euler aus dem 18. Jahrhundert und legt dar, wie bereits dieser die Wichtigkeit realer, anschaulicher Beispiele für den Aufbau adäquater Vorstellungen über Mathematik betonte. Euler beschrieb die Lernprozesse noch ohne den Begriff der Grundvorstellungen. Erst in späteren Jahren wurde dann der gedankliche Brückenbau der Lernenden zwischen abstrakter Mathematik in symbolischer Schreibweise und realen Situationen mit dem Begriff des Aufbaus adäquater Vorstellungen verbunden. Die verschiedenen Didaktiker und Mathematiker legten dabei unterschiedliche Schwerpunkte, so bekam z.B. die visuelle Wahrnehmung bei Herbart grosse Bedeutung, während der Eigenaktivität der Lernenden bei Kühnel besondere Wichtigkeit beigemessen wurde. Pestalozzis Arbeiten aber führten dazu, dass die Begriffe «Vorstellungen» und «Anschauungen» in Zusammenhang mit den Lernprozessen von Mathematik immer grössere Verbreitung fanden (vom Hofe, 1995, S. 31).

Der Begriff der «Grundvorstellungen» findet sich dann erstmals bei Oehl (1962). Der Autor beschreibt, wie die «Einsicht», also das Verständnis für den mathematischen Sachverhalt im Unterricht aufgebaut werden soll und die «Übungen zur Erhaltung und Stärkung der Einsicht» darauf folgen sollten. Dabei erläutert er, dass «die gleichen Grundvorstellungen, die bei der Einführung zur Gewinnung der Einsicht führten», zu späterer Zeit «bei der Lösung von Sachaufgaben wieder wirksam» werden (Oehl, 1962, S. 102 f.). Diese Grundvorstellungen der Einsicht sollen nach dem Aufbau immer wieder «lebendig gehalten» und angesprochen werden. Darauf folgend wird dieses Prinzip am Beispiel des Einmaleins illustriert, bei dem nebst Übungen zur Fertigkeit auch immer wieder grundsätzliche Fragen des Verständnisses, der

Interpretation und der Übertragung der Rechenaufgabe in reale Situationen behandelt werden sollen. Dieses Prinzip blieb bis in die heutige Zeit erhalten und spiegelt sich auch in neuerer Literatur wider, so u.a. auch in Gaidoschik (2016a).

Manche Autoren verwenden den Begriff der Grundvorstellung in enger Verwandtschaft zum Begriff «Grundverständnis» (Oehl, 1962; Bender, 1991; vom Hofe, 1995). Vielfach kommen anstelle von Grundvorstellungen mit gleicher Bedeutung auch andere Begrifflichkeiten, so z.B. «Modellvorstellungen», «grundlegende Vorstellungen» oder «Schemata» u.a. zur Anwendung.

Anwendungen und heutige Verwendung des Begriffs

In der heutigen mathematikdidaktischen Literatur wird der Begriff der Grundvorstellungen verwendet, um grundlegende Lernprozesse von Kindern zu beschreiben. Die Grundvorstellungen werden «gedankliche Werkzeuge» genannt, welche zu Zahlen und Strategien aufgebaut werden können. Sind sie aufgebaut, so ermöglichen sie Übersetzungsprozesse zwischen Darstellungsebenen (Wartha, 2011). Es geht dabei darum, grundlegende, wesentliche und allgemein gültige Konzepte zu festigen. Das Bestimmungswort «Grund-» steht dafür, dass es sich um allgemein verbindliche, fundamentale Begriffe handelt, die auch in der realen Lebenswelt verankert sein sollen (Bender, 1991, S. 48).

Grundvorstellungen müssen aufgebaut werden in einer Art und Weise, dass sie später «aktiviert» werden können. Sie sollen es dem lernenden Kind erlauben, mathematische Strukturen in realen Begebenheiten wie Bildern, Handlungen oder Geschichten wiederzuerkennen. Die Grundvorstellungen bilden also eine gedankliche Verankerung oder ein gedankliches Werkzeug, welches es dem Kind oder der lernenden Person erlaubt, die mathematische Struktur in anderen (realen) Situationen wiederzuerkennen. In diesem Sinne ermöglichen sie den Übergang zwischen Realität und Mathematik und agieren so als Bindeglieder zwischen diesen beiden Welten (vom Hofe, 1995, S. 81). Prediger (2010) bezeichnet sie als «Übersetzungsscharniere» zwischen Mathematik und Lebenswelt.

Da diese gedanklichen Werkzeuge immer auch eine Prägung durch das jeweilige lernende Kind (das Individuum) haben, wird das Modell zuweilen durch den subjektiven Erfahrungsbereich des Lernenden (des Individuums) ergänzt. Die Grundvorstellungen werden dann herangezogen, um Beziehungen zwischen Realität, Individuum und Mathematik zu beschreiben.

Das Ziel des Aufbaus von Grundvorstellungen ist es, den Lernenden in einem Ausmass, das seinen individuellen Möglichkeiten entspricht, am Kern des mathematischen Begriffs teilhaben zu lassen, d.h. diesen zu «verstehen» (vom Hofe, 1995, S. 125). Damit ergeben sich offensichtliche Parallelen zum Konstrukt des Operationsverständnisses.

Auch wenn die Grundvorstellungen grundlegende Verstehens-Prozesse beschreiben sollen, hat das Konzept doch auch Anwendung weit über den Primarschulunterricht hinaus gefunden, namentlich in verschiedenen Gebieten des späteren und fortgeschrittenen Mathematikunterrichts, so z.B. in der Wahrscheinlichkeitsrechnung (Malle & Malle, 2003), ferner in Gebieten

der Differentialrechnung innerhalb der Analysis (Greefrath et al., 2016; Malle, 2003) sowie in verschiedenen anderen Themenbereichen.

Grenzen des Begriffs

Der Begriff der Grundvorstellungen wird wie dargelegt verwendet, um einen Bezug zwischen Realität und mathematischer Symbolik zu beschreiben. Offen bleibt die Frage, ob der Begriff auch für Mathematiklernprozesse anwendbar ist, bei denen keine offensichtliche Beziehung zur Realität mehr gegeben ist. Wenn in der höheren Mathematik der Alltagsbezug gänzlich fehlt, beispielsweise, wenn Vektorräume mit Dimensionen grösser als 4 oder gar unendlichdimensionale Vektorräume eingeführt werden, dann greift die in der Mathematikdidaktik übliche Formulierung («Bezug zur Realität») im ursprünglichen Sinne nicht mehr. Im eben genannten Beispiel könnte es eigentlich als durchaus angemessen angesehen werden, von Grundvorstellungen und von einem Grundverständnis für die neuen Begrifflichkeiten zu sprechen, doch wären die verwendeten Begriffe in solchen Anwendungsbereichen vermutlich anders zu umschreiben. Die Begriffe sind aber eingeführt worden, um Verstehens-Prozesse in der Schule zu beschreiben, und die eben genannte Schwierigkeit tritt dort nicht oder nur in sehr geringem Ausmass auf. Allenfalls klingt sie bei der Behandlung des Begriffs des «Unendlichen» in der Sekundarstufe II an.

Bei der Lektüre mathematikdidaktischer Literatur fällt auf, dass die Klassierung der Grundvorstellungen auch zu einfachen Inhalten wie z.B. der Multiplikation keineswegs einheitlich ausfällt. Die Einteilung mathematischer Themen in Grundvorstellungen ist keine exakte Wissenschaft. Vielmehr handelt es sich um Kategorisierungen mit didaktischen Absichten, also mit dem Zweck, Lernprozesse zu vereinfachen. Die Ergebnisse solcher Prozesse können unterschiedlich ausfallen.

Zu diskutieren wäre auch die Frage, ob die mehrfache Verwendung des Begriffs «Grundvorstellung» als individuelle Schnittstelle zwischen Mathematik und Realität einerseits sowie zur Unterscheidung verschiedener Aufgabentypen andererseits (z.B. bei den Grundoperationen wie zum Beispiel der Multiplikation) sinnvollerweise wirklich mit demselben Begriff erfolgen sollte. Vielleicht wären unterschiedliche Begriffe für unterschiedliche Dinge – so wie sie in der englischen Sprache verwendet werden – angemessener. Mulligan & Mitchelmore (1997) teilen multiplikative Situationen beispielsweise in vier Gruppen ein und nennen diese vier Situationen dann schlicht Kategorien («categories»). Die Einteilung erinnert zwar an jene in die verschiedenen Grundvorstellungen aus dem deutschsprachigen Raum, weist aber auch Unterschiede zu dieser auf. Insbesondere besteht kein Anspruch darauf, einen Begriff einzuführen, mit dem auch in anderen Gebieten der Mathematik Einteilungen vorgenommen werden können.

Die Nützlichkeit des Begriffs der Grundvorstellungen wird also auch dadurch eingeschränkt, dass dieser (zumindest bisher) ausserhalb der deutschsprachigen Literatur keine Verbreitung gefunden hat. Entsprechende deutschsprachige Werke und Kategorisierungen sind und bleiben schwierig zu übersetzen.

Der Grundvorstellungsbegriff scheint also keine Notwendigkeit darzustellen. Dies wird belegt durch die Tatsache, dass auch sehr grosse Sprachregionen ohne eine solche Begrifflichkeit auskommen. Er wird aber in der deutschsprachigen Literatur − zumindest bisher − als nützlich angesehen und häufig verwendet.

Da er in der deutschsprachigen Literatur weit verbreitet ist und grosse Teile der deutschsprachigen Literatur (z.B. auch bei der Beschreibung von Operationsverständnis) mit diesem Begriff arbeiten, wird er in dieser Arbeit ebenfalls verwendet und mit anderen Begrifflichkeiten in Zusammenhang gebracht.

2.2.5 Grundvorstellungen und Operationsverständnis

Der Begriff der «Grundvorstellungen» und das Konstrukt «Operationsverständnis» werden in der mathematikdidaktischen Literatur verschiedentlich in Zusammenhang zueinander gebracht, häufig aber auch nebeneinander und unabhängig voneinander verwendet. In diesem Kapitel werden die Gemeinsamkeiten und die Unterschiede der beiden Begriffe diskutiert. Dabei wird aufgebaut auf der Vorarbeit, die beim Vergleich dieser beiden Begriffe in Prediger (2009) bereits vorgenommen wurde.

Gemeinsamkeiten der beiden Begriffe

Die Ähnlichkeit und Verwandtschaft der beiden Konzepte «Grundvorstellungen» und «Operationsverständnis» wurde bereits durch die sprachliche Beschreibung der beiden Konzepte in den bisherigen Ausführungen deutlich gemacht. Augenfällig ist, dass dasselbe Wort «Grund-» nicht nur in den «Grundvorstellungen», sondern auch in den «Grundoperationen» vorkommt und so die fundamentale und zentrale Bedeutung beider Konzepte sogar mit ähnlichem Sprachgebrauch betont wird.

Werden die beiden Begriffe «Grundvorstellungen» und «Operationsverständnis» miteinander in Zusammenhang gebracht, so kann dies explizit oder implizit geschehen. Beim expliziten Vorgehen werden beide Begriffe verwendet und dabei vergleichend und im Bestreben auf Klärung der Unterschiedlichkeit nebeneinander gestellt, so z.B. in Prediger (2009). Häufiger aber ist eine implizite Verzahnung der beiden Begriffe. Es werden dabei Aspekte des einen Begriffs hinzugezogen, ohne dass dieser explizit genannt wird. So werden z.B. einzelne der vier Grundrechenarten hinzugezogen, um Aspekte des Grundvorstellungsbegriffs zu erläutern. Oehl (1962, S. 103) erläutert beispielsweise den Grundvorstellungsbegriff mit Hilfe der Multiplikation und auch andere Autoren ziehen immer wieder die Grundrechenarten herbei, um Aspekte des Grundvorstellungsbegriffs zu beleuchten (Bender, 1991, S. 55 f.). Gaidoschik (2016a, S. 38 f., 71) beschreibt den Aufbau von Grundvorstellungen für die Multiplikation in Worten, die in anderen Arbeiten für die «Förderung von Operationsverständnis» benutzt werden.

Die enge Verzahnung der beiden Begriffe zeigt sich auch darin, dass sie in der Literatur in kombinierter Form verwendet werden, z.B. dann, wenn von «Vorstellungen von Operationen»

oder vom «Aufbau von Operationsvorstellungen» die Rede ist (Benz, 2018). Moser Opitz (2013, S. 142) schreibt im Zusammenhang mit dem allgemeinen Lernen von Mathematik vom «Aufbau von Vorstellungen». Der «Aufbau von Mathematisierungsfähigkeiten» stellt einen spezifischen solchen Vorstellungsaufbau dar.

Für eine Klärung der Begriffsverwendung ist es hilfreich, sich zuerst einmal bewusst zu werden, dass in der historischen Entwicklung des Begriffs der «Grundvorstellung» der Aspekt des «Grundverständnisses» für den Sachverhalt immer schon eine Rolle gespielt hat. Besonders deutlich kommt dies in Bender (1991) zum Ausdruck. Der Autor schreibt durchwegs von «Grundvorstellungen und Grundverständnis» und verwendet diesen Begriff immer in dieser kombinierten sprachlichen Form. Der Autor hält in seinem Beitrag fest, dass «die Verständnisse in einer engen Wesensverwandtschaft mit den Vorstellungen» stehen und dass «Verständnis nicht ohne Vorstellungen» und «Vorstellungen nicht ohne Verständnis» möglich sind (Bender, 1991, S. 55). Bei der Verwendung des Begriffs «Grundvorstellung» schwingt also stets auch der Aspekt des Verstehens mit.

Eine weitere Parallele zwischen den beiden Begriffen ergibt sich durch die Festlegung auf die wesentlichen Übersetzungsrichtungen des Mathematisierens («zum Term hin») und des Interpretierens («vom Term weg») für den Begriff des Operationsverständnisses, wie er in Kapitel 2.2.3 beschrieben wurde.

Als zentraler Aspekt von Operationsverständnis wird die Fähigkeit eines Kindes angesehen, zwischen Sachsituation (sei sie nun sprachlicher oder bildlicher Art) und symbolisch-mathematischer Sprache hin und her zu übersetzen.

Grundvorstellungen hingegen betrachten wir als die gedankliche Verankerung, die beim Kind aufgebaut und aktivierbar sein muss, um die mit Operationsverständnis beschriebenen Übersetzungsleistungen zwischen Realität und Mathematik zu leisten. Mit dieser gedanklichen Verankerung oder diesem «Übersetzungsscharnier» können die Kinder die wesentliche mathematische Struktur in Sachsituationen wiedererkennen und/oder die mathematische Struktur auf Sachsituationen korrekt anwenden (Prediger, 2010).

Reduziert man Operationsverständnis solchermassen auf die wesentlichen Übersetzungsleistungen und fasst das Konzept der Grundvorstellungen auf die obige Art zusammen, dann zeigt sich die Nähe der Begriffe besonders deutlich.

Sind Grundvorstellungen aufgebaut und aktivierbar, dann wird Operationsverständnis sichtbar. Oder umgekehrt: Wird Operationsverständnis sichtbar, so müssen zuvor Grundvorstellungen aufgebaut und aktiviert worden sein. Somit wird mit dem «Aufbau von Grundvorstellungen» und der «Förderung von Operationsverständnis» derselbe Vorgang beschrieben. So betrachtet erklärt sich auch die Begriffsverwendung in Gaidoschik (2016a, S. 38 ff.). In dieser Arbeit ist vom «Aufbau von Grundvorstellungen» die Rede an Stellen, in denen in anderen Werken die Begrifflichkeit «Förderung von Operationsverständnis» verwendet wird.

Verschiedenheit der Begriffe

Trotz dieser begrifflichen Nähe gibt es Unterschiede zwischen dem Begriff der «Grundvorstellungen» und dem Konstrukt des «Operationsverständnisses». Prediger (2009) nennt zwei dieser Differenzen und hält fest, dass Operationsverständnis als Teilaspekt der Grundvorstellungen betrachtet werden kann, da sich der Begriff des Operationsverständnisses nur auf die vier Grundrechenarten, der Begriff der Grundvorstellungen hingegen auf viele weitere Gebiete der Mathematik bezieht. Der Begriff der Grundvorstellungen hat also einen deutlich grösseren Geltungsbereich und die grössere Reichweite als der Begriff des Operationsverständnisses.

Auch Freesemann (2014, S. 42) bringt das Konstrukt Operationsverständnis in Zusammenhang mit Modellvorstellungen. Sie erwähnt, dass jeder Grundoperation verschiedene Modellvorstellungen zugrunde liegen. Das Verständnis der Kinder muss sowohl für die Operationen als auch für die Modellvorstellungen vorhanden sein, damit die Kinder verständig mit Rechenaufgaben umgehen können.

Prediger (2009, S. 220) beschreibt zudem, dass das Operationsverständnis eine Erweiterung des Grundvorstellungskonzepts darstellt, da beim Operationsverständnis auch graphische Darstellungen hinzugenommen werden. In diesem Sinne wiederum ist das Konstrukt des Operationsverständnisses umfangreicher als das Konzept der Grundvorstellungen.

Der erste dieser beiden genannten Unterschiede ist unbestritten. Das Konstrukt Operationsverständnis bezieht sich auf die vier Grundrechenarten, während das Grundvorstellungskonzept von sehr viel weitreichenderer Bedeutung ist und viele weitere Gebiete der Mathematik unter dem Gesichtspunkt von Grundvorstellungen betrachtet werden können.

Der zweiten genannten Differenz könnte mit dem Argument begegnet werden, dass auch graphische Darstellungen (welcher Art auch immer) Teil der realen Welt sind und unter den Geltungsbereich der Grundvorstellungen – die zwischen realen Situationen, der Individualität des Betrachters und dem symbolischen Formalismus der Mathematik «vermitteln» – fallen.

Eine zusätzliche Differenz zwischen den beiden Begriffen ergibt sich beim Einbezug der Granularität der Betrachtungsweise. Während beim Operationsverständnis immer die vier Grundrechenarten oder eine Teilmenge davon untersucht wird, zerfallen diese Grundrechenarten bei Betrachtung durch die Brille der Grundvorstellungen in verschiedene Typen, so z.B. in Situationen räumlich-simultaner, zeitlich-sukzessiver und kombinatorischer Art bei der Multiplikation und in vergleichbarer Weise für die andern drei Grundrechenarten. Das Konzept der Grundvorstellungen erlaubt also eine Aufteilung der Grundrechenart und eine der jeweiligen Grundrechenart angemessenere, gegebenenfalls auch feinere Betrachtung durch eine noch etwas schärfere Brille.

Im Konstrukt Operationsverständnis erscheint die Bündelung zu den vier Grundrechenarten, also die gemeinsame Betrachtung quasi «in einem Korb», zumindest aus mathematikdidaktischer Hinsicht als diskussionswürdig. Dies ist darum der Fall, weil die vier Grundrechenarten doch eigentlich sehr unterschiedliche Operationen und Vorgänge beschreiben. Auch in didaktischer Hinsicht unterscheiden sich die vier Operationen, was nur schon durch die sehr unter-

schiedlichen Anforderungen an die Kinder deutlich wird. Die vier Operationen unterscheiden sich deutlich im Hinblick auf den Schwierigkeitsgrad bei Erstbehandlung im Unterricht. Diese Unterschiedlichkeit äussert sich auch im Hinblick auf den Zeitpunkt der Erstbehandlung im Unterricht und der ersten Empfänglichkeit der Kinder für ihren Inhalt. Für die Addition beispielsweise ist ein erstes frühes Verständnis bereits lange vor dem Schuleintritt erkennbar. Die Multiplikation weist vom fachlichen Gehalt und Anspruchsniveau eher Parallelen zum Stellenwertverständnis auf (Hurst & Hurrel, 2016). Inhaltlich geht es sowohl bei der Multiplikation als auch beim Stellenwertverständnis um Bündelungen einzelner Einheiten. Die Addition ist hingegen eine viel «leichtere» und für die Kinder zugänglichere Grundoperation. Die gemeinsame Betrachtung der Addition und der Multiplikation − sowie auch der anderen beiden Operationen − unter der Bezeichnung Operationsverständnis kommt wohl eher aus der Alltagssprache und der Tatsache, dass es für genau diese vier Operationen ein weit verbreitetes Zeichen gibt, welches in der frühen Primarschulstufe eingeführt wird. Inhaltlich und strukturell sind die vier Grundrechenarten − auch wenn sie «paarweise Umkehroperationen» voneinander sind − sehr unterschiedliche Konstrukte.

Das Konzept der Grundvorstellungen ist im Hinblick auf die Zusammenfassung unterschiedlicher Themen flexibler ausgelegt als das Konstrukt Operationsverständnis. Bei den Grundvorstellungen werden die unterschiedlichen Grundrechenarten nicht in einem «künstlichen» Korb verpackt. Sie werden − ganz im Gegenteil − separat und mit Fokus auf die jeweiligen Inhalte und deren mathematische Substanz betrachtet.

Zu bemerken ist aber, dass die Unterteilung einer einzelnen Grundrechenart in ihre Grundvorstellungen durchaus unterschiedlich vorgenommen werden könnte. Die übliche Einteilung (räumlich-simultan, zeitlich-sukzessiv, kombinatorisch) kann z.B. ergänzt werden durch eine weitere «proportionale» Grundvorstellung oder eine Grundvorstellung des «multiplikativen Vielfachen».

Padberg (2009, S. 122 f.) benennt nebst den oben erwähnten Grundvorstellungen zur Multiplikation fünf weitere «Klassen von multiplikativen Kontextaufgaben», die jeweils durch entsprechende Beispiele illustriert werden. Es handelt sich bei den fünf weiteren Klassen um den «multiplikativen Vergleich», das «multiplikative Ändern», die «Proportionalität», die «Verkettung von Vervielfältigungsoperatoren» und die «formelhafte Multiplikation von Grössen».

Eine leicht unterschiedliche Einteilung in vier multiplikative Situationen ergibt sich mit den Kategorien «äquivalente Gruppen», «multiplikativer Vergleich», «Rechteckfläche» und «kartesisches Produkt» wie beschrieben in Mulligan & Mitchelmore (1997) bzw. in Ruwisch (Ruwisch, 1999, S. 123 ff.). In der letztgenannten Arbeit wird dargelegt, dass die erste dieser vier Kategorien in der Mathematikdidaktik im deutschsprachigen Raum aufgeteilt wird in die beiden Grundvorstellungen räumlich-simultaner und zeitlich-sukzessiver Art. Die vierte Kategorie entspricht wiederum der kombinatorischen Situation bzw. Grundvorstellung. Die Autorin erläutert auch die Gründe, weshalb die kombinatorische Grundvorstellung im Anfangsunterricht zur Multiplikation noch nicht verwendet werden sollte (Ruwisch, 1999, S. 130).

Sofern man das Konstrukt Operationsverständnis aber nicht nur auf die beiden Darstellungswechsel des Mathematisierens und des Interpretierens reduziert und es in Anlehnung an die

Verwendung von «operation sense» im englischen Sprachraum verwendet, ergibt sich eine weitere inhaltliche Differenz zum Konzept der Grundvorstellungen. Das Operationsverständnis erfasst dann auch Nuancen und Sensibilitäten, die bei den Grundvorstellungen meist nicht mitformuliert sind.

Fazit

Zusammenfassend kann festgehalten werden, dass das Konzept der Grundvorstellungen weitreichender, aber auch etwas flexibler ist als das Konstrukt Operationsverständnis. Bei Einschränkung der Betrachtungen (Grundvorstellung nur für die vier Grundrechenarten) und bei Anwendung der Operationalisierung und Definition von Operationsverständnis, wie sie in der deutschsprachigen Literatur möglich ist (Reduktion auf die beiden Darstellungswechsel gemäss Kapitel 2.2.3), stehen sich die beiden Begriffe aber sehr nahe. Dies untermauert auch die Aussage, dass Vorstellungen nicht ohne Verständnis und Verständnis nicht ohne Vorstellungen möglich sind (Bender, 1991). Bei Betrachtung der auf die Grundrechenarten eingeschränkten Grundvorstellungen und Verwendung obiger Begrifflichkeiten kann also gefolgert werden, dass es sich beim «Aufbau von Grundvorstellungen» und bei der «Förderung von Operationsverständnis» um dieselbe Aktivität handelt. Wird Operationsverständnis beobachtet, so sind Grundvorstellungen vorhanden und aktiviert worden. Dasselbe gilt umgekehrt: Operationsverständnis wird genau dann beobachtet, wenn der Aufbau von Grundvorstellungen gelungen ist und diese auch aktiviert werden können (Wartha, 2011; Wartha & Schulz, 2013).

Das Konstrukt des Operationsverständnisses sollte aber − wie bereits erwähnt − nicht nur auf Gelingen oder Nichtgelingen einzelner Darstellungswechsel reduziert werden. Die verschiedenen Darstellungswechsel mögen als Indiz dienen; das Konstrukt Operationsverständnis hat aber häufig eine weitere und umfassendere Bedeutung, die sich auch in der mathematikdidaktischen Literatur von Arbeit zu Arbeit unterscheiden kann.

Der Begriff «Operationsverständnis» erlaubt neben der Diskussion in der Mathematikdidaktik auch eine pragmatische und leicht nachvollziehbare Besprechung eines der wichtigen Phänomene des Mathematiklernens der Kinder in den ersten Primarschuljahren. Der Begriff spielt eine wesentliche Rolle im Unterricht und natürlich auch im Lehramtsstudium für künftige Lehrpersonen der Unterstufe. Aus diesem Grunde ist es angemessen, eine eigene und von ihrer sprachlichen Ausgestaltung her ein Stück weit auch «selbsterklärende» Begrifflichkeit zu verwenden. Mit dem Begriff «Operationsverständnis» kann dieses Phänomen im Lehramtsstudium und in der Praxis auf pragmatische Art und Weise besprochen werden. Dies auch dann, wenn die Begrifflichkeit unter Zuhilfenahme eines anderen Konstrukts (wie eben demjenigen der Grundvorstellungen) ersetzt werden könnte bzw. die beobachteten Phänomene allenfalls sogar noch präziser beschrieben werden könnten.

2.2.6 Strukturierungsfähigkeiten und Operationsverständnis

Häufig wird in der Literatur auch das Arbeiten der Kinder mit Struktur sowohl in der mathematisch-symbolischen Schreibweise als auch in anderen Darstellungsformen mit dem Konstrukt Operationsverständnis in Zusammenhang gebracht. Die Kinder sollen nicht nur einzelne Terme bildlich darstellen und auf diese Art zwischen Darstellungen wechseln können. Sie sollen beispielsweise auch darlegen können, wie sich die Änderung eines Terms in der bildlichen Darstellungsform niederschlägt. Auch sollen sie erläutern können, wie die strukturelle Änderung einer bildlichen Darstellung (z.B. eines Punktefelds) oder die andersartige Betrachtung einer solchen bildlichen Darstellung sich in der symbolischen Schreibweise zeigt. Zur Illustration dieses Sachverhalts sei an die einführenden Beispiele in Kapitel 2.1.3 erinnert, welche in Abbildung 2 bzw. in Abbildung 3 dargestellt sind.

Das Nutzen von Beziehungen zwischen Rechenaufgaben bzw. von Strukturen bei bildlichen Darstellungen durch Kinder wird in der Literatur explizit mit Operationsverständnis in Zusammenhang gebracht. In verschiedenen Beschreibungen aus der wissenschaftlichen Literatur wird dieser Aspekt nebst dem Darstellungswechsel und dem Konzept der Grundvorstellungen für eine umfassende Beschreibung von Operationsverständnis herangezogen (siehe Abbildung 10).

Abbildung 10: Aspekte von Operationsverständnis gemäss (pikas-kompakt.dzlm.de)

Von Strukturierungsfähigkeiten wird in der Literatur immer wieder in Zusammenhang mit grundsätzlichem Verstehen und mit Flexibilität der Kinder berichtet. Dabei wird zum Teil auch explizit auf die Struktur von Malrechnungen Bezug genommen. Ein strukturorientierter Umgang mit Anschauungsmitteln ergibt sich dabei nicht von selbst aus den Anschauungsmitteln, sondern muss durch eine entsprechende Unterrichtskultur begleitet werden. Kinder müssen dabei lernen, Anschauungsmittel aktiv zu deuten und Strukturen in diese hinein zu lesen (Söbbeke, 2005; Beutler, 2013; Steenpass, 2014; Benz, 2018).

Söbbeke (2005) entwickelt ein Konstrukt «visuelle Strukturierungsfähigkeit», welches als Bindeglied zwischen Anschauungsmitteln und internen, mentalen Repräsentationen des Kindes («im Geiste des Kindes») zu verstehen ist. In dieser Umschreibung zeigt sich auch die in Kapitel 2.2.4 beschriebene Nähe des Konstrukts zu den Grundvorstellungen. Auch Strukturierungsfähigkeiten «vermitteln» zwischen Vorstellungen; sie stehen ebenfalls mit Darstellungswechseln in Zusammenhang. Wie die Grundvorstellungen können sie als Bindeglied zwischen Mathematik, strukturierter bildlicher Darstellung und individueller Gedankenwelt des lernenden Kindes betrachtet werden.

Die visuelle Strukturierungsfähigkeit enthält aber noch eine weitere Komponente. Es handelt sich dabei um die Fähigkeit des Kindes, ein Diagramm visuell zu strukturieren (Söbbeke,

2005, S. 67 ff.). Die Nähe zum Konstrukt Operationsverständnis kommt zur Geltung, wenn diese Strukturierung durch das Beschreiben des Anschauungsmaterials mit mathematischen Termen und Symbolen vorgenommen wird. Die Kinder beschreiben dann das Material mit Hilfe eines Darstellungswechsels zwischen Bild und Term.

Das Konstrukt der «visuellen Strukturierungsfähigkeit» wird in der Literatur präzisiert und zerlegt in mehrere Ebenen, die sich durch einen Anstieg der Flexibilität bei der Nutzung von Beziehungen und Umdeutungen voneinander unterscheiden (Söbbeke, 2005, S. 131 ff.) .

Damit erhält dieses Konstrukt eine feinere Unterteilung und Beschreibung als das schlichte «gelungen/nicht gelungen» bei der Betrachtung des Darstellungswechsels.

Der Dreiklang zwischen Darstellungswechsel, Gebrauch von Grundvorstellungen und Arbeit mit Struktur und Strukturierungsfähigkeit wird in der didaktischen Literatur auch entsprechend beschrieben:

Zu einem umfassenden Operationsverständnis gehört die Berücksichtigung und Förderung der folgenden drei Aspekte:

- Grundvorstellungen zur Rechenoperation
- Fähigkeit zum Darstellungswechsel
- Erkennen und Nutzen von Beziehungen und Strukturen zwischen Aufgaben

Abbildung 11: Drei Aspekte von umfassendem Operationsverständnis (primakom.dzlm.de)

Steenpass (2014) schliesst an das theoretische Konstrukt «visuelle Strukturierungskompetenz» an und baut darauf auf. Sie entwickelt ein Konstrukt «rahmungsbasierte Deutungskompetenz», diskutiert dieses und beobachtet, wie die Kinder als Lernende in der Auseinandersetzung mit Anschauungsmitteln ihr eigenes Verständnis erwerben. Untersucht wurden einzelne Kinder aus der 3. Schulklasse mit Hilfe von Interviews im Rahmen einer Interventionsstudie. Nachgewiesen wird dabei, dass sich durch geeignete Interventionen eine flexible Deutungskompetenz fördern lässt. Dies ist ein Indiz dafür, dass die Gestaltung des Unterrichts auf die Förderung der Deutungs- und Strukturierungsfähigkeiten der Kinder einen bedeutenden Einfluss hat. Für die Entwicklung eines gesicherten Verständnisses bei den Kindern ist es wichtig, dass die Lehrperson die Vielschichtigkeit von Deutungsprozessen im Auge behält. Im Unterricht sollten komplexe Deutungsanforderungen thematisiert, Umdeutungen und Umstrukturierungen vorgenommen und auf die Vielfalt von Deutungen geachtet werden. Gerade mehrdeutige und offene Aufgabenformate erwiesen sich in dieser Studie als wirksam (Steenpass, 2014, S. 61 ff., 281 ff.). Dieser Aspekt der Strukturierungstätigkeit und der Flexibilität im Umgang mit entsprechenden Anschauungsmitteln kann herangezogen werden, um das Konstrukt Operationsverständnis wie oben dargelegt zu ergänzen und in angemessener Weise zu fassen.

2.3 Das Konstrukt Operationsverständnis

In diesem Kapitel werden die theoretischen Überlegungen zum Konstrukt Operationsverständnis um weitere Aspekte ergänzt. In Kapitel 2.3.1 wird die Vereinigung der verschiedenen Operationen innerhalb eines einzelnen Konstrukts besprochen. Beachtung findet dabei insbesondere der sehr unterschiedliche Schwierigkeitsgrad dieser Operationen. In Kapitel 2.3.2 wird die Abgrenzung gegenüber dem Begriff «operation sense» im englischsprachigen Raum vorgenommen. Schliesslich wird in Kapitel 2.3.3 die für diese Arbeit vorgenommene Konzeptualisierung des Konstrukts dargestellt.

2.3.1 Die vier einzelnen Grundrechenarten und deren vereinigte Betrachtung

Die Schwierigkeiten der Kinder mit den Grundoperationen bzw. mit dem Wechsel zwischen Darstellungen zu diesen Operationen zeigen sich nicht in allen vier Grundrechenarten im selben Masse. Insbesondere beim Multiplizieren und Dividieren, also bei den beiden Operationen zweiter Stufe, treten bereits im kleinen Zahlenraum verstärkt Probleme auf (Bönig, 1995; Schäfer, 2005; Moser Opitz, 2005, 2013; Freesemann, 2014).

Verschiedene arithmetische Fähigkeiten werden von den Kindern bereits im Vorschulalter entwickelt und aufgebaut. Es wurde beispielsweise bemerkt, dass der Erwerb eines frühen Verständnisses für die Addition (die «leichteste» der vier Grundrechenarten) bereits in den sehr frühen Kindheitsjahren stattfinden kann. Dieses frühe Verständnis bleibt dann aber vorerst auf sehr kleine Anzahlen beschränkt (Geary, 1994; Steiner, 1997; Gerlach, 2007; Purpura & Lonigan, 2013; Benz, 2018).

Die Subtraktion hat gegenüber der Addition einen erhöhten Schwierigkeitsgrad. Auch stellen nach deren Behandlung die Multiplikation und das Verständnis für diese Grundrechenart die Kinder wiederum vor neue Schwierigkeiten. Trotz erhöhtem Schwierigkeitsgrad gegenüber den Operationen erster Stufe wurde aber beobachtet, dass auch Kinder im Vorschulalter in der Lage sind, einfache multiplikative Aufgaben richtig zu bearbeiten (Carpenter et al., 1993; Sherin & Fuson, 2005, S. 354).

Als anspruchsvollste der vier Grundrechenarten wird in der Fachliteratur die Division genannt (Oehl, 1962; Schäfer, 2005; Ruwisch, 1999; Barmby et al., 2009; Gaidoschik, 2010; Moser Opitz, 2013; Ehlert et al., 2013; Schulz et al., 2019; Burtscher, 2019).

Diese Aussagen erscheinen plausibel und spiegeln sich auch in der Reihenfolge der Behandlung der vier Grundrechenarten im Laufe der ersten drei Schuljahre wider. Die Behauptungen werden aber auch durch empirische Evidenz untermauert. So stellt Schäfer (2005) das Ergebnis einer Messung des Operationsverständnisses dar, welches zeigt, dass der grösste Sprung im Schwierigkeitsgrad zwischen den Operationen erster und zweiter Stufe stattfindet. Dieselbe Untersuchung belegt auch, dass Mängel im Operationsverständnis mit zählendem Rechnen einhergehen und dass sich kombiniert rechenschwache Kinder auch mit dem Operationsverständnis schwerer tun als Kinder mit isolierten Rechenschwächen.

Ebenfalls empirisch belegt werden die Abstufungen des Schwierigkeitsgrads in Schulz et al. (2019). Die Autoren entwickeln in dieser Arbeit ein Kompetenzmodell für Operationsverständnis, das von vier Schwierigkeitsstufen («levels») ausgeht. Die vier Schwierigkeitsstufen entsprechen nicht den einzelnen Grundoperationen, doch gehen die Art der jeweiligen Grundoperation und die verschiedenen Grundvorstellungen zu den einzelnen Grundoperationen massgeblich in dieses Modell ein. Die Autoren arbeiten mit dem Rasch-Modell der «Items Response Theory», entwickeln die Schwierigkeitswerte für verschiedenen «Items» und weisen nach, dass diese Schwierigkeitswerte die Einteilung der «Items» in die vier Schwierigkeitsstufen («levels») des Kompetenzmodells untermauern.

Die Multiplikation als «einfachere» der beiden Operationen zweiter Stufe kann − zumindest von einzelnen Kindern − mit frühen denkerischen Mitteln und ohne systematische Erarbeitung im Unterricht verstanden und angemessen bewältigt werden. Padberg (2009, S. 114) nennt verschiedene Studien, die belegen, dass Kinder multiplikative Situationen mit unterschiedlichen Strategien schon früh richtig bearbeiten können. Der Autor belegt «hohe Vorkenntnisse» der Kinder vor Behandlung dieser Grundoperation im zweiten Schuljahr. Er gibt einzelne Strategien an, welche von den Kindern zur Aufgabenlösung verwendet werden. Auch Mulligan & Mitchelmore (1997) berichten davon, dass viele Kinder Multiplikationsaufgaben schon lange vor deren Behandlung im Unterricht lösen können. Die Autoren weisen aber auch darauf hin, dass sie in ihrer Untersuchung festgestellt haben, dass einige Kinder selbst die Division nicht immer als schwieriger wahrnehmen als die Multiplikation.

Bemerkenswert ist es, dass das eigene Urteil der Kinder über ihre mathematischen Fähigkeiten anders ausfallen kann. Die Kinder schätzen gemäss einzelnen Untersuchungen die Subtraktion im Vergleich zur Multiplikation zum Teil als schwieriger ein (Moser Opitz, 2013, S. 260).

Vereinigung der vier Operationen zu einem Konstrukt

In verschiedenen Beiträgen und Forschungsarbeiten sowohl theoretischer als auch empirischer Art werden sämtliche vier Grundrechenarten nacheinander und in aufbauender Reihenfolge besprochen und behandelt. In manchen Arbeiten wird entweder auf die Operationen erster Stufe oder auf die beiden «schwierigeren» Operationen zweiter Stufe fokussiert. Unter letzteren sind auch solche, in denen nur eine dieser beiden Operationen betrachtet wird, also entweder die Multiplikation oder aber die Division.

Nur sehr vereinzelt gab es bisher quantitative Untersuchungen zum Zusammenhang zwischen dem Operationsverständnis für die einzelnen Grundrechenarten. Royar (2013) zeigt z.B. mit Hilfe einer explorativen Faktorenanalyse der Daten eines Testinstruments, dass einzelne zu den vier Grundrechenarten passende Faktoren isoliert werden können. Er stellt dabei aber auch fest, dass damit «nicht völlig unterschiedliche Faktoren erfasst werden». Dies kann als datenbasierte Evidenz dafür gedeutet werden, dass das Konstrukt Operationsverständnis sinnvoll ist.

Ob es in mathematikdidaktischer Hinsicht nützlich ist, die vier Grundrechenarten in ein gemeinsames Konstrukt («unter einen Hut») zusammenzufassen, kann angesichts des Forschungsstandes aber nicht abschliessend beantwortet werden. Zu dieser Frage äussern sich wenige Autoren. An vielen Stellen wird aber auf die unterschiedliche Schwierigkeit und den unterschiedlichen Abstraktionsgrad der Operationen hingewiesen, was den Zusammenzug zu einem gemeinsamen Konstrukt auch als fragwürdig erscheinen lässt.

Die Vorgehensweise, die bei vielen empirischen Untersuchungen gewählt wird, ist die getrennte und meist qualitative Untersuchung der einzelnen Grundrechenarten, häufig mit Hilfe von Interviews mit Kindern. Selbstredend werden die Konstrukte «Operationsverständnis Addition» (oder auch «Operationsverständnis der Addition»), «Operationsverständnis Subtraktion» usw. eingeführt. Das «Operationsverständnis Multiplikation» wird dabei wie in dieser Arbeit auch «Multiplikationsverständnis» genannt (Bönig, 1995; Gaidoschik, 2016a; Schäfer, 2005). Definitionen und Vorgehensweisen werden dabei meist von einer auf die andere Grundrechenart übertragen. In diesem Sinne wird die Fähigkeit zum Darstellungswechsel auf alle vier einzelnen Konstrukte als Indiz für Operationsverständnis übertragen. Zu untersuchen wäre aber sicherlich die Frage, ob diese Fähigkeit bei allen vier einzelnen Konstrukten dieselbe Relevanz hat. Bei eingehender Betrachtung ergeben sich dabei allerlei unerwartete Schwierigkeiten. Ein Beispiel für eine solche Schwierigkeit ist das Gestalten von angemessenen bildlichen Darstellungen zur Subtraktion. Viele Illustrationen zur Subtraktion sind mehrdeutig und können auf verschiedene Arten interpretiert werden. Häufig ist eine sprachliche Ergänzung notwendig, wenn ein Bild auf eindeutige Art und Weise einem Subtraktionsterm zugeordnet werden soll. Zu untersuchen und zu diskutieren wäre also die Frage, inwieweit die Zusammenfassung der vier Konstrukte zu einem einzigen Konstrukt sinnvoll ist.

2.3.2 «Operation Sense» − die andersartige Betrachtung im englischsprachigen Raum

Im englischsprachigen Raum erscheint das Phänomen Operationsverständnis unter der Begrifflichkeit «operation sense». In der Literatur aus diesem Sprachraum wird das Phänomen ebenfalls mittels «Darstellungswechsel» beschrieben und diskutiert. Wie im deutschsprachigen Raum wird auch in der englischen Literatur auf die Vorteile sowohl von Interviews als auch von qualitativen gegenüber quantitativen Untersuchungen hingewiesen (Huinker, 1993). In der Literatur des englischsprachigen Raums wurde − in Analogie zum deutschsprachigen Raum − das Operationsverständnis bisher ebenfalls weniger intensiv erforscht als z.B. der «number sense» (Geary, 1994, S. 43; Berch, 2005; A. Baroody et al., 2006).

Eine begriffliche Differenz ergibt sich durch den Sprachgebrauch, denn die Begrifflichkeit «operation sense» (Operations-Sinn oder -Gefühl) gibt dem Phänomen nicht nur einen anderen Klang, sondern auch weitere Nuancen und zusätzlichen Inhalt. Die Begriffsverwendung in der englischen Sprache gibt dem Konstrukt ausdrücklicher als in der deutschen Sprache die Nuance des «Empfindungsvermögens» und des «Gefühls» (Schäfer, 2005, S. 198). Es klingen in der englischen Sprache Fähigkeiten an, die in der deutschsprachigen Literatur zwar ebenfalls erwähnt werden, jedoch seltener in der Operationalisierung von Operationsverständnis Eingang finden. So werden beispielsweise als Merkmale von Leistungsstärke im Fach Ma-

thematik nebst der «Fähigkeit zum Darstellungswechsel» auch «mathematische Fantasie» oder «mathematische Sensibilität» genannt (Spiegel & Selter, 2013, S. 100 ff.). Derartige Fähigkeiten klingen in der Begrifflichkeit «operation sense» an, während Operationsverständnis in der deutschsprachigen Literatur in stärkerem Ausmass auf den Darstellungswechsel reduziert wird als «operation sense» im englischsprachigen Raum. Besonders deutlich kommt dies dann zum Ausdruck, wenn der Darstellungswechsel nicht nur als Indiz für Operationsverständnis genannt wird, sondern dieses als die Fähigkeit zum Darstellungswechsel definiert wird (Gerster & Schultz, 2004). Es kann sein, dass auch die enge Verwandtschaft zum Konzept der Grundvorstellungen zu dieser eingeschränkteren Verwendung im deutschsprachigen Raum beigetragen hat. Auch fehlt in der deutschen Sprache ein geläufiger Begriff, mit dem «operation sense» übersetzt werden könnte, was zusätzlich zu einer etwas anderen inhaltlichen Ausgestaltung der Begriffe in den beiden Sprachräumen beigetragen haben dürfte.

Operationsverständnis wird bei der Übersetzung aus dem deutschsprachigen Raum in die englische Sprache trotz diesen Nuancen wie zum Beispiel in Schulz et al. (2019) typischerweise mit «operation sense» übersetzt und die beiden Begriffe somit zumindest in übersetzten Arbeiten aus dem deutschsprachigen Raum als Synonyme verwendet.

Übersetzungsschwierigkeiten kommen insbesondere auch beim Begriff der Grundvorstellungen vor, der in der englischsprachigen Literatur nicht verwendet und anscheinend auch nicht vermisst wird. Unterschiedliche Situationen werden in englischer Sprache entweder als «situations» oder als «categories» beschrieben. Die Einteilung in diese Kategorien kann sich von derjenigen im deutschsprachigen Raum unterscheiden (Schulz et al., 2019; Sun et al., 2019, S. 5).

1. A conceptualization of the base components of the process.
2. Familiarity with properties which the operation is able to possess.
3. Relationship with other operations.
4. Facility with the various symbol systems associated with the operations.
5. Familiarity with operation contexts.
6. Familiarity with operation facts.
7. Ability to use the operation without concrete or situational referents.
8. Ability to use the operation on unknown or arbitraty inputs.
9. An ability to relate the use of the operation across different mathematical objects.
10. An ability to move back and forth between the above conceptions.

Abbildung 12: Die zehn Aspekte von «operation sense» nach (Slavit, 1999)

Bei Betrachtungen im deutschsprachigen Raum ist es zum Teil hilfreich, sich diese Nuance – den Operations-«Sinn» − und die damit verbundenen weicheren Fähigkeiten der «Sensibilität» bzw. des «Gefühls», die in der englischen Sprache explizit anklingen, zusätzlich mitzudenken. In englischsprachigen Publikationen wird dies zum Teil sehr deutlich gemacht. Slavit (1999) berichtet von einer Untersuchung mit 6- bis 9-jährigen Kindern zum «operation sense». Der Hauptfokus der Untersuchung lag dabei auf der Addition. In dieser Arbeit ist auch eine Aufzählung von zehn Aspekten des «operation sense» enthalten, die dazu dienen sollen, die Bedeutung des Begriffs zu erfassen. Die genannten Aspekte werden in dieser Arbeit ausführlich erläutert. Selbst diese Erläuterungen sind aber mit der Vorbemerkung versehen, dass über jeden einzelnen der zehn Aspekte noch viel mehr gesagt werden könnte. Der Autor kommentiert die Liste mit der abschliessenden Bemerkung «the above list should not be considered exhaustive, and there is much more to operation sense than a list, no matter how exhaustive it may be» (Slavit, 1999, S. 254 ff.).

Die sprachliche Ähnlichkeit der Begriffe «number sense» und «operation sense» bekommt im englischsprachigen Raum also auch inhaltliche Bedeutung, die in der deutschen Sprache so nicht mehr anklingt. Zum Teil werden die beiden Begriffe in der englischen Literatur gemeinsam behandelt. «Number and operation sense» werden gemeinsam als ein intuitives Gefühl für das Zusammenspiel von Zahlen («an intuitive feel for how numbers are related and behave») genannt (A. Baroody et al., 2006, S. 188). Das Konstrukt «operation sense» erhält dadurch eine stärker ausgeprägte Weite, wobei auch mitspielt, dass der Begriff «number sense» in der englischen Sprache auf vielfältige Art und Weise gebraucht wird. In Purpura & Lonigan (2013) wird beispielsweise davon berichtet, dass für die Beschreibung von «number sense» eine Aufzählung von über 30 verschiedenen Fähigkeiten besteht.

Aufgrund dieser fehlenden Einheit unter Forschern ist somit auch die Vergleichbarkeit von Ergebnissen aus unterschiedlichen Studien nicht immer gegeben bzw. nur in eingeschränktem Masse sinnvoll.

Zu beachten bleibt ausserdem, dass nicht nur bei der Begriffsverwendung des Operationsverständnisses, sondern auch bei Untersuchungen der Entwicklung von Rechenstrategien grosse Differenzen zwischen Sprachräumen auftreten können (Gaidoschik, 2010, S. 18, 52 ff.).

Ergebnisse aus Untersuchungen aus anderen Sprachräumen können folglich nicht ohne Vorbehalte und Einschränkungen auf den deutschsprachigen Raum übertragen werden. Umso wichtiger erscheint es aber, dem Konstrukt «Operationsverständnis» und der Thematik auch im deutschsprachigen Raum ausreichend Aufmerksamkeit zu schenken. Die in dieser Arbeit vorgenommene Untersuchung soll dazu beitragen, die im deutschsprachigen Raum bestehende Forschungslücke im Bereich der quantitativen Untersuchungen zu dieser Thematik zu verkleinern.

2.3.3 Das Konstrukt Operationsverständnis – eine Zusammenfassung

In den bisherigen Kapiteln wurden die Begriffe erläutert, welche in Zusammenhang mit dem Konstrukt Operationsverständnis verwendet werden. Verschiedene Konstrukte und Konzepte werden mit Operationsverständnis in Zusammenhang gebracht und das Konstrukt wird in der Literatur unterschiedlich beschrieben. Besonders deutlich wird dies ersichtlich bei einem Vergleich mit der Verwendung des Begriffs «operation sense» im englischsprachigen Raum. Allerdings unterscheidet sich die Begriffsverwendung auch innerhalb der deutschsprachigen Literatur. An dieser Stelle wird das Konstrukt Operationsverständnis noch einmal zusammenfassend umschrieben und sodann für die weitere Verwendung in dieser Arbeit entsprechend verwendet.

Operationsverständnis bezeichnet ein Grundverständnis für die vier Grundrechenarten, welches die Kinder in den ersten drei Primarschuljahren erwerben und aufbauen sollten. Dieses Grundverständnis kann erfasst werden durch die Fähigkeit der Kinder, flexibel und gut begründet zwischen verschiedenen Darstellungsarten der Grundrechenarten hin und her zu wechseln. Das Konstrukt Operationsverständnis bezeichnet daher genau diese Fähigkeit.

Für die Erfassung von Operationsverständnis sind aber diejenigen Darstellungswechsel besonders wesentlich, in denen eine Darstellungsart in mathematisch-symbolischer Schreibweise (in Form eines Terms) vorliegt. Deshalb werden für die Konstrukt-Beschreibung nur diejenigen Darstellungswechsel herangezogen, bei denen entweder die Ausgangsdarstellung (Darstellungswechsel «vom Term weg») oder die Enddarstellung (Darstellungswechsel «zum Term hin») eine mathematisch-symbolische Darstellung ist.

Bei diesen Darstellungswechseln ist darauf zu achten, dass vielfältige und unterschiedliche Darstellungen verwendet werden und die Kinder den Darstellungswechsel entsprechend flexibel vollziehen können. Dies bedeutet beispielsweise, dass sie Ausgangssituationen unterschiedlicher Art erfassen können. Während die Fähigkeit zum Darstellungswechsel für sich alleine formuliert noch keine inhaltlichen Differenzen zwischen Ausgangsdarstellungen aufgreift, werden solche durch das Konzept der Grundvorstellungen auch inhaltlich beschrieben. Die Kinder sollten für die Multiplikation beispielsweise in der Lage sein, nicht nur räumlich-simultane, sondern auch zeitlich-sukzessive Ausgangssituationen in ihrer multiplikativen Struktur zu erfassen und in entsprechende Terme zu übersetzen. Die Unterscheidung dieser Ausgangssituationen wird in der deutschsprachigen Literatur mit dem Konzept der Grundvorstellungen in Verbindung gebracht. Die Kinder sollten also in der Lage sein, den Darstellungswechsel bei Ausgangssituationen zu den verschiedenen Grundvorstellungen für die vier Grundrechenarten vollziehen zu können. Auch sollten sie in der Lage sein, Veränderungen an der Ausgangssituation in Veränderungen am passenden Term zu übersetzen. Sie sollten nicht nur die veränderte Ausgangssituation (z.B. das Punktefeld mit einer zusätzlichen Spalte) als gänzlich neue Situation erneut «übersetzen», sondern den Veränderungsvorgang in eine Veränderung am Term (z.B. «einen Faktor um 1 erhöhen») übersetzen können. Diese Art der Strukturnutzung und des Umgangs mit strukturierten Ausgangssituationen wird in der Literatur mit dem Konstrukt der Strukturierungsfähigkeit beschrieben. Die Kinder sollen also Darstellungswechsel so vollziehen können, dass sie auch Strukturierungen anwenden und für ihre Übersetzungsprozesse zwischen Darstellungen nutzen können.

Das Konstrukt Operationsverständnis bezeichnet also die Fähigkeit, zwischen Darstellungen (von denen eine in mathematisch-symbolischer Schreibweise vorliegt) flexibel zu wechseln und dabei Situationen zu unterschiedlichen inhaltlichen Grundvorstellungen bearbeiten sowie auch Strukturierungsfähigkeiten einsetzen und nutzen zu können.

Diese Formulierung erfolgt im Bewusstsein, das der Begriff «Verständnis» immer – so wie auch im Zusammenhang mit Operationsverständnis − viele Nuancen und Facetten aufweist und die entsprechende Formulierung in mancher Hinsicht auch starr erscheinen mag. Deutlich ersichtlich wird dies bei einem Vergleich mit den «weiteren» Beschreibungen, wie sie im englischsprachigen Raum verwendet und in Kapitel 2.3.2 ausgeführt wurden.

Die Formulierung erlaubt aber eine verhältnismässig klare Operationalisierung des Konstrukts, welche die Erfassung für ein Erhebungsinstrument ermöglicht. Dies erlaubt eine Umsetzung, wie sie in Kapitel 4.4 beschrieben wird.

Ebenso lassen sich mit Hilfe der vorliegenden Formulierung für das Konstrukt Operationsverständnis entsprechende Unterrichts-Interventionen (in dieser Arbeit in Form von Fördereinheiten) zielgerichtet entwickeln und einsetzen. Bei der Beschreibung des Aufbaus des Förderpakets in Kapitel 4.3 wird deshalb auf die hier vorgenommene Operationalisierung zurückgegriffen. Die besonderen Betonungen der einzelnen Aspekte (Darstellungswechsel, Grundvorstellungen bzw. Strukturierungsfähigkeiten) des Konstrukts werden dort der besseren Übersicht und Verständlichkeit halber auch farblich dargestellt.

Die oben gewählte Formulierung des Konstrukts Operationsverständnis sowie die Verwendung von Darstellungswechseln, Grundvorstellungen und Strukturierungsfähigkeiten ermöglichen also eine Umsetzung bzw. Operationalisierung des Konstrukts in einer Studie, wie sie in dieser Arbeit nun beschrieben wird.

3 Operationsverständnis – Bedeutung, Befunde und Förderung

In diesem Kapitel erfolgt eine Zusammenstellung von Erkenntnissen zum Operationsverständnis. Das Konstrukt wird beleuchtet in seiner Bedeutung für das Erlernen der Mathematik bzw. des Rechnens in der Schule. Beleuchtet werden bisher vorliegende Befunde, Schwierigkeiten der Kinder und Möglichkeiten zur Förderung.

3.1 Operationsverständnis in seiner Bedeutung für das Rechnenlernen

Der Aufbau eines tragfähigen Operationsverständnisses stellt für einige Kinder eine bedeutend grössere Herausforderung dar als für andere. Besonders leistungsstarke Kinder meistern diese Thematik in kurzer Zeit und mit einer Leichtigkeit, die dazu führen kann, dass die Schwierigkeiten anderer Kinder übersehen werden. Es besteht eine Versuchung für die Lehrpersonen, die leistungsstarken Kinder im Unterricht nicht zu langweilen, was dazu führen kann, dass ernsthafte Lücken beim Operationsverständnis bei anderen Kindern nicht angemessen behandelt werden.

Schwächen beim Operationsverständnis sind im Kontext von Rechenschwächen zu beleuchten. Die Themenkreise, in denen sich Rechenschwäche äussern kann, sind vielfältig. Die Schwierigkeiten der Kinder sind häufig vernetzt und äussern sich in unterschiedlichen Bereichen der Primarschulmathematik. Das Operationsverständnis ist eines von mehreren Teilgebieten, das im Hinblick auf Rechenschwächen bei den Kindern Beachtung finden sollte.

In diesem Kapitel soll einerseits beleuchtet werden, welches die Voraussetzungen für den Aufbau eines angemessenen Operationsverständnisses sind. Andererseits soll die Thematik im Kontext von Rechenschwächen beleuchtet werden. Aus diesen Betrachtungen ergibt sich auch die hohe didaktische sowie gesellschaftliche Relevanz der Thematik dieser Arbeit.

3.1.1 Voraussetzungen für Operationsverständnis

Am Operationsverständnis wird mit den Kindern im Mathematikunterricht in den ersten Primarschuljahren gearbeitet. Die Bearbeitung der Grundoperationen, auch schon derjenigen der ersten Stufe, erfordert von den Kindern aber denkerische Leistungen, die nicht notwendigerweise mit dem Schuleintritt schon erbracht werden können. Bei jeglicher Förder- und Unterrichtstätigkeit ist es aber relevant, die Lernenden bzw. die Kinder dort abzuholen, wo sie stehen, und gegebenenfalls auch die fehlenden Voraussetzungen für den Lernerfolg zu erkennen und anzugehen (Gaidoschik, 2016b).

Einige der dafür relevanten Themenfelder werden hier aufgeführt. Auch wird deren Bedeutung für das Operationsverständnis skizziert.

Umfassendes Zahlverständnis

Für die Arbeit am Operationsverständnis ist u.a. ein umfassendes Zahlverständnis bei den Kindern eine wichtige Voraussetzung. Die Kinder sollten in Zahlen jeweils auch den Kardinalzahlaspekt erkennen und den Zählprozess nicht nur als ein «Korrektes Aufsagen» einer Zahlenreihe verstehen (Gaidoschik, 2016b). Grundlegende Zahl- und Operationsvorstellungen werden in der Literatur auch als «Basis» für verständnisvolles Rechnen genannt (Büchter et al., 2019; Häsel-Weide & Nührenbörger, 2018). Dass solche Voraussetzungen nicht selbstverständlich sind und vor allem auch bei Kindern mit intellektueller Beeinträchtigung eine Hürde darstellen können, wurde schon verschiedentlich beobachtet (Moser Opitz, 2013, 2016).

Teil-Teil-Ganzes-Konzept

Die Kinder müssen für die Bewältigung der frühen Primarschulmathematik lernen und verstehen, dass eine Anzahl (als Ganzes) aus einzelnen Teilen zusammengesetzt sein kann. Diese Einsicht wird in der Literatur als Teil-Teil-Ganzes-Konzept (TTG-Konzept) bezeichnet. Die Ausprägung des Teil-Teil-Ganzes-Konzepts bei den Kindern ist wesentlich bzw. notwendig für weitere Fortschritte in den mathematischen Leistungen (Focke, 2004; Weisshaupt & Peucker, 2009; Gaidoschik, 2010, S. 115 ff.; Langhorst et al., 2011; Benz, 2018).

Dieses Grundverständnis und entsprechende Strukturierungsfähigkeiten der Kinder werden als Voraussetzung für die Ausbildung von Operationsverständnis angesehen. Ein strukturiertes Sehen der Kinder wird als Grundlage für weitere mathematische Inhalte in der Literatur aufgeführt und nachgewiesen (Beutler, 2013; Benz, 2018).

In mathematikdidaktischen Forschungsarbeiten wurde verschiedentlich dargelegt, dass für die Entwicklung des Operationsverständnisses die alleinige Arbeit an den Grundoperationen mit den Kindern nicht ausreichend ist, sondern die wesentlichen Grundlagen für deren Verständnis vorgängig ebenfalls gelegt sein müssen. Zu eben diesen Grundlagen zählt nebst dem umfassenden Zahlverständnis u.a. auch das TTG-Konzept (Fritz et al., 2018, S. 36).

Rechenstrategien und Ablösung vom zählenden Rechnen

Auf den Zusammenhang zwischen Operationsverständnis und der Rechen- bzw. Zählstrategie der Kinder wird in der Literatur verschiedentlich hingewiesen, so z.B. in Schäfer (2005). In der Literatur wird ein Anteil an Kindern von 20% bis 33% genannt, der sich mit dieser Ablösung schwer tut (Rechsteiner-Merz, 2015, S. 58).

Die Ablösung vom zählenden Rechnen wird zu den Haupthürden im Mathematikunterricht der ersten Primarschuljahre gezählt. Der Schritt vom zählenden Rechnen zu raffinierteren Rechenstrategien ist eng verbunden mit dem TTG-Konzept und eine der Voraussetzungen für verständiges Rechnen, auch für die Addition als einfachste der vier Grundrechenarten. Eine Aufgabe des frühen Mathematikunterrichts ist es, die Kinder dabei zu unterstützen, flexible und nicht-zählende Rechenstrategien zu erwerben (Benz, 2018).

Das zählende Rechnen und die mangelhafte Ablösung von dieser Rechenstrategie erschwert u.a. auch ein angemessenes Verständnis für das Stellenwertsystem (Büchter et al., 2019; Wittich, 2017). Dieses Verständnis ist aber wiederum eine der Voraussetzungen für den weiteren arithmetischen Kompetenzaufbau (Gaidoschik, 2015b). Erst die Ablösung vom zählenden Rechnen macht also verständiges Rechnen in den Grundoperationen und somit auch Operationsverständnis möglich (Pfister, 2016, S. 24 f.).

Verwandtschaft mit anderen Themenfeldern der Primarschulmathematik

Beim Zusammenspiel des Operationsverständnisses mit den Rechenstrategien der Kinder klingt an, wie eng die Verwandtschaft zwischen dem Konstrukt Operationsverständnis und verschiedenen anderen Themenfeldern der Primarschulmathematik ist. Dieses enge Zusammenspiel wird auch im kommenden Kapitel noch klarer ersichtlich werden, denn gerade auch bei Rechenschwächen der Kinder wird die Verwandtschaft der verschiedenen Themengebiete in der Mathematik sichtbar.

Besonders deutlich tritt dieses Zusammenspiel auch in der englischsprachigen Literatur auf, in der Aspekte des Operationsverständnisses zum Teil nicht mit dieser Begrifflichkeit beschrieben werden und in andere Konstrukte einfliessen (siehe auch Kapitel 2.3.2). Aunio & Räsänen (2016) ist eine Arbeit, die diese Schwierigkeit unterstreicht. Die Autoren untersuchen die wichtigsten Faktoren bei der Entwicklung der mathematischen Fähigkeiten von Kindern im Alter zwischen fünf und acht Jahren. Sie unterscheiden dabei vier wesentliche Faktoren, nämlich «number sense», «understanding mathematical relations», «counting skills» und «basic skills in arithmetics». Sie versprechen sich von dieser Einteilung eine bessere Diagnose der Kinderfähigkeiten («assessment of numerical skills of children») sowie Hilfestellungen bei der Unterrichtsgestaltung. Wenn auch das Operationsverständnis sich in dieser Einteilung am ehesten in den «basic skills in arithmetics» wiederfindet, so gibt es doch zumindest keine klare Zuordnung zu dieser Fähigkeit der Kinder. Auch das Verständnis mathematischer Relationen (so wie es die Autoren beschreiben) umfasst Aspekte des Operationsverständnisses. Auch gibt es inhaltliche Überschneidungen mit den «counting skills». Die Autoren beschreiben selbst in ihrer Arbeit auch diese Beziehungen zwischen den einzelnen Fähigkeiten (Aunio & Räsänen, 2016, S. 15 f.).

3.1.2 Operationsverständnis und Rechenleistung

In bisherigen Untersuchungen gab es Indizien dafür, dass die Fähigkeit zum Darstellungswechsel nicht nur mit dem Verständnis der Kinder, sondern auch mit deren Rechenleistung in Zusammenhang steht (Moser Opitz, 2005; Gaidoschik, 2010, S. 95 ff.; Ladel, 2011a; Royar, 2013). Es gab aber auch Studien, die diese Vermutung nicht oder nur teilweise belegen konnten (Freesemann, 2014). Verschiedentlich wird auch darauf hingewiesen, dass die Kinder ein mangelndes Operationsverständnis zu verbergen wissen und sich dieses daher nicht – oder zumindest nicht unmittelbar − beim Rechnen bemerkbar macht.

Auf eine Reihe von Schwierigkeiten beim Umgang mit den beiden Operationen zweiter Stufe wird auch in Moser Opitz (2013, S. 111 ff., 197 ff.) hingewiesen. Die empirische Untersuchung in dieser Arbeit offenbart «Widersprüchliches» zur Multiplikation. Einerseits wurden Aufgaben des kleinen Einmaleins gut gelöst, andererseits zeigten sich aber Mängel beim Multiplikationsverständnis (beim «Mathematisieren»). Als Erklärung hierfür werden der Zugang zur Multiplikation über die Summe gleicher Summanden und auch die möglicherweise mangelnde Einsicht ins Stellenwertsystem angegeben.

In manchen Arbeiten zum Einmaleins wird aber auf die Notwendigkeit eines vorgängig bereits erarbeiteten und gefestigten Verständnisses für die Operation bzw. auf die Notwendigkeit für den vorgängigen Aufbau von Grundvorstellungen verwiesen (Smith & Smith, 2006; Scherer & Moser Opitz, 2010, S. 117 ff.; Gaidoschik, 2015a, 2016a).

Das Operationsverständnis im Kontext von Rechenschwäche

Erarbeiten sich die Kinder in der Einführungsphase der jeweiligen Grundoperation kein angemessenes Operationsverständnis und bildet sich dieses auch im Laufe des späteren Unterrichts nicht aus, so kann es sein, dass den Kindern das spätere Lernen von Mathematik schwerer fällt (Gaidoschik, 2016a; Sun et al., 2019).

Manche Kinder sind in der Lage, ein mangelndes Grundverständnis für die Rechenoperationen zu verbergen. Ein Weg dies zu tun, besteht darin, die Ergebnisse von Rechenaufgaben auswendig zu lernen. Die Kinder können so zu vielen Rechenaufgaben das korrekte Ergebnis nennen, was zum Beispiel bei Aufgaben des kleinen Einmaleins beobachtet werden kann. Nicht immer ist dieses Lösen von Rechenaufgaben aber mit einem intakten Grundverständnis für die Bedeutung der Operation verbunden (Moser Opitz, 2013).

Bleiben solche Lücken im Grundverständnis der Kinder unentdeckt, so besteht die Gefahr, dass sich die Mängel bei den Kindern verfestigen und dass sich Rechenschwächen ausbilden. Das rezepthafte Auswendiglernen von Ergebnissen ohne Grundverständnis wird die Kinder beim Problemlösen immer wieder vor Schwierigkeiten stellen. Insbesondere bei mehrstufigen Aufgaben werden die Strategien der betroffenen Kinder nicht mehr zum Ziel führen (Moser Opitz, 2005).

In Sun et al. (2019) wird berichtet, wie Mängel im Umgang mit den Grundoperationen auch Rechenleistungen in späteren Gebieten der Mathematik negativ beeinflussen. Als betroffene

Themengebiete werden das Bruchrechnen, das Verständnis für Brüche und Proportionen sowie die Algebra genannt. In derselben Arbeit wird ausgeführt, dass das Operationsverständnis und die Fähigkeit, entsprechende Textaufgaben in arithmetische Operationen umzuwandeln, eine erste Gelegenheit für die Kinder darstellt, den Transfer zwischen realen Situationen und der abstrakten symbolischen Welt der Mathematik zu vollziehen. Gelingt den Kindern diese Transferleistung, so ist damit auch für spätere − ähnlich gelagerte − Anforderungen innerhalb der Mathematik ein Baustein gelegt. Mit einer entsprechenden Förderung kann also ein Fundament gelegt werden, von dem die Kinder auch später und in komplexeren Gebieten der Mathematik noch profitieren können.

Mögliche Folgen von Rechenschwächen

Rechenschwächen können nicht nur das spätere Lernen von Mathematik und das Gelingen der späteren schulischen Laufbahn erschweren (Rittle-Johnson, 2017; Schipper, 2002; Sun et al., 2019), sondern sie stehen auch mit Problemen im späteren Leben in Zusammenhang (Ise et al., 2012; Gaidoschik, 2016a). Es wird berichtet von Schwierigkeiten der Betroffenen im Umgang mit den persönlichen Finanzen, vom Misslingen der begonnenen beruflichen Ausbildung, von Einschränkungen bei der Berufswahl und Schwierigkeiten im Berufsleben oder auch von einer Korrelation zwischen Leistungsfähigkeit im Fach Mathematik und späterem Berufserfolg, Höhe des späteren Einkommens oder sonstigem Erfolg im Leben (Geary, 1994; Purpura & Lonigan, 2013; Pfister, 2016; Rittle-Johnson, 2017; Gleissberg, 2018; Moser Opitz, 2005; Ehlert et al., 2013).

Auch wird verschiedentlich berichtet von möglichen psychischen Folgen von Rechenschwächen bzw. von einem negativen Einfluss auf das Selbstvertrauen und das mathematische Selbstbild der betroffenen Kinder (Spiegel & Selter, 2013; Gaidoschik, 2015a). Es wird in der Fachliteratur aber auch darauf hingewiesen, dass viele solche Aussagen auf Alltagsbeobachtungen und Praxiserfahrungen beruhen und nicht immer wissenschaftliche Folgerungen und Verallgemeinerungen zulassen (Moser Opitz, 2013).

Trotz dieser einschränkenden Bemerkung zeigen aber die denkbaren Folgen von Rechenschwächen bei Kindern die hohe Relevanz der Thematik in eindrücklicher Weise auf. Diese hohe Relevanz rechtfertigt Arbeiten, die darauf abzielen, die Thematik Rechenschwäche besser zu verstehen und zu erfassen. In dieser Arbeit wird nun das Operationsverständnis erforscht, welches − wie in diesem Kapitel dargelegt – einen Themenkomplex darstellt, der mit Rechenschwächen von Kindern in Zusammenhang steht.

Es wurden mit den Fördereinheiten (siehe Kapitel 4.3) Ansätze zur Behebung bzw. Vermeidung von Lücken entwickelt, die sich beim Multiplikationsverständnis der Kinder sonst einstellen könnten. Auch wurde die Wirksamkeit dieser Fördereinheiten in Form von Unterrichtsinterventionen erfasst und quantifiziert (siehe Kapitel 6).

Die Relevanz der Thematik rechtfertigt aber sicherlich auch weitere Arbeiten, in denen andere, aber verwandte Themen auf ähnliche Art und Weise untersucht werden.

Operationsverständnis als eine Hürde im frühen Mathematikunterricht

Als Hauptthemenfeld für Rechenschwierigkeiten wird in der Primarschulstufe die mangelhafte Ablösung vom zählenden Rechnen beschrieben (Schipper, 2002, 2005; Scherer & Moser Opitz, 2010; Gaidoschik, 2010).

Schwächen beim Operationsverständnis gehören aber – wie z.B. auch Schwierigkeiten beim Verständnis für das Stellenwertsystem − zu den weiteren Themenfeldern, in denen sich eine frühe Rechenschwäche des Kindes manifestieren kann (Akinwunmi & Deutscher, 2014; Focke, 2004; Freesemann & Breucker, 2014; Gaidoschik, 2015a, S. 36 ff.; Ise et al., 2012; Kuhnke, 2013, S. 17; Pfister et al., 2015; Scherer & Moser Opitz, 2010; Schipper, 2002). In vielen dieser Arbeiten wird das Operationsverständnis als Themenfeld genannt bzw. die Thematik der Rechenschwäche erwähnt, um die Wichtigkeit und die Bedeutung der Förderung von Operationsverständnis zu untermauern. Auch in Stöckli (2019) wird davon ausgegangen, dass Kinder mit einer Rechenschwäche über ein eingeschränktes Operationsverständnis verfügen.

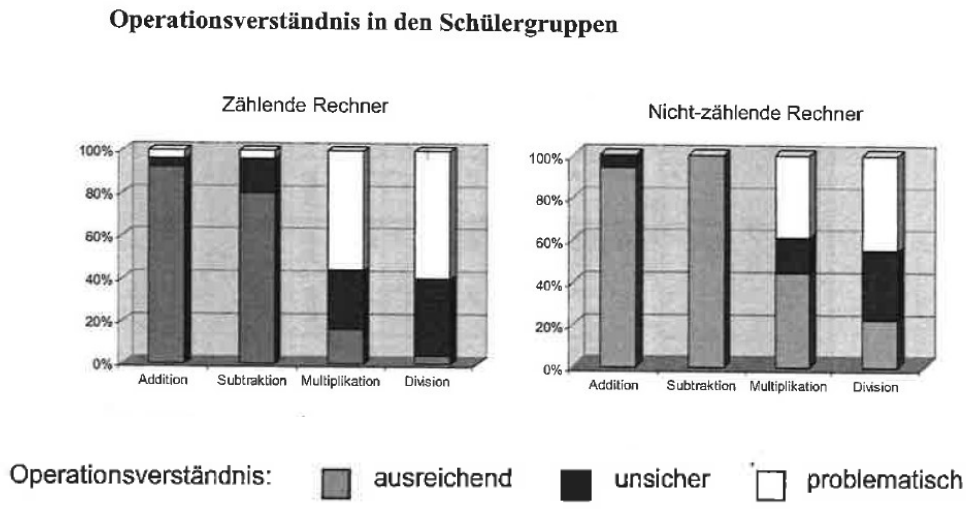


Abbildung 13: Operationsverständnis und Schülergruppen (zählend/nicht-zählend) aus (Schäfer, 2005)

Die Themengebiete der Rechenschwierigkeiten betroffener Kinder sind untereinander verknüpft und greifen ineinander, so dass auch das Operationsverständnis im Zusammenhang mit andern Themenfeldern der frühen Schulmathematik beleuchtet und untersucht werden sollte (Gaidoschik, 2015a). In einzelnen Arbeiten wurde auch nachgewiesen, dass das Operationsverständnis mit dem zählenden Rechnen in Zusammenhang steht bzw. dargelegt, auf welche Art das umfassende Zahlverständnis auch beim Operationsverständnis eine Rolle spielt (Gaidoschik, 2010, 2015a; Venkat & Mathews, 2019). Schäfer (2005) stellt diesen Zusammenhang anhand der Ergebnisse einer empirischen Studie graphisch dar (siehe Abbildung 13).

Manche Studien weisen im Übrigen nach, dass in den verschiedenen Bereichen, in denen Rechenschwächen sich typischerweise manifestieren, durch entsprechende Förderung auch im Klassenverband erfolgreich mit den Kindern gearbeitet werden kann. Die entsprechenden

Schwächen der Kinder im Fach Mathematik können somit auch durch Ausgestaltung des Unterrichts behandelt werden (Pfister et al., 2015).

Anteil betroffener Kinder

Die Angaben zum Anteil rechenschwacher bzw. hochgradig rechenschwacher Kinder in der Literatur sind unterschiedlich. Die meisten Autoren machen aber Angaben im tiefen zweistelligen Prozentbereich. Gaidoschik (2015a) schätzt den Anteil hochgradig rechenschwacher Kinder auf 6%, den Anteil förderungsbedürftiger Kinder auf 15%. Spiegel & Selter (2013) quantifizieren den Anteil der Kinder mit diagnostizierten Rechenstörungen mit 2% bis 20%. Die Angaben sind verbunden mit dem Hinweis, dass diese Zahlen andernorts anders geschätzt werden und auch abhängig seien von der verwendeten Definition einer Rechenstörung bzw. Rechenschwäche. Auf die Unterschiedlichkeit der Zahlenwerte beim Studium verschiedener Quellen wird auch in Pfister (2016) hingewiesen.

Als Erklärung für die unterschiedlichen Angaben sind nebst Unterschiedlichkeiten zwischen Sprachräumen oder natürlichen Schwankungen zwischen verschiedenen Untersuchungen auch unscharf definierte Begriffe anzuführen. In Moser Opitz (2013) wird auf die unterschiedliche Verwendung der Begrifflichkeiten in der Fachliteratur verwiesen. Die Verwendung des Begriffs «Rechenschwäche» variiert in der Literatur derart stark, dass einzelne Didaktiker auch schon vorgeschlagen haben, gänzlich auf den Begriff zu verzichten (Schipper, 2002).

Royar (2011) nennt einen Anteil betroffener Kinder fürs Operationsverständnis. Bis zu 20% der untersuchten Kinder in dritten Schulklassen seien nicht in der Lage, einfachsten Rechenausdrücken eine Bedeutung zuzumessen, die von Experten für adäquat gehalten wird.

Es wird in der Literatur auch darauf hingewiesen, dass die genauen Prozentzahlen weniger relevant seien als die Frage, wie denn die jeweiligen Kinder in optimaler Weise zu fördern seien.

3.1.3 Rechenschwäche und Operationsverständnis als Folge vielfältiger Einflüsse

Als mögliche Ursachen für Rechenschwäche werden nebst der Ausgestaltung des Unterrichts auch vielfältige andere Faktoren genannt. Die Gründe für Rechenschwächen bei einzelnen Kindern liegen also mit Sicherheit nicht nur im Unterricht, aber auch nicht nur beim Kind selbst. Es wird berichtet von genetischer Disposition bei den Kindern, von individuellen Unterschieden beim Gehirn und bei den Sinnesorganen und Nervenbahnen, von frühkindlichen Hirnfunktionsstörungen sowie von neurologischen und anderen Erkrankungen als Einflussfaktoren (Gaidoschik, 2015b, 2016b, S. 9; Jacobs & Petermann, 2003). Der schulische Unterricht hat aber dennoch hohe Relevanz (Gaidoschik, 2010, S. 55 f., 2016b). Mit gutem schulischen Unterricht kann einer allfälligen Rechenschwäche vorgebeugt oder aber eine solche frühzeitig erkannt und behandelt werden. Auch ist belegt, dass die Leistungen der Kinder mit dem Fachwissen der Lehrpersonen in Zusammenhang stehen und dass unterrichtsbezogene Faktoren wie zum Beispiel die Zeitdauer der Behandlung des Operationsverständnisses oder

anderer Themen einen erheblichen Einfluss auf die Entwicklung mathematischer Kompetenzen v.a. der rechenschwachen Kinder haben können (Gaidoschik, 2016b; Moser Opitz, 2009; Wittich, 2017, S. 43).

In der Literatur werden die Gründe für Rechenschwäche den drei Faktoren Schule («didaktische Risikofaktoren»), Familie («familiäre und soziale Risikofaktoren») und Kind («individuelle Risikofaktoren») zugeordnet (Gaidoschik, 2015a; Schipper, 2002; Spiegel & Selter, 2013). Innerhalb dieser drei Faktoren werden wiederum verschiedene Komponenten aufgeführt, die alle einen Einfluss haben können. Beispielsweise können unter dem Risikofaktor Schule die folgenden Komponenten eine negative Wirkung entfalten: Missachtung des Lernstands, Geschwindigkeit des Unterrichts, negative Verstärkungen, ungünstige Klassenstruktur, Notendruck, Lehrpersonenwechsel usw. (Gaidoschik, 2015a). Andernorts wird diese Aufzählung ergänzt durch die Ursachenfelder Lehrkraft, Unterrichtsmethode, Umgang mit Material, Lehrbuch, Mitschüler, Sprache und Gespräche auf der Meta-Ebene, Förderunterricht, Lehrerausbildung (Schipper, 2002).

Wie bereits dargelegt sind auch die Kombinationen, in denen das Operationsverständnis verbunden mit anderen Schwächen auftritt, vielfältig und daher nicht einfach zu erforschen. In der Literatur wird festgehalten, dass Rechenschwächen bei den Kindern sehr unterschiedlich und individuell in Erscheinung treten können und dass es letztlich so viele Rechenschwächen gibt wie betroffene Kinder (Gaidoschik, 2015a).

Im Folgenden werden viele weitere Einflussfaktoren, welche auf das Operationsverständnis der Kinder einwirken, besprochen und dargelegt. Es zeigt sich, dass nicht nur Rechenschwächen, sondern auch das Operationsverständnis komplexe Konstrukte sind, die von vielerlei äusserst verschiedenen Faktoren beeinflusst werden.

Unterschiede zwischen Schulklassen

Vereinzelt wurde festgestellt, dass hinsichtlich Operationsverständnis signifikante Unterschiede zwischen Schulklassen bestehen (Royar, 2013). Diese Beobachtung lässt vermuten, dass die Art der Behandlung der Grundoperationen im Unterricht einen wesentlichen Einfluss auf die Entwicklung und Ausprägung des Operationsverständnisses der Kinder hat. Dies ist eine These, die auch andernorts in der Literatur zu finden ist (Gaidoschik, 2016b).

Unterschiede zwischen Schulen und zwischen Schulklassen nennt auch Gleissberg (2018) mit Bezug auf verschiedene Examensarbeiten.

Ruwisch (2001, S. 179 ff.) beobachtete Klasseneffekte, als Kinder zu einem vorgegebenen Multiplikationsterm ein Bild gestalten mussten. Die Unterschiede zwischen den Klassen waren in dieser Untersuchung grösser als jene zwischen den Jahrgangsstufen. Bei der Interpretation dieser Beobachtung wird auf den starken Einfluss der Unterrichtsgestaltung hingewiesen.

Auch Moser Opitz (2013, S. 282 f.) berichtet, dass ein Teil der Leistung der Kinder durch die Schulklassenzugehörigkeit erklärt wird. Ebenso berichtet Stern (1997) von Leistungsunter-

schieden in den arithmetischen Kompetenzen, die durch die Zugehörigkeit zur Schulklasse erklärt werden.

Zusammenhang mit dem Fachwissen der Lehrperson

Lernfortschritte der Kinder im Unterricht wurden verschiedentlich schon mit dem fachlichen Wissen oder auch mit den didaktischen Fähigkeiten ihrer Lehrpersonen in Zusammenhang gebracht (Ball et al., 2005; Krauss et al., 2008; Jordan et al., 2008; Moser Opitz, 2009; Baumert et al., 2010; Kreis et al., 2012; Kunter et al., 2013; Moser Opitz, 2013; Krauthausen & Scherer, 2014). Der Einfluss der Lehrperson kann als ein (wenn auch nicht als einziger) erklärender Faktor für die oben erwähnten Unterschiede zwischen Schulklassen dienen.

Einfluss des Geschlechts

Da in der Schweiz verschiedentlich schon Unterschiede zwischen den Geschlechtern im Hinblick auf das Lernen von Mathematik beobachtet wurden (Moser Opitz, 2013, S. 179), soll auch diese Variable in der hier vorliegenden Arbeit am Beispiel des Operationsverständnisses untersucht werden. Ähnliche Ergebnisse wie in früheren Studien, die daraufhin deuten, dass die Leistungen der Jungen gegenüber denjenigen der Mädchen besser sind, sollten zumindest nicht überraschen.

Die bisherigen Erklärungsversuche solcher Differenzen verwenden das tiefere Selbstkonzept der Mädchen sowie die Einstellung bzw. die Vorurteile der Lehrpersonen als erklärende Variablen (Moser Opitz, 2013, S. 180).

Das Geschlecht aller Kinder wurde in der Studie erfasst und konnte für die Auswertungen verwendet werden.

Ein Blick auf Studien aus anderen Kulturkreisen

Ein Blick in Untersuchungen aus anderen Kulturkreisen (auch ausserhalb des deutschen und englischen Sprachraums) zeigt auf, dass die Vergleichbarkeit der Forschungsergebnisse nicht nur durch den Wechsel des Sprachraums erschwert wird. Es können sich weitere Komplikationen ergeben, wenn nebst der Sprache auch die Lehrmittel, das Schulsystem oder z.B. auch das verwendete Zahlensystem anders ausgestaltet sind. Eine Studie, an der beispielhaft sichtbar wird, wie sich ein andersartiges Zahlensystem auswirken kann, ist Zhang et al. (2019).

Aufgrund solcher Schwierigkeiten sind die Untersuchungen aus dieser Arbeit grundsätzlich auf den deutschsprachigen Raum zu beziehen.

In diesem ersten theoretischen Teil wird vorwiegend mit Erkenntnissen aus dem deutsch- und zum Teil auch aus dem englischsprachigen Raum gearbeitet, in welchen zwar Unterschiede bestehen, aber wo zumindest dasselbe Zahlensystem verwendet wird.

3.1.4 Bedeutung und Besonderheiten der Grundrechenart «Multiplikation»

Die Grundrechenart der Multiplikation bzw. das Verständnis der Kinder für diese spezifische Operation bildet das Hauptaugenmerk der vorliegenden Untersuchung. Die Auswahl wurde wie in Kapitel 2.3.1 dargelegt aufgrund des Schwierigkeitsgrads und aufgrund des Behandlungszeitpunktes im Unterricht so vorgenommen.

Wegen der zentralen Bedeutung dieser Operation für diese Arbeit wird die Multiplikation in diesem Kapitel gesondert und im Hinblick auch auf ihre Besonderheiten vertieft betrachtet und diskutiert.

Multiplikation als wiederholte Addition

Eine wichtige Beobachtung besteht in der Tatsache, dass es bei der Multiplikation im Gegensatz zu den Grundoperationen erster Stufe nicht um das Zusammenfügen oder Auseinandernehmen von Mengen geht, sondern um eine Art von «Replikation» (Barmby et al., 2009, S. 218). Dies macht die Multiplikation zu einer Grundoperation höherer Stufe. Die multiplikativen Denkstrukturen werden als kritischer Moment im Lernprozess der Kinder beschrieben und als wesentlich für mathematische Lernfortschritte in der zweiten Hälfte der Primarschulzeit angesehen. Betont wird dabei auch, dass die Kinder für weiteren Fortschritt in ihren mathematischen Leistungen die Multiplikation nicht nur als wiederholte Addition betrachten dürfen (Fischbein et al., 1985; Hurst, 2015; Venkat & Mathews, 2019, S. 97). Obwohl die wiederholte Addition für die Multiplikation als diejenige Rechenstrategie beschrieben wird, die am wenigsten fehleranfällig ist (Mulligan & Mitchelmore, 1997, S. 320), wird für die lernenden Kinder ein Ablöseprozess von dieser einen Strategie bzw. eine Repertoire-Erweiterung notwendig, um in verschiedenen Situationen erfolgreich rechnerisch tätig zu bleiben.

Die Multiplikation bleibt aber ein «complex set of concepts» mit vielen Verbindungen und die Autoren, die sich damit beschäftigen, weisen auch darauf hin, dass ihre Arbeiten nur einzelne Aspekte beleuchten können (Hurst, 2015; Hurst & Hurrel, 2016). Dieselben Autoren stellen nach eingehender Auseinandersetzung fest, dass die Komplexität des mathematischen Gehalts beim Unterrichten der Multiplikation überraschend hoch sei: «the greatest learning was in terms of realising the complexity of the mathematics involved in teaching about multiplicative thinking» (Hurst & Hurrel, 2017).

Auch in weiterführender und nicht primär auf die Grundschule bezogener mathematikdidaktischer Literatur wird betont, dass die Multiplikation sich nicht restlos als wiederholte Addition beschreiben lasse. Confrey & Smith (1995) belegen, dass die wiederholte Addition viele multiplikative Situationen weder angemessen noch ausreichend beschreibe. Die Autoren entwickeln in ihrer Arbeit mit «splitting» ein alternatives gedankliches Modell.

Vergnaud (1983) hält fest, dass multiplikative Strukturen nur teilweise auf additiven Strukturen beruhen. Multiplikative Strukturen haben auch eine «eigene, intrinsische Organisation», die nicht auf additive Aspekte reduziert werden könne. Der Autor arbeitet bei seiner Analyse

mit dem Konzept der «conceptual fields» zur Beschreibung multiplikativer Strukturen (Vergnaud, 1983, 1996).

In Fischbein et al. (1985) wird ebenfalls auf die Schwierigkeit eingegangen, die sich ergibt, wenn sich die wiederholte Addition als Modellvorstellung für die Multiplikation bei den Lernenden verfestigt. Anhand einer Studie mit 623 Kindern in italienischen Schulen wurde nachgewiesen, dass die Kinder in höheren Schulklassen in Schwierigkeiten geraten können, falls die Operatoren nicht mehr ganzzahlig sind. Eine einfache Modellvorstellung wie die wiederholte Addition kann die Aufgabenbearbeitung der Kinder in einzelnen Situationen vereinfachen und beschleunigen, sie kann diese aber auch verlangsamen, erschweren oder sogar blockieren. Das usprünglich im Unterricht erarbeitete Modell kann sich bei den Kindern zum Teil als äusserst zäh und dauerhaft unvollständig erweisen. Die Kinder benötigen nebst den ursprünglichen Modellen auch effiziente Denkstrategien, die es ihnen erlauben, mit den erworbenen Modellen flexibel umzugehen (Fischbein et al., 1985, S. 14 ff.).

Bei der Untersuchung der Nützlichkeit der «wiederholten Addition» als Basis für den Aufbau des Multiplikationsverständnisses bestehen Unterschiede zwischen der deutsch- und der englischsprachigen Grundschuldidaktik (Wittmann, 2015). Unterschiedliche Haltungen gibt es auch bei den Betrachtungen zum kombinatorischen Aspekt der Multiplikation.

Kommutativität

Eine weitere zentrale Eigenschaft der Multiplikation ist die Kommutativität (für die Kinder meist «Tauschaufgabe» genannt). Die Einsicht, dass und weshalb die Multiplikation kommutativ ist, bildet einen Bestandteil eines tragfähigen Operationsverständnisses und ist hilfreich auch bei der späteren Arbeit der Kinder am Einüben des Einmaleins (Barmby et al., 2009). Für die Einführung des Grundverständnisses für diese Rechenart wird aber vorerst eine klare Unterscheidung von Multiplikator und Multiplikand empfohlen (Gaidoschik, 2015a). Die Grundoperation sollte unter Ausklammerung des Ergebnisses eingeführt werden und somit wird die Gleichheit des Ergebnisses der Tauschaufgaben auch erst später zum Thema gemacht. Auch bei der Verwendung der wiederholten Addition als erster Modellvorstellung zur Erarbeitung der Multiplikation ist es für die Kinder nicht auf Anhieb ersichtlich, dass sich die Multiplikation später als kommutativ herausstellen wird (Fischbein et al., 1985, S. 6).

Arbeitsstrategien der Kinder

Anghileri (1989) untersuchte das Verhalten von Kindern und ihre Strategien beim Lösen von Multiplikationsaufgaben. Die Autorin beobachtete, dass für die Multiplikation immer wieder Zählstrategien in irgendeiner Form zum Einsatz kamen. Der Abruf von Ergebnissen («use of multiplication facts») zum Lösen von Malaufgaben ist zwar eine effiziente Arbeitsstrategie, braucht aber bei vielen Kindern Zeit. Sie wurde vorwiegend von den Kindern verwendet, deren Leistungen in Mathematik über dem Durchschnitt lagen. Die meisten Kinder wendeten verschiedene Strategien an, 78% arbeiteten mit mindestens drei verschiedenen Strategien.

Spätere Untersuchungen an einer anderen Grundoperation belegten dann, dass unterrichtsbedingte Faktoren einen Einfluss auf die Strategien der Kinder haben können (Anghileri, 1989; Anghileri et al., 2002).

In Mulligan & Mitchelmore (1997) wurden bei einer Untersuchung in Australien ebenfalls Arbeitsstrategien der Kinder erfasst. Beschrieben werden sowohl für die Multiplikation als auch für die Division die «intuitiven Modelle», welche die Kinder entwickeln und auf die verschiedenen Aufgaben anwenden. Die Autoren beschreiben 12 verschiedene Rechenstrategien und drei zentrale intuitive Modelle, welche von den Kindern auf die Multiplikation angewendet werden. Die drei Modelle sind Zählstrategien, wiederholte Addition und multiplikative Operation. Die Entwicklung und Anwendung dieser Strategien durch die Kinder scheint komplex zu sein und zu belegen, dass die Lernprozesse nicht linear verlaufen. Die Kinder entwickeln während der Lernphase ein sich mit der Zeit ausdehnendes Repertoire an unterschiedlichen Modellen. Die Anwendung der Lösungsstrategien hängt dabei u.a. auch davon ab, auf welche Art eine Aufgabe semantisch präsentiert wird. Die Darstellung beeinflusst die Bearbeitung, denn die Bearbeitung bei (mathematisch) äquivalenten Aufgaben war unterschiedlich, wenn diese anders präsentiert wurden. Die Wahl der Arbeitsstrategie durch die Kinder erfolgt also nicht in konsistenter Art und Weise, sie ist aber einer Entwicklung unterworfen und macht Fortschritte. Die Autoren beobachteten eine Entwicklung und Ausdehnung des Strategierepertoires über die Zeit bei beiden Operationen zweiter Stufe. Auch in dieser Studie wurde festgestellt, dass Abrufstrategien bei leistungsstarken Kindern verbreiteter sind als bei leistungsschwächeren.

Sherin & Fuson (2005) haben Rechenstrategien der Kinder für einstellige Multiplikationsaufgaben eingeteilt und zueinander in Verbindung gebracht. Der Studie lagen 230 Interviews mit Kindern in der dritten Schulklasse zugrunde. Die Autoren halten mit Verweis auf verschiedene Vorarbeiten fest, dass sich beim Bemühen um solche Einteilungen eine grosse Unterschiedlichkeit ergeben hatte. Auch unterlagen die Rechenstrategien der Kinder verschiedenen Variabilitäten. Die Rechenstrategien der Kinder waren zum Beispiel abhängig vom Kontext, in dem Aufgaben gestellt wurden. Sie waren aber auch abhängig vom Unterricht und veränderten sich somit über die Zeit. Auch wurden sie von der Wahl der Zahlenwerte beeinflusst, die für die jeweiligen Aufgabenstellungen verwendet wurden. Die Autoren kommen zum Schluss, dass «zahlenspezifische» Rechenressourcen («number-specific computational resources») für die Veränderung der Rechenstrategien ausschlaggebend sind. Als Fazit ihrer Einteilung und Analyse schlagen sie vor zu prüfen, ob die Multiplikation und die Division nicht auch gleichzeitig im schulischen Unterricht behandelt werden könnten.

Didaktischer Nutzen der Multiplikation für andere Themenkreise

Die Multiplikation bildet natürlicherweise die Grundlage für das Verstehen der Division, welche die schwierigste der vier Grundoperationen darstellt. Im Lehrplan geht daher die Einführung und Behandlung der Multiplikation der Division voraus.

Darüber hinaus wird der Multiplikation aber auch Bedeutung zugeschrieben bei anderen Verstehens-Prozessen, so z.B. beim Verstehen der Proportionalität oder auch beim algebraischen Denken (Hurst & Hurrel, 2016).

Das multiplikative Denken wird in der Literatur beschrieben als kritischer Moment im Lernprozess der Kinder bei ihrer Entwicklung von Verständnis im Fach Mathematik (Hurst, 2015). Der Autor zählt das multiplikative Denken zu den «big ideas» der Mathematik.

In Hino & Kato (2019) wird mit Untersuchungen in Japan nachgewiesen, dass das Multiplikationsverständnis bzw. das Verstehen der Multiplikation («children's understanding of mathematics») einher geht mit besserem proportionalen Argumentieren («proportional reasoning»). Die Autoren beschreiben Unterrichtsstrategien, die beim Aufbau des Multiplikationsverständnisses hilfreich sind. Strukturelle Veränderungen von Malaufgaben sowie Reflexion über die getätigten Veränderungen spielen dabei eine wesentliche Rolle. Die Nützlichkeit und die Vorteile dieses vertieften Verständnisses bei den Kindern ziehen sich dann durch verschiedene Themen der späteren Schulmathematik.

3.2 Befunde zum «Operationsverständnis»

In diesem Kapitel werden die Ergebnisse bisheriger Untersuchungen wiedergegeben. Es zeigt sich insbesondere auch, dass die bisherigen Ergebnisse mehrheitlich aus Forschungsarbeit mit Hilfe von Interviews stammen. Quantitative Untersuchungen sind bisher verhältnismässig selten durchgeführt worden.

3.2.1 Untersuchungen unter Verwendung von Interviews

Wie bereits dargelegt wird in der Literatur häufig die Fähigkeit der Kinder zum Transfer zwischen verschiedenen Darstellungsarten für die Einschätzung bzw. als Indiz für die Ausprägung ihres Operationsverständnisses herangezogen. Diese Fähigkeit erlaubt eine empirische Erfassung des Konstrukts Operationsverständnis und eine für Forschungszwecke notwendige Operationalisierung.

In den meisten Fällen erfolgten entsprechende Untersuchungen bisher unter Anwendung qualitativer Methoden. Typischerweise werden mit Hilfe von ausführlich dokumentierten Interviews mit Kindern deren Leistungen, Aussagen und Gedanken beim Transfer zwischen Darstellungen beschrieben (Bönig, 1995; Schäfer, 2005; Kuhnke, 2013; Burtscher, 2019). Solche Interviews zeigen auf, wie der Darstellungswechsel mit der individuellen Gedankenwelt des jeweiligen Kindes eng zusammenhängt und von dieser mitgeprägt wird.

In den vielfältigen und zahlreich vorliegenden Interviews mit den Kindern kommen auch die zum Teil grossen und erstaunlichen Schwierigkeiten der Kinder beim Transfer zwischen Darstellungen ans Licht. Obwohl gemäss schweizerischen und deutschen Lehr- und Bildungsplänen Kinder zum Ende des zweiten Schuljahrs über ein gesichertes Operationsverständnis verfügen sollten, zeigen empirische Studien, dass viele Kinder Rechenterme nicht mit Handlungen, Bildern oder Sachsituationen verknüpfen können (Bönig, 1995; Schäfer, 2005; Burkhardt, 2008; Royar, 2013). So konnte gezeigt werden, dass es im dritten Schuljahr ungefähr einem Drittel der Kinder Schwierigkeiten bereitet, wenn sie die Bedeutung einer Grundrechenart anhand einer Handlung bzw. mit Hilfe von Wendeplättchen erklären sollen (Royar, 2013). Dieses Resultat stimmt mit den Feststellungen in Burkhardt (2008) überein. Es wurde auch berichtet, dass rechenschwache Kinder Zahlenoperationen in Ziffernschreibweise nicht mit sachgerechten Vorstellungen von Mengen assoziieren können und umgekehrt nicht in der Lage sind, einfache Mengenoperationen durch Ziffern und Rechenzeichen auszudrücken (Focke, 2004).

Aufgrund der Erkenntnisse aus solchen Interviews und Untersuchungen des Lernverhaltens von Kindern wird auch vor einer vorschnellen Einschätzung bzw. Beurteilung des Operationsverständnisses als «richtig» oder «falsch» gewarnt. Auch eine Einteilung des Vorgehens der Kinder in «gelungenen» und «nicht gelungenen» Darstellungswechsel ist problematisch (Kuhnke, 2013; Freesemann, 2014). Es werden vorsichtigere Formulierungen vorgeschlagen oder eine abgestufte Beurteilung, z.B. als «ausreichendes», «unsicheres» bzw. «problematisches» Operationsverständnis, wie in Schäfer (2005, S. 205) verwendet, oder aber das Vorge-

hen der Kinder beim Darstellungswechsel eingehend untersucht und beschrieben (Kuhnke, 2013).

In Kuhnke (2013) wird für die Multiplikation anhand einer Analyse von 15 Einzelinterviews der Prozess des Darstellungswechsels analysiert. In der Arbeit wird darauf hingewiesen, dass diese prozessorientierte Perspektive auf die Vorgehensweise der Kinder bisher selten eingenommen wurde. Der Darstellungswechsel erwies sich als komplexes Gefüge und war geprägt von einem stetigen Wechselspiel zwischen Deutungen und Verbindungsherstellungen zwischen verschiedenen Deutungen. Bei dieser Untersuchung zeigte sich, dass der Fokus der Kinder auf den einzelnen Faktoren (den Einzelelementen), dem Ergebnis oder auf der Relation der Faktoren liegen kann. Bei einem Fokus auf die einzelnen Faktoren nennen die Kinder Darstellungen als zueinander passend, wenn sie die beiden Zahlenwerte der Faktoren (z.B. in einem Bild) wiedererkennen. Liegt der Fokus auf dem Ergebnis, so nennen die Kinder ein Bild als zum Term passend, wenn das Ergebnis der Multiplikation im Bild als Zahlenwert ersichtlich ist. Bei einem Fokus auf die Relation wird die «relationale Information abgeglichen», d.h. die Kinder achten auch auf die Relation, in der die einzelnen Elemente zueinanderstehen. Dieser Fokus der Kinder kann sich entwickeln und durch Beschulung bzw. durch Unterrichtsinterventionen verändern. Auch in dieser Arbeit wird angesichts der Beobachtungen davor gewarnt, einen Darstellungswechsel vorschnell als «gelungen» oder «nicht gelungen» einzuschätzen. Für die Einschätzung von Operationsverständnis ist es nicht ausreichend, einzelne Darstellungswechsel zu beobachten. Notwendig ist es, verschiedene Darstellungswechsel zu analysieren und bei dieser Analyse auch auf den Prozess zu achten (Kuhnke, 2013, S. 263 ff.) .

Die Beobachtungen und die Erkenntnisse aus solchen Interviews erklären auch, weshalb in diesem Themengebiet hauptsächlich mit qualitativen Studien gearbeitet wird und eher wenige quantitative Forschungsmethoden zur Anwendung kommen. Viele Beobachtungen und Denkvorgänge der Kinder sind komplex und schwer vorherzusehen. Die Denkvorgänge der Kinder mögen unerwartet, aber durchaus korrekt sein. Diese Komplexität der Untersuchungen erschwert eine Einteilung der Kinderleistungen in «richtig» und «falsch» und dementsprechend auch eine für quantitative Zwecke wünschenswerte Codierung in 1 und 0.

Im Rahmen der vorliegenden Arbeit kann solchen Forderungen nach eingehender Analyse der Gedankengänge des einzelnen Kindes aufgrund der Anzahl Kinder und der damit verbundenen Testökonomie nicht immer nachgekommen werden. Trotzdem konnten mit dieser Arbeit interessante Informationen gewonnen werden, welche die mehrheitlich qualitativen Arbeiten im Bereich des Operationsverständnisses auf sinnvolle Art und Weise ergänzen.

3.2.2 Interventionsstudien und quantitative Untersuchungen

Einige Interventionsstudien behandelten in den vergangenen Jahren das Operationsverständnis, allerdings sind derartige Studien nur in geringer Anzahl vorhanden (Freesemann, 2014; J. Walter et al., 2001).

Unter den vorliegenden Interventionsstudien gibt es solche mit einer geringen Schülerzahl (J. Walter et al., 2001) und auch einige, in denen das Operationsverständnis nicht die zentrale Rolle einnahm, sondern neben verschiedenen andern Themen untersucht wurde (Freesemann, 2014; Pfister, 2016, S. 96 f.; Pfister et al., 2015). In Freesemann (2014) zeigten sich nur geringe Effekte der Interventionen bzw. nur in den ausdrücklich geförderten Themengebieten, nicht aber bei der Rechenleistung der Kinder. Die Ergebnisse sind also bisher uneinheitlich und ergänzungsbedürftig. Ein wichtiger Aspekt für die Wirksamkeit allfälliger Interventionen scheint gemäss heutigem Forschungsstand die Zeitdauer des Einsatzes der jeweiligen Interventionen zu sein (Focke, 2004, S. 240; Slavin, 2008; Freesemann, 2014; Wittich, 2017). Dies ist ein Aspekt, der in der geplanten Arbeit Beachtung finden wird.

Im Übrigen zeigen Hasemann & Stern (2002) mit einer Interventionsstudie in der 2. Klasse, dass abstrakte Trainingsprogramme gerade für schwache Kinder auf dieser Stufe vorteilhaft sein können.

In Lamprecht (2017) wurde eine Interventionsstudie zum Multiplikationsverständnis mit 18 Klassen und ca. 340 Kindern angekündigt. Von den Ergebnissen dieser Studie wurde in Lamprecht (2020) berichtet. In der Arbeit wurden die Wirkungen eines Unterrichts-Settings mit Einzelförderung (Setting A; n=8) für einzelne Kinder mit Förderbedarf verglichen mit einem Unterrichts-Setting im Klassenverband (Setting B; n=124) und mit einer Kontrollgruppe (Setting C; n=108). Es wurden sowohl quantitative als auch qualitative Untersuchungen vorgenommen. In der quantitativen Analyse wurden die Lernentwicklungen der Kinder in den drei Bereichen «Angemessenheit», «Grundvorstellungen» und «Ergebnisrichtigkeit» untersucht. Die Messungen erfolgten zu drei Messzeitpunkten um auch Langzeiteffekte zu erfassen. Es zeigten sich sowohl bei der «Angemessenheit» als auch bei den «Grundvorstellungen» und der «Ergebnisrichtigkeit» signifikante positive Effekte der Förderung im Setting A. Diese Effekte zeigten sich sowohl nach der Intervention als auch als Langzeit-Effekte beim dritten Messzeitpunkt. Bei den Kindern mit Förderbedarf im Setting B zeigte sich ein zusätzlicher signifikanter positiver Effekt bei der «Angemessenheit» gleich nach der Intervention.

In Schulz et al. (2019) wird ein Kompetenzmodell mit 4 Stufen («levels») für die Fähigkeiten der Kinder beschrieben und verwendet. Die verschiedenen Stufen unterscheiden sich in den Grundrechenarten, die von den Kindern beherrscht werden müssen, sowie in der Vielfalt der Grundvorstellungen, welche die Kinder zur Verfügung haben müssen. Verschiedene Items und ihr Schwierigkeitsgrad, der mit Logit-Werten im Rasch-Modell erfasst wird, belegen die Validität des vierstufigen Kompetenzmodells.

Als adäquate Lernaktivität wird wie in der vorliegenden Arbeit auch das Einüben des Darstellungswechsels dargelegt (Schulz et al., 2019, S. 9). Des Weiteren wird darauf hingewiesen, dass für Kinder mit einem tiefen Kompetenzlevel Signalwörter entscheidend sein können für ihre Fähigkeit, die entsprechenden Items korrekt zu lösen, was natürlich für ein tieferes Ver-

ständnis der Inhalte nicht der erwünschte Zustand des Lernenden darstellt. In der Arbeit wird auf quantitative Art und Weise die Erkenntnis belegt, dass die Division die schwierigste der vier Grundrechenarten darstellt.

In Venkat & Mathews (2019) wird eine Interventionsstudie beschrieben, die im Jahr 2017 in Südafrika in Schulklassen des 7. Jahrgangs durchgeführt wurde, und die zum Ziel hatte, den Lernfortschritt beim «Multiplicative Reasoning (MR)» auszuwerten. Die Studie umfasst mehrheitlich Kinder mit schwächeren Leistungen in Mathematik, was durch die Auswahl der Schulen und Klassen bewusst so gewählt war. Nebst multiplikativen Aufgaben waren auch Aufgaben zur Division Teil des Tests und der Förderung in den Interventionsklassen. Die Interventionen waren von verhältnismässig kurzer Dauer und umfassten 4 Lektionen; der zeitliche Abstand zwischen pre- und post-Test betrug 6 Wochen. Die Interventionen fanden in einem Setting mit Ähnlichkeiten zum kooperativen Setting statt, insbesondere wurden Einzel- und Paararbeitsphasen am Interventionsmaterial verwendet. Es wurden Fortschritte bei den Leistungen der Kinder im Bereich des multiplikativen «Reasonings» (Begründens) festgestellt. Die Interventionen führten dazu, dass die Kinder besser Modelle entwerfen und diese sinnvoller beschreiben und verwenden konnten. Die Autoren kommentieren entsprechend kritisch die Rahmenbedingungen der Lehrpläne in Südafrika, welche hohen Wert auf Rechenleistungen («fluent calculation») legen, z.B. bei der Division vierstelliger Dividenden durch dreistellige Divisoren. Diese Art von Lehrplänen trägt die Gefahr in sich, dass das Argumentieren, Modellieren und eigentliche Verstehen der Kinder zu Gunsten von Rechenkünsten vernachlässigt wird. Die Absicht der Autoren ist es nun, zu prüfen, ob sich mit noch kürzeren Interventionen ebenfalls Erfolge erzielen lassen.

In Stöckli (2019, S. 51 ff.) wird von verschiedenen Interventionsstudien zur Förderung der arithmetischen Kompetenzen sowohl in der Unter- als auch in der Mittelstufe berichtet.

Zum Teil wurde im Rahmen von Interventionsstudien auch beobachtet, dass der Effekt einer Förderung sich erst langfristig zeigt. In Pfister et al. (Pfister et al., 2015) konnte gleich nach den Interventionen noch keine Wirkung nachgewiesen werden kann, mit zeitlicher Verzögerung zeigte sich dann aber doch eine solche. Auch wurde darauf hingewiesen, dass der Unterricht in inklusiven Klassen komplex sei und daher noch viel Entwicklungs- und Forschungsarbeit benötigt werde (Pfister et al., 2015, S. 65).

3.3 Förderung von Operationsverständnis in unterschiedlichen Settings

In diesem Kapitel werden zuerst unterschiedliche Formen der Förderung von Operationsverständnis behandelt. Im Anschluss daran werden spezifische Aspekte wie z.B. die Sprachsensibilität (in Kapitel 3.3.3) oder auch der Einsatz von Material (in Kapitel 3.3.4) besprochen.

3.3.1 Behandlung im regulären Unterricht

Alle Kinder sollten im Laufe ihrer Schulzeit Verständnis für die vier Grundoperationen entwickeln. Der Umgang mit den Grundoperationen wird im Laufe ihres Lebens in vielerlei Zusammenhängen immer wieder von Bedeutung sein. Lücken im Verständnis der Grundoperationen können (wie in Kapitel 3.1 beschrieben) zu gravierenden Folgen in der späteren schulischen Laufbahn, aber auch im späteren Leben führen.

Aus diesem Grund werden die Grundoperationen auch im regulären Unterricht behandelt und sind in allen Schulbüchern der ersten drei Primarschuljahre prominent vertreten (Brandenberg et al., 2013, 2015; Buschmeier et al., 2017, S. 64 ff.; Hess, 2018; Wittmann, 2011; Wittmann & Müller, 2008). Am Operationsverständnis wird selbstverständlich ebenfalls im regulären Schulunterricht gearbeitet. Die Intensität der Behandlung dieses Themas kann aber unterschiedlich sein und auch das Bewusstsein der Lehrperson für dessen Wichtigkeit ist unterschiedlich stark ausgeprägt. Dies wurde beobachtet, obwohl die Thematik im Lehramtsstudium und in der Ausbildung zur Primarschullehrperson an den pädagogischen Hochschulen in der Schweiz und auch in Deutschland eingehend behandelt wird (Baireuther, 1999; Gerster, 1994).

Gemäss dem Vorschlag der jeweiligen Lehrmittel zur Planung des Unterrichtsjahrs kommt der Arbeit am Operationsverständnis (bzw. dem Aufbau von Grundvorstellungen) für die jeweilige Operation beachtliche Zeit zu (Brandenberg et al., 2013, 2015; Wittmann, 2011; Wittmann & Müller, 2008; Hess, 2018). Die Operationen werden einzeln eingeführt und der Darstellungswechsel wird anhand verschiedener Darstellungsarten geübt. Allerdings lässt sich anhand der Jahresplanung des Lehrmittels nicht exakt eruieren, wie viel Unterrichtszeit dann effektiv in diese Thematik investiert wird. Auch ist nicht ersichtlich, in welchem Ausmass der Unterricht erfolgreich war, d.h. wie viele der unterrichteten Kinder in dieser Zeitspanne ein ausreichendes Operationsverständnis (bzw. tragfähige Grundvorstellungen zu den jeweiligen Operationen) aufbauen konnten.

Eine Übersicht zum Einfluss unterrichtlicher Aspekte bei der Ausbildung bzw. Behebung mathematischer Lernstörungen ist u.a. in Moser Opitz (2013, S. 32 ff.) zu finden.

Verschiedene Studien weisen auch nach, dass Lücken im Verständnis für die Grundoperationen noch lange Jahre nach Ersteinführung im Unterricht auftreten können (Radatz, 1991; Jacobs & Petermann, 2003; Weinhold Zulauf et al., 2003; Moser Opitz, 2005; Schipper, 2005; Ise et al., 2012).

3.3.2 Spezifische Förderung des Operationsverständnisses

Es erfolgt in diesem Kapitel eine Unterscheidung zwischen der individuellen Förderung einzelner Kinder und der Förderung der Kinder im Klassenplenum. Zusätzlich vertieft wird das kooperative Setting als eine mögliche Art der Förderung im Klassenplenum.

Einzelförderung

Es besteht in der didaktischen Literatur eine starke Übereinstimmung in der These, dass rechenschwache Kinder möglichst frühzeitig zu fördern sind. Eine erfolgversprechende Förderung sollte nach erfolgter Diagnose stattfinden. Sie sollte mit dem betroffenen Kind einzeln durchgeführt werden, spezifisch auf das betroffene Kind zugeschnitten sein und bezogen auf die jeweilige Problematik zielgerichtet ablaufen (Gerster & Schultz, 2004; Kretschmann, 2006; Lorenz, 2008, S. 98 f.; Schäfer, 2005; Fritz et al., 2018; Gaidoschik, 2015a).

Eine solche Einzelförderung sollte möglichst bald erfolgen, wenn sich Lücken im Verständnis des Kindes zeigen, die eine solche Förderung im Einzelsetting nahelegen. Ein frühzeitiges Erkennen der Schwächen und der Verständnislücken eines Kindes dient dazu, angemessene Massnahmen frühzeitig einzuleiten und die Entstehung umfangreicherer Verständnislücken zu verhindern. Insbesondere empfiehlt sich eine frühzeitige Überprüfung des Zahlverständnisses bei den Kindern (Gaidoschik, 2015a). Bleiben wesentliche Verständnislücken unerkannt, so droht sich die Problematik im Laufe der Schulzeit zu verschärfen.

Die Diagnose der jeweiligen Schwierigkeiten des einzelnen Kindes ist deshalb wichtig, weil die Schwierigkeiten bei verschiedenen Kindern sehr unterschiedlich gelagert sein können. Die «Hürden, die Kinder beim Erlernen des Rechnens überwinden müssen», sind zahlreich (Gaidoschik, 2015b, S. 25). Als Folge dieser Situation werden für eine zielgerichtete Förderung sehr unterschiedliche Behandlungsansätze notwendig.

Die Einzelförderung wird deshalb empfohlen, weil nur in diesem Setting das Kind entsprechend begleitet und der Lernfortschritt durch die Betreuungsperson präzise mitverfolgt werden kann (Bräuning & Schülke, 2010). In der Literatur wird bezogen auf den Gesamtunterricht z.T. lediglich auf die Wichtigkeit einer «differenzierten Analyse des Lernstands» verwiesen (Schäfer, 2005, S. 264), die eigentliche Förderung sollte dann aber in der Einzelarbeit mit dem Kind erfolgen (Gaidoschik, 2015a). In einzelnen Arbeiten wird auch argumentiert, Förderung sei ohne die persönliche Interaktion zwischen Kind und Förderperson undenkbar (Schipper, 2002).

Materialien zur Einzelförderung enthalten aus obigen Gründen darum häufig Instrumente zur Diagnose. Diese Instrumente werden ergänzt durch Unterrichts- bzw. Fördermaterial zur anschliessenden zielgerichteten Arbeit mit dem Kind oder durch das Kind (Scherer, 2004; Lorenz, 2008; Lorenz & Radatz, 2008; Selter & Akinwunmi, 2014; Freesemann & Breucker, 2014; Scherer, 2016; Gaidoschik, 2016a). Auch Fördermaterial, welches dem Aufbau von Grundvorstellungen für die arithmetischen Operationen dient, ist diesem Material zuzurechnen. In einigen dieser Arbeiten wird die Bedeutung des «verdeckten» Einsatzes von Material

zum Aufbau von Vorstellungen durch die Kinder besonders betont (Wartha, 2010; Wartha & Schulz, 2013; Gaidoschik, 2016a; Rottmann, 2018).

Unterrichtsintegrierte Förderung

Die Praxis der Einzelförderung findet allgemeine Anerkennung, führt aber auch dazu, dass bis heute kaum Antworten auf die Frage vorliegen, welche Aspekte des Unterrichts im Klassenverband die Entwicklung des Operationsverständnisses begünstigen oder hemmen.

Für die Anreicherung des Unterrichts im Klassenverband stehen ebenfalls bereits Materialien zur Verfügung (Kuhnke, 2012). Solche Unterlagen können die jeweils verwendeten Lehrmittel ergänzen und auch das Bewusstsein der Lehrpersonen für die Thematik schärfen. Einige Schulbücher enthalten auch Material, welches zur Förderung des Operationsverständnisses genutzt werden kann. In der Unterrichtspraxis kann so – gegebenenfalls durch den Einbezug eines zweiten Lehrmittels − durch die Lehrperson ergänzendes Material zum Einsatz gebracht werden. In der Schweiz eignen sich dafür natürlich insbesondere die gängigen Lehrmittel (Wittmann & Müller, 2008; Brandenberg et al., 2013; Hess, 2018).

Zu beachten ist bei der unterrichtsintegrierten Behandlung des Operationsverständnisses, dass das grundlegende Verständnis, dessen Entwicklung einigen Kindern Mühe bereitet, für die Mehrheit der Kinder aber keine sehr hohe Hürde darstellt. Bei der eingehenden und über längere Zeit andauernden Behandlung der Thematik im Klassenverband, z.B. durch intensives Üben des Darstellungswechsels, besteht somit auch die Gefahr, dass stärkere Kinder nicht ausreichend gefordert werden. Um dies zu vermeiden, ist eine geschickte Planung der unterrichtsintegrierten Behandlung des Themas notwendig. Der Unterricht muss so gestaltet sein, dass auch stärkeren Kindern eine für sie interessante Tätigkeit zukommt und sich diese nicht langweilen. Im Folgenden wird mit dem kooperativen Setting (siehe auch Kapitel 4.3.2) ein solcher Ansatz zur Unterrichtsgestaltung besprochen.

Die unterrichtsintegrierte Förderung unterscheidet sich also deutlich von der mehrheitlich empfohlenen Einzelförderung, auch wenn sie ähnliche Ziele verfolgt. Als einer der Vorteile der unterrichtsintegrierten Förderung kann angeführt werden, dass sie innerhalb des staatlich organisierten Unterrichts stattfindet. Dies wird von Autoren begrüsst, welche die Auslagerung der Förderarbeit an ausserschulische, auf kommerzieller Basis agierende «lerntherapeutische Zentren» kritisieren (Schipper, 2002). Gelingt eine solche unterrichtsintegrierte Behandlung einer kritischen Thematik, so tut sie dies auch, ohne dass rechenschwache Kinder identifiziert werden müssen, was zudem dem Selbstvertrauen der betroffenen Kinder dienen könnte. Dies erklärt die mancherorts formulierte Forderung, integrative Unterrichtsformen müssten weiterentwickelt werden und gezielte Förderung habe grundsätzlich im Rahmen des regulären Unterrichts stattzufinden (Pfister, 2016).

Für eine solche Förderung werden u.a. auch passende Lernumgebungen empfohlen. Häsel-Weide (2016) legt dar, dass Kinder für den Aufbau des Multiplikationsverständnisses Beziehungen zwischen Zahlen verstehen müssen, zwischen unterschiedlichen Darstellungen übersetzen können sowie Vorstellungen darüber aufbauen sollten, wie mit zusammengesetzten

Einheiten umzugehen sei. Die Autorin präsentiert in dieser Arbeit auch Unterrichtsvorschläge und illustriert solche am Beispiel «Punktefelder deuten». In diesen Unterrichtsvorschlägen arbeiten die Kinder kooperativ und mit einem hohen Anteil an Kommunikation untereinander. Die Bedeutung der Kommunikation für die Lernprozesse im Fach Mathematik wird in der Literatur vielfach genannt. Brandt (2019, S. 9 ff.) bietet eine Übersicht über die Literatur zu diesem Aspekt des Lernens.

3.3.3 Sprachsensibilität der Förderung und Bedeutung der Sprache in der Mathematik

Die Verwendung von Sprache und die Bedeutung der Sprache bei Lernprozessen in der Mathematik wurden bereits in vielen Studien untersucht. Dabei wurde mehrfach belegt, dass die Bedeutung der Sprache für den Mathematikunterricht weit über die Fähigkeit der Lesekompetenz hinaus geht (Dörfler, 1988; Schipper, 2002; Götze, 2007; Nührenbörger, 2009; Gaidoschik, 2010, S. 75; Morgan et al., 2014; Prediger et al., 2015; Götze, 2015; Prediger & Götze, 2017; Verboom, 2017; Götze, 2018; Tiedemann, 2018; Langer-Osuna, 2018; Ufer et al., 2020; Ufer & Bochnik, 2020; Leiss & Plath, 2020).

In diesem Kapitel werden nun einige Studienergebnisse beleuchtet, die für die Ausgestaltung dieser Arbeit von besonderer Bedeutung sein werden. Den sprachlichen Aspekten der Förderung wurde in den Fördereinheiten, die für diese Arbeit entwickelt wurden, hohe Bedeutung beigemessen. Immer wieder wurde in den Einstiegs- und Reflexionsphasen des Unterrichts in den Interventionsklassen das gemeinsame Vorgehen mit den Kindern besprochen und entweder durch die Lehrperson oder aber durch die Kinder selber verbalisiert.

Die Wichtigkeit einer solchen sprachlichen Begleitung wird in der Literatur verschiedentlich betont. Tiedemann (2020) untersuchte z.B. die Verwendung von Sprache durch Kinder auf der Grundschulstufe zum Beschreiben von Handlungen bzw. Materialhandlungen. In der Studie wird nachgewiesen, dass es nicht *das eine* Beschreiben gibt, sondern verschiedene Praktiken des Beschreibens vorkommen. Schwarzkopf (2019) legt dar, dass kollektives Argumentieren notwendig sei, um substantielle mathematische Lernprozesse gelingen zu lassen. Unter kollektivem Argumentieren wird dabei ein Diskurs verstanden, der von Begründen und Verstehen, aber auch von strittiger Auseinandersetzung geprägt ist. Die durch die Lehrperson angeregte Diskussion bzw. Auseinandersetzung soll dabei nicht zur Schwäche, sondern zur Stärke des Mathematikunterrichts werden. Dies insbesondere dann, wenn sich die Klasse argumentativ mit dem Unterrichtsstoff auseinandersetzt. Weitere Autoren verweisen auf die Wichtigkeit des «Kommunizierens» über die wahrgenommenen Strukturen oder über das verwendete Material bzw. über den Vorgang des «Argumentierens» durch die Kinder (Fetzer, 2015; Tiedemann, 2017; Benz, 2018). Auch das «Begründen» wird in der Literatur als ein Vorgang von grundlegender Bedeutung für jegliches mathematische Tun angesehen (Käpnick, 2014).

Manche Autoren betonen auch die Wichtigkeit der Sprache gerade bei der Arbeit am Operationsverständnis bzw. bei der Arbeit mit Darstellungswechseln. Duval (2006) betont, dass Unterricht in Mathematik durch den Einsatz von Sprache immer einem Arbeiten in mehreren

Darstellungsarten gleichkommt. In der Mathematik wird meist mit der mathematischen Symbolik und zugleich auch mit sprachlicher Begleitung gearbeitet. Durch sie kommt immer eine weitere Darstellungsart ins Spiel und der Unterricht wird durch den Einsatz der Sprache in natürlicher Weise zum Wechselspiel zwischen verschiedenen Darstellungen (Duval, 2006, S. 125).

Kuhnke (2013) kam bei einer Interviewstudie am Darstellungswechsel bei der Multiplikation zum Schluss, die Thematik eigne sich besonders für Gespräche mit, aber auch zwischen Kindern. In Royar (2013, S. 120 f.) wird ein angemessener und verständlicher Gebrauch von Sprache als Indiz für Operationsverständnis vorgeschlagen.

In der hier vorliegenden Arbeit kommt im Hinblick auf die Sprachlichkeit in den Interventionen (Fördereinheiten) besonders der behutsamen Einführung und Behandlung der sprachlichen Kurzform «mal» hohe Bedeutung zu (Gaidoschik, 2015a, S. 106 ff.). Darüber hinaus werden die Kinder im Rahmen der Fördereinheiten immer wieder eingeladen, ihre Arbeiten zu besprechen und ihr jeweiliges Vorgehen zu begründen. Das für diese Studie gewählte und in den Kapiteln 3.3.2 und 4.3.2 beschriebene «kooperative und kommunikative» Unterrichtssetting legt hohen Wert auf diesen Aspekt der Förderarbeit.

3.3.4 Einsatz von Material und Anschauungsmitteln für Förderzwecke

Für die Förderung des Operationsverständnisses zentral sind Illustrationen und der sinnvolle Einsatz von Material, z.B. in Form von Bildern und Rechengeschichten (Freesemann, 2014, S. 96 ff.). Mit solchem Material wird der Darstellungswechsel mit den Kindern geübt. Dabei sind auch verschiedene Modellvorstellungen und Grundvorstellungen zu berücksichtigen (Stöckli, 2019, S. 90).

Trotz dessen Wichtigkeit ist Material aber auch sparsam und gezielt einzusetzen, denn Anschauungsmaterial stellt insbesondere die rechenschwachen Kinder jeweils auch vor Schwierigkeiten (Lorenz, 2013; Rottmann, 2018, S. 77 ff.). Der Transfer zwischen verschiedenen Anschauungsmaterialien stellt für viele Kinder eine Hürde dar. Eine zu grosse Vielfalt an Modellen kann die Kinder auch zu zusätzlichen Verständnisschwierigkeiten führen, denn Darstellungen und Modelle stellen nicht nur Lernhilfen, sondern auch Lernstoff dar (Gerlach, 2007; Scherer & Moser Opitz, 2010; Kuhnke, 2013, S. 9 ff.; Dreher, 2013; Wagner & Wörn, 2013). Auch in Hiebert & Carpenter (1992, S. 70) wurde bereits darauf hingewiesen, dass der Einsatz von Materialien im Unterricht zweischneidig sein kann und bereits zu unterschiedlichen Ergebnissen geführt hat. Gaidoschik (2015b, S. 34) umschreibt die Schwierigkeit im Grundschulunterricht mit «Desorientierung durch Orientierung an zu vielen Darstellungen in zu kurzer Zeit». Söbbeke (2005, S. 65) umschreibt dies so, dass Diagramme und Anschauungsmittel nicht selbsterklärend seien, sondern aktiv gedeutet werden müssen. Die Übersetzungsprozesse zwischen verschiedenen Anschauungsmitteln stellen für die Kinder immer auch Anforderungen dar. Diesen Standpunkt nimmt auch Steenpass (2014, S. 12 ff.) ein. Erforderlich seien aktive Interpretationsleistungen der Kinder, damit Anschauungsmaterial zur Entwicklung von mathematischem Wissen führt.

In Acevedo Nistal et al. (2009, S. 628) wird der Darstellungswechsel auch kritisch beleuchtet. Darstellungen sollten von den Lernenden verbunden und zueinander in Verbindung gebracht werden können. Sofern dies aber nicht gelingt und die einzelnen Darstellungen ohne Zusammenhang «isoliert nebeneinander» stehen, erfordert jede weitere Darstellung eine zusätzliche Denk- und Gedächtnisleistung. Die Autoren stellen fest, dass der Einsatz multipler Darstellungen auch einen negativen Effekt auf die Mathematikleistung haben kann.

Mehrdeutigkeiten sind produktiv zu nutzen und die verschiedenen Darstellungsarten im Unterricht eingehend zu diskutieren und zu reflektieren. Sonst besteht die Gefahr, dass verschiedene Darstellungsarten im Unterricht unverbunden nebeneinander stehen und die Lernenden den Darstellungswechseln durch die Lehrperson nicht in ausreichendem Masse folgen können (Burkhardt, 2008, S. 31; Dreher, 2013, S. 218 f.).

Die Fähigkeit, Strukturen in Anschauungsmittel hinein zu deuten, Beziehungen herstellen zu können und die Anschauungsmittel flexibel zu deuten, wird in der Literatur als fundamentale Voraussetzung für den Umgang mit solchen Materialien beschrieben (Söbbeke, 2005, S. 65).

Material und dessen Einsatz erfordert – wie bereits in Kapitel 3.3.3 ausgeführt − zusätzlich ein angemessenes Sprechen darüber. Dieses Sprechen selbst ist bereits ein Arbeiten in einer anderen Darstellungsart, nämlich der Sprache, allerdings ein begleitendes und üblicherweise unterstützendes. Die denkerische Last von zu umfangreichem Materialeinsatz sollte aber auch darum nicht unterschätzt werden, weil im Unterricht ja immer auch noch der sprachliche Aspekt dazu kommt.

Ein gelungener Materialeinsatz kann aber auch der Entlastung dienen, nämlich dann, wenn Sprache und Material «aufeinandertreffen» (Tiedemann, 2017, S. 45 ff.).

Erschwert wird die angemessene Materialwahl dadurch, dass die Lernenden unterschiedliche Voraussetzungen, unterschiedliches Vorwissen und unterschiedliche Vorlieben und denkerische Anknüpfungspunkte in den Unterricht mitbringen. Es wurde schon beobachtet, dass die Lernenden mit individuellen Vorlieben ausgestattet seien, die sich auch auf die für sie optimale Wahl von Mitteln zur Veranschaulichung auswirken (Acevedo Nistal et al., 2009, S. 632). Offensichtlich ist, dass diese Beobachtung Lehrpersonen gerade im unterrichtsintegrierten Setting vor grosse Herausforderungen stellt.

Selbstverständlich ist auch die Nützlichkeit des jeweiligen Materials für didaktische Zwecke unterschiedlich einzuschätzen. In der Literatur wird der Materialeinsatz abhängig vom jeweiligen Fach und jeweiligen Gebiet innerhalb dieses Fachs einzeln diskutiert. Die Nützlichkeit von Graphen zur Illustration von Sachverhalten kann z.B. sehr unterschiedlich sein (Acevedo Nistal et al., 2009, S. 630). Genauso kann der Nutzen von Material zur Veranschaulichung des Grundverständnisses für die Multiplikation verschieden eingeschätzt werden.

Alltagsbilder und didaktisches Material

Sowohl die Verwendung von Alltagsbildern als auch die Verwendung von didaktischem Material hat für die Förderung der Kinder in der Schule sowohl Vor- als auch Nachteile. Alltags-

bilder enthalten häufig zusätzliche Distraktoren, welche die Aufmerksamkeit der Kinder auf sich ziehen und den didaktischen Wert des Bilds dadurch mindern können. Auf der anderen Seite können die Kinder bei Alltagsbildern aber im Optimalfall «an Bekanntes anknüpfen». In der Literatur wird der Einsatz der beiden Bildarten unterschiedlich kommentiert. Zum Teil wird angenommen, Kinder hätten weniger Mühe mit Alltagsbildern, zum Teil wird dem didaktischen Material der Vorzug gegeben (Burkhardt, 2008, S. 32). Die Vorliebe für didaktisches Material wird so begründet, dass dieses entlastend wirke und für die Kinder leichter zu verarbeiten sei, weil «der lebenswirkliche bildliche Kontext das Konstruieren von Bedeutung schwieriger machen» kann (Kuhnke, 2013, S. 268).

Wie schwierig die richtige Materialwahl bzw. Auswahl von Bildern ist, zeigt auch die Tatsache auf, dass diese Wahl stark vom jeweiligen Themengebiet geprägt ist. Während in einigen Themengebieten Diagramme und schematische Bilder zu bevorzugen sind, sieht die bestmögliche Wahl in anderen Gebieten der Mathematik wieder anders aus (Thomas, 2008, S. 10). Und auch innerhalb eines einzelnen Gebiets sind sich die Autoren – wie oben dargelegt − nicht immer einig.

Zu beachten ist, dass gerade Alltagsbilder auch mit unterschiedlichen Distraktoren beladen sein können. Alltagsbilder knüpfen eher an Alltagserfahrungen der Kinder an; dieser Anknüpfungspunkt kann aber auch ein anderer sein als derjenige, den sich die Lehrperson für ihre didaktischen Bemühungen wünschen würde.

Die Arbeit mit dem Material sollte aber auf jeden Fall zielgerichtet und auf Strukturierungsaktivität der Kinder ausgerichtet sein (Söbbeke, 2005; Steenpass, 2014).

Auch wenn der Materialeinsatz für didaktische Zwecke rege diskutiert wird, so gibt es doch einen Konsens bei der Behandlung der Multiplikation im Unterricht. Die Nützlichkeit von Punktefeldern für didaktische Zwecke ist in der Mathematikdidaktik gerade für den Unterricht der Multiplikation unumstritten (Barmby et al., 2009; Harries & Barmby, 2007; Wittmann, 2015, S. 240 f.; Venkat & Mathews, 2019). Dies gilt, obschon die typischen Rechtecks-Darstellungen zum Teil auch kritisch diskutiert werden (D. Walter & Schulz, 2019, S. 50 f.). In Barmby et al. (2009) werden Punktefelder als Arbeitsmittel zwar grundsätzlich empfohlen, deren Vor- und Nachteile aber auch diskutiert.

In dieser Arbeit kommen beide Arten von Bildern und von Material zum Einsatz. Die Fördereinheiten enthalten hohe Mengen an Alltagsbildern, aber auch verschiedene Übungen mit didaktischem Material, so z.B. mit Punktefeldern, Punktebildern und Wendeplättchen. Besonders das Arbeitsmittel der Punktefelder, das vor allem auch für die Multiplikation sinnvoll eingesetzt werden kann, wird in den Fördermaterialien intensiv verwendet. Mit ihm lassen sich die für das Operationsverständnis wesentlichen Strukturierungen besonders vorteilhaft mit den Kindern üben.

4 Empirischer Teil 1: Forschungsprojekt und methodisches Vorgehen

An dieser Stelle wird in Kapitel 4.1 zuerst das gesamte Forschungsprojekt beschrieben, in welches die vorliegende Arbeit als Teiluntersuchung eingebettet war. Kapitel 4.2 widmet sich dann demjenigen Teil des Forschungsprojekts, welcher den Inhalt dieser Arbeit bildet.

Es folgt in den Kapiteln 4.3 und 4.4 eine eingehende Beschreibung des für diese Arbeit entwickelten Materials. In Kapitel 4.5 werden dann die Auswertungsmethoden besprochen, die im späteren Verlauf dieser Arbeit eingesetzt werden.

4.1 Einbettung in ein umfassendes Forschungsprojekt

Das Forschungsprojekt, in welches diese Studie eingebettet war, hat diese Arbeit in verschiedener Hinsicht stark geprägt. Aus diesem Grund erfolgt an dieser Stelle eine genaue Beschreibung des gesamten Forschungsprojekts.

4.1.1 Beschreibung des Forschungsprojekts

Das Forschungsprojekt, in dessen Rahmen diese Interventionsstudie durchgeführt und ausgewertet wurde, trägt den Namen MALKA («Mathe lernen und kooperieren von Anfang an»). Es handelt sich dabei um eine Längsschnittstudie über zwei Schuljahre.

Im Forschungsprojekt MALKA wurden Kinder während den beiden Schuljahren 2018/2019 sowie 2019/2020 untersucht. Das Hauptaugenmerk lag dabei auf denjenigen Kindern, die im Schuljahr 2018/19 die 1. Klasse und im Anschluss daran im Schuljahr 2019/2020 die 2. Klasse besucht haben. Das Forschungsprojekt hatte zum Ziel, die Ablösung vom zählenden Rechnen und den Aufbau des Operationsverständnisses zu erforschen.

Im ersten Schuljahr bildete die Förderung flexibler Rechenstrategien das Hauptthema des Forschungsprojekts. Die Kinder wurden dabei mittels eigens für das Projekt vorbereiteter Fördereinheiten unterrichtsintegriert und zum Teil auch in Kleingruppen gefördert. Die Durchführung der unterrichtsintegrierten (kooperativen) Fördereinheiten wurde gefilmt und qualitativ ausgewertet (Diener, in Vorbereitung). Zudem wurden einzelne ausgewählte Kinder in kleineren Gruppen auch individuell gefördert. Die Durchführung dieser Förderung wurde ebenfalls ausgewertet (Hofmann-Villiger, in Vorbereitung). Die Wirkung der verschiedenen Förderungen auf den Lernfortschritt der Kinder wurde quantitativ ausgewertet (Leuenberger, in Vorbereitung).

Im zweiten Schuljahr war die Förderung des Multiplikationsverständnisses das Hauptthema des Forschungsprojekts. Diese Förderung und vor allem auch die Auswertung ihrer Wirksamkeit bilden den Inhalt der vorliegenden Arbeit. Dieser Teil des Forschungsprojekts wird in Kapitel 4.2 beschrieben.

In beiden Schuljahren dienten verschiedene Schulklassen als Kontrollgruppe. In den Klassen der Kontrollgruppe wurde keine mathematische Förderung vorgenommen. Dafür aber wurden

in der Kontrollgruppe soziale Interventionen zur Förderung der Interaktionen zwischen Kindern vorgenommen und die Intensität der Freundschaften in der Klasse mit der Dichte der jeweiligen Klassennetzwerke erfasst (Nesme, in Vorbereitung).

Das Forschungsprojekt MALKA umfasste insgesamt drei leitende Professorinnen, eine Professorin mit Habilitationsprojekt, eine leitende Doktorandin sowie fünf weitere Doktorandinnen und Doktoranden und verschiedene studentische Hilfskräfte, die mehrheitlich am Institut für Erziehungswissenschaften der Universität Zürich tätig waren. Zwei Masterstudierende haben ihre Masterarbeit zu Themen aus dem Forschungsprojekt MALKA geschrieben (Gloor, 2018; Strasser, 2020).

Die involvierten Hochschulen waren die Universität Zürich, die Pädagogische Hochschule St. Gallen, die Pädagogische Hochschule der FHNW in Muttenz, die Hochschule für Heilpädagogik HfH in Zürich sowie die Pädagogische Hochschule PHZH in Zürich.

Das Forschungsprojekt MALKA war ein Nachwuchsförderungsprojekt im Bereich der Fachdidaktiken und wurde finanziert von der staatlichen Institution swissuniversities sowie von den beteiligten Hochschulen.

Design des Projekts MALKA

Im Forschungsprojekt MALKA wurden zu vier verschiedenen Zeitpunkten Kompetenzerfassungen vorgenommen. Die vier Testzeitpunkte dieser Kompetenzerfassung werden im Folgenden mit T1, T2, T3 und T4 abgekürzt. Eine Übersicht über die verschiedenen Kompetenzen, die zu diesen Zeitpunkten erfasst wurden, findet sich in Kapitel 4.1.3.

Zwischen den Kompetenzerfassungen wurden die Kinder auf unterschiedliche Art und Weise unterrichtet und gefördert. Die Schulklassen, die sich am Forschungsprojekt beteiligt haben, wurden dafür in vier Gruppen (I, II, III und IV) eingeteilt.

Im ersten Schuljahr (2018/19) wurden die Kinder in den Klassen der Gruppen I, II und III auf unterschiedliche Art und Weise im Fach Mathematik gefördert. In sämtlichen Klassen der Gruppe I wurden im Laufe des Schuljahrs kooperative Fördereinheiten eingesetzt, in den Klassen der Gruppe II wurden einzelne Kinder individuell gefördert und in der Gruppe III fanden beide Formen der Förderung (kooperative & individuelle) Anwendung. Die Klassen in der Gruppe IV dienten dabei − mit Bezug auf die mathematischen Fähigkeiten − als Kontrollgruppe. In der Gruppe IV kam eine Förderung der sozialen Interaktion ohne Bezug zur Mathematikleistung zum Einsatz. Die verschiedenen Arten der Förderung im ersten Schuljahr des Forschungsprojekts sind im grünen Teil der Abbildung 14 schematisch dargestellt.

Die Testzeitpunkte T2 und T3 wurden aufgeteilt, was den Umfang der Erhebungen zum Testzeitpunkt T2 reduzierte. Die Aufteilung hatte den zusätzlichen Vorteil, dass die v.a. für das zweite Schuljahr relevante Erhebung zum Operationsverständnis zum Testzeitpunkt T3 in das zweite Schuljahr zu liegen kam. Dadurch wurden die durch den Schuljahrswechsel bedingten Änderungen des Bestands (Projektausstiege einzelner Klassen, Zu- und Wegzüge von Kindern in den Projektklassen) im getesteten Bestand bereits berücksichtigt.

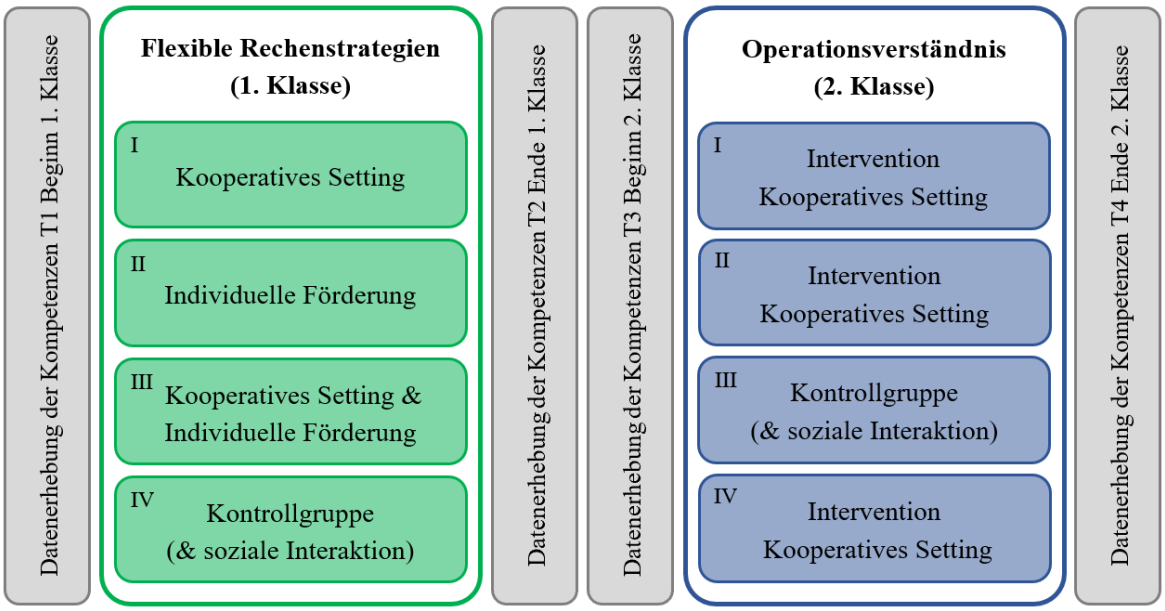


Abbildung 14: Einteilung der Klassen in Gruppen (I, II, III und IV) und Testzeitpunkte T1, T2, T3 und T4

In der zweiten Klasse (Schuljahr 2019/20) wurde die Einteilung in die vier Gruppen I, II, III und IV beibehalten. In den Gruppen I, II und IV wurde mit spezifischen Fördereinheiten (Unterrichtslektionen, siehe Kapitel 4.3) im Mathematikunterricht interveniert. Auch diese Fördereinheiten fanden in unterrichtsintegrierter Form im kooperativen Rahmen (im «kooperativen Setting») im Gesamtunterricht statt. Die verwendeten Fördereinheiten (Unterrichtsbausteine) waren betreffend Struktur und Ablauf vergleichbar mit den kooperativen Fördereinheiten, die in den Gruppen I und III im vorangehenden Schuljahr zum Einsatz gekommen waren. Die Art der Intervention unterschied sich dabei für die Gruppen I, II und IV im zweiten Schuljahr nicht. Die Gruppe III diente im zweiten Schuljahr als Kontrollgruppe. In dieser fand wiederum eine Förderung der sozialen Interaktion ohne Bezug zur Mathematikleistung der Kinder statt. Die Gruppeneinteilung für das zweite Schuljahr ist im blauen Teil von Abbildung 14 ebenfalls schematisch dargestellt.

In beiden Schuljahren wurden die Förderungen von den jeweiligen Kompetenzerfassungen zu Beginn und am Ende des jeweiligen Schuljahrs umrahmt. Die Kompetenzerfassungen fanden jeweils vor (in den Monaten August bis Oktober 2019) bzw. nach (in den Monaten Juni und Juli 2020) Abschluss sämtlicher Interventionen in den Klassen statt. Die Interventionen wurden in den Monaten November 2019 bis Mai 2020 – mit der in der Einleitung erwähnten Unterbrechung aufgrund der Schulschliessung wegen der Covid-19-Situation in der Schweiz – umgesetzt.

Für die vorliegende Arbeit und die damit verbundenen Untersuchungen zum Operations- bzw. Multiplikationsverständnis wurden lediglich die Kinder der zweiten Klassen des Schuljahrs 2019/2020 betrachtet (blauer Teil in Abbildung 14). Für die Auswertung werden aber für diese Arbeit auch die Ergebnisse einzelner Kompetenzerfassungen aus den Testzeitpunkten T1 und T2 verwendet.

4.1.2 Gesamtstichprobe

Die Stichprobengrösse war für die einzelnen Untersuchungen aus mehreren Gründen unterschiedlich. An dieser Stelle werden die Grössen der Stichproben und die Gründe für Schwankungen in deren Grösse dargelegt.

Rekrutierung und ursprüngliche Stichprobengrösse

Das Projekt MALKA hat sich über die Schuljahre 2018/19 sowie 2019/20 erstreckt. In dieser Studie haben Primarschulklassen aus insgesamt 15 Kantonen aus den deutschsprachigen Regionen der Schweiz mitgearbeitet.

Die Primarschulklassen wurden im Frühjahr 2018 rekrutiert. Die Lehrpersonen konnten sich für die Mitarbeit mit ihren Klassen in diesem Forschungsprojekt freiwillig melden. Es ist somit davon auszugehen, dass besonders motivierte, belastbare und an fachdidaktischen Fragen interessierte Lehrpersonen mit ihren Klassen im Forschungsprojekt mitgearbeitet haben. Die Gesamtstichprobe ist insofern nicht gänzlich durch Zufall entstanden. Es handelt sich also um eine nicht-probabilistische Stichprobe, die sowohl als Gelegenheitsstichprobe als auch als Quotenstichprobe bezeichnet werden könnte. Aufgrund der Struktur der Stichprobe mit Kindern in einzelnen Schulklassen ist auch der Begriff der Klumpenstichprobe zutreffend (D. H. Rost, 2007, S. 104 ff., 112; Bortz & Schuster, 2016, S. 81; Bortz & Döring, 2016, S. 314).

Das Ziel der Rekrutierung im Frühjahr 2018 war es, insgesamt 60 Schulklassen für die Teilnahme zu gewinnen. Erfreulicherweise war das Echo auf die Anfrage unter den Lehrpersonen ausgesprochen positiv, so dass das Forschungsprojekt mit insgesamt 80 Schulklassen ins erste Schuljahr starten konnte.

Tabelle 1: Grösse der Gesamtstichprobe (Anzahl Schulklassen)

ursprünglich geplant	Beginn 1. Schuljahr	Beginn 2. Schuljahr	Testzeitpunkt T4
60	80	68	46

Unter den Klassen waren nebst der Mehrheit der Klassen mit Kindern im ersten Schuljahr auch einige Mischklassen. Bei den Mischklassen gab es vorwiegend solche mit Kindern im ersten und im zweiten Schuljahr, aber auch einige wenige mit Kindern im ersten, zweiten und dritten Schuljahr. Ursprünglich hatten mehr als 80 Lehrpersonen ihr Interesse an der Mitarbeit im Forschungsprojekt gemeldet. Um die Gesamtstichprobe zu Schuljahrsbeginn im August 2018 auf 80 Schulklassen zu begrenzen, wurden dann bei den Mischklassen nur diejenigen berücksichtigt, die im Schuljahr 2018/19 eine hohe Anzahl Kinder in der ersten Schulklasse aufwiesen. In einigen der nicht-berücksichtigten Schulklassen konnte Interventionsmaterial für die spätere Phase des Projekts erprobt werden.

Die 80 Klassen, mit denen das Projekt zu Beginn des Schuljahrs 2018/19 anfing, wurden in vier Gruppen zu je 20 Klassen eingeteilt. Bei der Einteilung der Klassen in die vier Gruppen I, II, III und IV wurde auf Verschiedenes geachtet. Insbesondere sollte sich eine gute Streuung

von Schulen in urbanen und ländlichen Gegenden und Schulen mit vielen und wenigen Kindern mit Migrationshintergrund auf die vier Gruppen ergeben. Wünsche von Lehrpersonen (wenn z.B. zwei Lehrpersonen im selben Schulhaus dasselbe Material bearbeiten wollten) wurden − so weit möglich und für den Forschungszweck sinnvoll – berücksichtigt. Auf die Pflege des Bestands und das Eingehen auf die Wünsche und Rückmeldungen der Lehrpersonen wurde stets viel Wert gelegt. Dies geschah im Wissen darum, dass die Forschungsarbeit nebst vielen Anregungen, mehrfachen externen Rückmeldungen zu den Kinderleistungen und fachdidaktischem Material und Mehrwert für die Lehrpersonen auch eine spürbare zusätzliche Belastung darstellen würde.

Nebst der freiwilligen Mitarbeit der Lehrpersonen war auch diejenige der einzelnen Kinder bzw. ihrer Erziehungsberechtigten eine Voraussetzung für die Teilnahme jedes einzelnen Kindes im Projekt. Alle Eltern und Erziehungsberechtigen wurden zu Beginn der Forschungsarbeiten eingeladen zu entscheiden, ob ihr Kind am Forschungsprojekt und den Testungen mitmachen will und darf. Dafür wurde allen Eltern und Erziehungsberechtigten durch die jeweilige Lehrperson eine Einverständniserklärung übergeben, welche sie unterzeichnen konnten. Kinder, für welche diese unterzeichnete Einverständniserklärung nicht vorlag, nahmen im Rahmen des Projekts an keiner der Kompetenzerfassungen teil.

Die Klassenanzahl im Forschungsprojekt wurde ab Beginn des Schuljahrs 2018/19 nicht mehr durch Hinzuziehen neuer Schulklassen erhöht. Hingegen wurden sowohl Zuzüge von Kindern aus anderen Schulklassen in die «MALKA-Klassen» als auch Wegzüge durch das Projektteam erfasst.

4.1.3 Erhebungsinstrumente

Die verschiedenen im Forschungsprojekt eingesetzten Erhebungsinstrumente werden in diesem Kapitel einzeln beschrieben und erläutert. Insbesondere das Instrument für die allgemeinen mathematischen Fähigkeiten wird bei den Auswertungen im Längsschnitt für diese Arbeit eine wesentliche Rolle spielen.

Allgemeine mathematische Fähigkeiten

Die allgemeinen mathematischen Fähigkeiten wurden im Projekt MALKA mit dem Erhebungsinstrument BASIS-MATH erfasst (Moser Opitz et al., 2016, 2019, 2020). Abhängig von der Schulklasse und dem Zeitpunkt (Beginn bzw. Ende des Schuljahrs) wurden unterschiedliche Versionen dieses Erhebungsinstruments eingesetzt. Für die Kinder zu Beginn des ersten Schuljahrs kam das Instrument BASIS-MATH 0+ zum Einsatz. Für Kinder am Ende des ersten oder zu Beginn des zweiten Schuljahrs wurde das Instrument BASIS-MATH 1+ eingesetzt. Kinder am Ende des zweiten Schuljahrs wurden vom Instrument BASIS-MATH 2+ (Moser Opitz et al., 2020) erfasst. Dieses Erhebungsinstrument ist ein Verfahren, welches spezifisch auf rechenschwächere Kinder ausgerichtet ist und darauf zielt, entsprechenden Förderbedarf zu diagnostizieren. Es werden mit verschiedenen Aufgaben zentrale Kompetenzen

der Grundschulmathematik erfasst. In den Bereichen Arithmetik und Sachrechnen wird der Umgang der Kinder mit Zahl und Mass, Operationen und Rechenverfahren geprüft (Moser Opitz et al., 2019, 2020). Einige der eingesetzten Versionen waren zum Erhebungszeitpunkt noch nicht publiziert; eine Publikation ist in Anlehnung an die bereits veröffentlichten Versionen für höhere Schulklassen aber vorgesehen. Das Erhebungsinstrument BASIS-MATH ist auf die Verhältnisse in der Schweiz zugeschnitten und wurde bereits mehrfach erprobt und validiert.

In einzelnen Mischklassen kamen für Kinder ausserhalb der für diese Arbeit wesentlichen Hauptkohorte auch andere Versionen des BASIS-MATH-Erhebungsinstruments zum Einsatz. Dies war beispielsweise der Fall für die erfassten Kinder, die sich (in Mischklassen) im dritten Schuljahr befanden. Mit diesen Kindern wurde am Ende des dritten Schuljahrs die Version BASIS-MATH3+ (Moser Opitz et al., 2019) verwendet.

MALKA-Test

Eigens für das Forschungsprojekt MALKA wurde im Rahmen einer Promotion und einer Masterarbeit ein «MALKA-spezifischer» (und daher auch so genannter) Einzeltest entwickelt (Gloor, 2018; Leuenberger, in Vorbereitung). Dieser Test diente der Erfassung des spontanen Fokus auf Zahlwerte (SFON; «spontaneous focus on numerosity»), der Anzahlerfassung durch die Kinder, deren Fähigkeiten beim Kopfrechnen und deren Rechen- und Zählstrategien. Der Inhalt des Instruments ist in den jeweiligen Arbeiten ausgeführt.

Dieses Erhebungsinstrument wurde zu den Testzeitpunkten T1 und T2 eingesetzt. Die Kinder wurden von der testleitenden Person dafür im Einzelsetting befragt und getestet. Die Durchführung fand mit Hilfe eines Laptops am Computer statt. Die Ergebnisse dieser Befragung wurden zum Zeitpunkt der Befragung durch die testleitende Person direkt elektronisch erfasst.

Arbeitsgedächtnis

Im Rahmen des Projekts MALKA wurde im Einzelsetting mit den Kindern auch ihr Arbeitsgedächtnis erfasst. Für diesen Zweck kam mit der Arbeitsgedächtnistestbatterie AGTB 5-12 ein standardisiertes Instrument zum Einsatz (Hasselhorn et al., 2012).

Die Kinder mussten sich bei dieser Erhebung einen optischen Pfad (in Form von Punkten auf einer Arbeitsunterlage) merken und diesen anschliessend wiederholen. In einem zweiten Teil der Erhebung wurden die Kinder aufgefordert, eine Zahlenfolge, die ihnen vorgesagt wurde, rückwärts zu wiederholen. Die Testleitung führte die jeweiligen Interviews mit den Kindern einzeln durch und erfasste ihre Leistungen laufend in elektronischer Form.

Diese Erhebung fand ausschliesslich zum Testzeitpunkt T1 statt und wurde zu späteren Testzeitpunkten nicht mehr wiederholt. Die Ergebnisse liegen daher für Kinder, die später in die beteiligten Klassen dazu gestossen sind (sogenannte «Zuzüge in die MALKA-Klassen»), nicht vor.

Kognitive Fähigkeiten

Die kognitiven Fähigkeiten der Kinder wurden mit dem standardisierten Grundintelligenztest CFT 1-R erfasst (Weiss & Osterland, 2012).

Diese Erhebung fand in der Gruppe in der Form «paper and pencil» statt. Die Erhebung bestand aus zwei Aufgabenteilen mit zwei bzw. fünf Minuten Bearbeitungszeit. Die Kinder bearbeiteten bei dieser Kompetenzerhebung ein Testheft. Dieses wurde von den testleitenden Personen nach Durchführung der Erhebung eingesammelt und korrigiert. Die Ergebnisse wurden nach einer unabhängigen Zweitkorrektur von einer Hilfskraft in elektronische Form umgewandelt.

Auch diese Erhebung fand nur zum Testzeitpunkt T1 statt und wurde nicht wiederholt.

Soziale Interaktionen

Die Kinder der MALKA-Klassen wurden zu den Testzeitpunkten T1, T2 und T4 zu ihren Freundschaften innerhalb der Klasse befragt. Diese Befragung fand im Einzelsetting zwischen der testleitenden Person und dem Kind statt. Zu den Zeitpunkten T1 und T2 wurde diese Befragung mit allen MALKA-Klassen durchgeführt, bei T4 noch mit deren 36 Klassen. Die Reduktion der Anzahl Klassen bei T4 war eine Folge der Covid-19-Situation im Frühjahr 2020.

Aufgrund dieser Befragung wurde ein Klassennetzwerk gebildet, welches in Nesme (in Vorbereitung) beschrieben wird. Die Auswertung des Längsschnitts wird in Garrote et al. (in Vorbereitung) beschrieben.

Multiplikationsverständnis (Vor- und Nachtest)

Im Rahmen der vorliegenden Arbeit wurde eigens für das Projekt MALKA ein passendes Instrument zur Erhebung des Multiplikationsverständnisses entwickelt. Dieses Erhebungsinstrument besteht aus einem Vor- und einem Nachtest. Es wurde zu den Testzeitpunkten T3 (mit 68 Klassen) und T4 (mit 46 Klassen) im Verlauf des zweiten Schuljahrs eingesetzt.

Diese Erhebung fand in der Klasse in Form eines «paper and pencil»-Tests statt. Klassen mit einer Kinderanzahl von 20 oder mehr wurden für die Testdurchführung in zwei Halbklassen aufgeteilt. Bei kleineren Klassen fand die Erhebung im Plenum statt.

Die Testdurchführung zum Testzeitpunkt T4 erfolgte unter Beachtung eines strengen Schutzkonzepts, dessen Anwendung aber keinen Einfluss auf die Ergebnisse haben sollte. Die Schutzmassnahmen bestanden z.B. aus einem Mindestabstand von 2 Metern, den die testleitende Person zu den Kindern einhalten musste, Korrektur mit Handschuhen usw.

Der Inhalt dieses Tests und die Entwicklungsarbeit, die zu diesem Erhebungsinstrument geführt haben, werden in Kapitel 4.4 dieser Arbeit beschrieben.

Strukturdaten zu den einzelnen Kindern

Von sämtlichen Kindern, die am Projekt teilnehmen wollten und durften, wurden auch Strukturdaten erfasst. Diese wurden von der Lehrperson eingeholt und basierten auf der jeweiligen Einschätzung der Lehrperson. Die Strukturdaten umfassten die unten aufgeführten Variablen, welche mit Hilfe der unten angegebenen Skala von der Lehrperson eingeschätzt wurden.

Die Erhebung dieser Strukturdaten erfolgte im ersten Beobachtungsjahr (im März 2019). Lückenhafte Daten und Daten für neu zugezogene Kinder wurden im zweiten Schuljahr (zwischen Februar 2020 und Juli 2020) erhoben. Mit dieser Nacherhebung wurde der bestehende Strukturdatensatz ergänzt. Die erfassten Strukturdaten sind in Anhang 8.1.1 wiedergegeben.

Überblick über den Einsatz der einzelnen Erhebungsinstrumente

Die Kompetenzerfassungen im Forschungsprojekt MALKA fanden zu vier unterschiedlichen Testzeitpunkten T1, T2, T3 und T4 statt. Eine Übersicht über die einzelnen Testzeitpunkte und die jeweils eingesetzten Erhebungsinstrumente ist in Anhang 8.1.2 wiedergegeben. In der Übersicht in Anhang 8.1.2 sind die Tests zu den einzelnen Testzeitpunkten für die Hauptkohorte der Kinder in der 2. Klasse des Schuljahrs 2019/20 (bzw. 1. Klasse des Schuljahrs 2018/19) aufgeführt. Für die weiteren Kinder aus den Mischklassen kamen zusätzliche Instrumente zum Einsatz, die für die vorliegende Arbeit keine Relevanz haben.

4.1.4 Durchführung der Erhebungen

Die Erhebungen wurden mehrheitlich von studentischen Hilfskräften, zum Teil aber auch von Doktorandinnen und Doktoranden des Forschungsprojekts vorgenommen. Die Mehrheit der studentischen Hilfskräfte wurde von der Universität Zürich gestellt. Die Anzahl an eingesetzten Hilfskräften war unterschiedlich und hing vom Testzeitpunkt, von der Anzahl parallel stattfindender Erhebungen zum jeweiligen Testzeitpunkt, von der Anzahl Klassen sowie von der Verfügbarkeit der Hilfskräfte ab.

Bei den Testzeitpunkten T1 und T4 waren die Erhebungen verhältnismässig umfangreich, da mehrere Erhebungsinstrumente eingesetzt wurden. Zum Testzeitpunkt T2 war der Aufwand etwas geringer als zum Testzeitpunkt T1. Zum Testzeitpunkt T3 kam nur ein Instrument zum Einsatz, weshalb dies zeitlich und im Hinblick auf die eingesetzten Ressourcen die leichteste der vier Erhebungen war. Zum Testzeitpunkt T4 wurde bei verschiedenen Klassen auf die Erhebungen verzichtet, die Erhebungsrunde war aber begleitet von zusätzlichem organisatorischen Aufwand in Zusammenhang mit den Schutzmassnahmen aufgrund der Covid-19-Situation (Kontaktaufnahme mit sämtlichen Lehrpersonen, Erstellung und Einhaltung eines Schutzkonzepts, Besorgung von Schutzmaterialien für die Hilfskräfte, Schulung der Hilfskräfte betreffend Schutzmassnahmen usw.)

Die einzelnen Erhebungen wurden zu allen Testzeitpunkten ohne Unterschied zwischen Klassen der Kontroll- und Klassen der Interventionsgruppe des jeweiligen Jahrs durchgeführt.

Gruppen- und Einzelerhebungen

Die Erhebungen in der Gruppe fanden in Form von «paper and pencil»-Tests in der Gesamtklasse oder in Halbklassen statt. Eine Testdurchführung in der Gruppe dauerte dabei meist rund 30 bis 45 Minuten und erstreckte sich so über eine Unterrichtslektion. Eine Ausnahme bildete die Erhebung der kognitiven Fähigkeiten zum Testzeitpunkt T1, welche ebenfalls in der Gruppe stattfand, jedoch insgesamt nur rund 10 Minuten dauerte.

Bei einzelnen der Erhebungen in der Gruppe wurden die Klassen in Halbklassen aufgeteilt (z.B. bei derjenigen mit dem Instrument BASIS-MATH-0+ zum Testzeitpunkt T1), bei anderen Erhebungen wurde mit mehreren Versionen des Testheftes gearbeitet (so z.B. bei der Erhebung mit dem Instrument BASIS-MATH-1+ am Ende des ersten Schuljahrs). Bei einzelnen Erhebungen wurde die Aufteilung in Halbklassen lediglich bei besonders grossen Klassen vorgenommen (so z.B. beim Vor- und Nachtest zum Multiplikationsverständnis zu den Testzeitpunkten T3 und T4).

Die Erhebungen im Einzelsetting fanden einzeln mit den jeweiligen Kindern und häufig «am Laptop» statt. Die Ergebnisse der Befragung bzw. der Testungen wurden dabei laufend mit Hilfe einer eigens vorbereiteten FileMaker-Maske von der Testleitung in elektronischer Form erfasst. Diese Art der Durchführung fand sowohl bei der sozialen Befragung als auch beim MALKA-Test zu allen drei Testzeitpunkten T1, T2 und T4 Anwendung. Auch die Erhebung des Arbeitsgedächtnisses zum Testzeitpunkt T1 verlief auf diese Art.

Testleiterschulungen und Instruktionen für die Testleitung

Die Anzahl eingesetzter Testleiterinnen und Testleiter betrug rund 20 Personen pro Testzeitpunkt. Die Testleiterinnen und Testleiter wurden im Rahmen einer Schulungsveranstaltung einige Wochen vor ihrem Einsatz während mehrerer Stunden instruiert und auf ihren Einsatz vorbereitet.

Zusätzlich zur Testleitungsschulung wurden allen Testleiterinnen und Testleitern sämtliche Instruktionen für die Erhebungen in ausformulierter, detaillierter Form schriftlich abgegeben. Die Testleiterinnen und Testleiter wurden aufgefordert, die Erhebungen vor dem Einsatz mit einem Kind im Einzelsetting zu üben.

Bei allen Gruppentests wurden die Arbeiten der Kinder von der Testleitung eingesammelt und danach codiert. Für diesen Zweck wurde den Testleiterinnen und Testleitern vorgängig auch eine detaillierte Codieranleitung für die jeweiligen Erhebungen übergeben. Diese Anleitung wurde an den Schulungsveranstaltungen zusammen mit den Anweisungen zur Testdurchführung eingehend besprochen und falls notwendig zusätzlich erläutert.

Bei den Erhebungen im Einzelsetting wurden die Daten jeweils direkt am Laptop mit einer vorgefertigten FileMaker-Maske erfasst.

Die Hilfskräfte wurden für ihren Aufwand im Rahmen des Forschungsprojekts entschädigt.

Zweitkorrektur, Datenübertragung und Datenkontrolle

Die bearbeiteten Unterlagen (Testhefte) aller Gruppentests wurden von einer zweiten Hilfskraft gegenkorrigiert. Sodann wurden die Ergebnisse der Gruppentests von einer Hilfskraft mit Hilfe von eigens für diesen Zweck vorbereiteten FileMaker-Dateien in ein elektronisches Format übertragen («abgetippt»). Diese Übertragungsarbeit wurde wiederum von einer weiteren Hilfskraft in Form von Stichproben überprüft. Die elektronischen FileMaker-Dateien wurden dann in weitere Softwarepakete (SPSS, RStudio, Microsoft Excel) umgewandelt und die Daten entsprechend weiterverarbeitet (Koller et al., 2012; Field, 2013; Luhmann, 2015).

Rückmeldungen an die Lehrpersonen zu den Leistungen der Kinder

Zu allen vier Testzeitpunkten wurden den Lehrpersonen Rückmeldungen zu den mathematischen Leistungen der Kinder und zu den Erkenntnissen aus den sozialen Befragungen zugestellt. Diese Rückmeldungen erfolgten in einem Zeitrahmen von rund ein bis zwei Monaten nach der Erhebung und beschrieben die Leistung jedes einzelnen Kindes. Beim Erhebungsinstrument für das Multiplikationsverständnis wurde dafür eine einfache Beschreibung in Worten verwendet, nämlich das Kind habe «sehr viele», «viele» oder «wenige» Aufgaben richtig gelöst.

Alle diese Rückmeldungen waren verbunden mit dem deutlichen Hinweis an die Lehrpersonen, dass die Ergebnisse nicht zur Beurteilung der Kinder verwendet werden dürften. Trotzdem fanden die Lehrpersonen diese Rückmeldungen hilfreich und schätzten die externe Sicht auf die Leistung ihrer Kinder, die sich – gemäss Angaben der Lehrpersonen – häufig mit ihrer eigenen Sicht gedeckt habe. Die Rückmeldung bei den sozialen Befragungen bestand in einer graphischen Darstellung des «Klassennetzwerks», welcher Informationen zur Dichte und zur Struktur der Freundschaften in der Klasse entnommen werden konnten.

Anonymisierung des Datenaustausches

Die Rückmeldungen erfolgten stets in anonymisierter Form unter Verwendung des fünfstelligen Identifikationscodes, mit dem alle Kinder im Forschungsprojekt identifiziert werden konnten. Die Lehrpersonen verfügten über die jeweilige Klassenliste mit der Zuordnung der Codes zu den einzelnen Kindern. Auf allen Unterlagen, die vom Projektteam an die Lehrpersonen gingen, wurde nur dieser Code verwendet. Listen und Daten wurden im Forschungsprojekt getrennt aufbewahrt, die Listen nur auf Papier, die Daten in elektronischer Form.

Weitere Rückmeldungen

Eltern und Erziehungsberechtigte wurden eingeladen, für Rückmeldungen zu ihrem Kind mit der jeweiligen Lehrperson in Kontakt zu treten. In Einzelfällen und auf explizite Anfrage hin wurden Rückmeldungen auch an interessierte Eltern und Erziehungsberechtigte gegeben, die sich direkt ans Projektteam gewandt hatten.

4.2 Darstellung der Untersuchung und Vorgehen

Ziel der Untersuchungen im zweiten Schuljahr ist die Beantwortung der Forschungsfragen dieser Arbeit gemäss Kapitel 4.2.1.

Der Frage nach dem Lernzuwachs bezüglich der Rechenleistungen in Mathematik wurde mit Hilfe einer Messung dieser Rechenleistungen zu zwei verschiedenen Zeitpunkten nachgegangen. Das Instrument, mit dem diese allgemeinen mathematischen Leistungen («Rechenleistungen») erfasst wurden, ist das Erhebungsinstrument BASIS-MATH (siehe Kapitel 4.1.3).

Zwei Erhebungen mit dem Instrument BASIS-MATH sind für die hier vorliegende Arbeit und für den Lernfortschritt in der 2. Klasse von besonderer Relevanz. Es sind dies die Erhebung für die Hauptkohorte (Kinder in der 1. Klasse des Schuljahrs 2018/19) am Ende des ersten Schuljahrs zum Testzeitpunkt T2 mit dem Instrument BASIS-MATH 1+ sowie die Erhebung am Ende des zweiten Schuljahrs zum Testzeitpunkt T4 mit dem Instrument BASIS-MATH 2+.

Ergänzend konnte mit Hilfe des Erhebungsinstruments für das Multiplikationsverständnis untersucht werden, ob in den Interventionsklassen des 2. Schuljahrs (in denen das Multiplikationsverständnis gefördert wurde), ein höherer Lernzuwachs bezüglich des Multiplikationsverständnisses messbar geworden war als in den Klassen der Kontrollgruppe. Zu diesem Zweck wurde für diese Arbeit ein entsprechendes Instrument entwickelt (Florin et al., 2020; Strasser, 2020). Dieses Instrument wurde in Form eines Vortests vor den Interventionen zu Beginn der 2. Schulklasse zum Testzeitpunkt T3 und später in Form eines Nachtests nach den Interventionen am Ende der 2. Schulklasse zum Testzeitpunkt T4 eingesetzt. Die Analyse dieses Instruments erfolgt in Kapitel 5. Die Ergebnisse der Längsschnittmessung mit diesem Instrument werden in den Kapiteln 6.2 und 6.3 wiedergegeben und dann in Kapitel 7.1 diskutiert.

4.2.1 Forschungsfragen

Der Fokus der hier vorliegenden Arbeit lag auf den Kindern, die im Schuljahr 2019/20 die 2. Primarschulklasse besucht haben. In diesem Schuljahr wurde mit den Kindern im Mathematikunterricht unter anderem die Grundrechenart der Multiplikation (das «Malrechnen») eingeführt.

Für Untersuchungen des Operationsverständnisses und die Einbettung in eine Interventionsstudie eignen sich aufgrund ihres Schwierigkeitsgrads die Operationen zweiter Stufe im besonderen Masse. Die Voraussetzungen, welche die Kinder zu Beginn der Interventionsphase mitbringen, sind einheitlicher als bei der Addition und bei der Subtraktion, bei welchen schon bei Schuleintritt grosse Unterschiede auftreten. In dieser Arbeit wird das Augenmerk auch deshalb auf die Multiplikation gelegt. Die Untersuchungen zum Operationsverständnis beziehen sich hauptsächlich auf das Operationsverständnis für die Multiplikation. Dies geschieht aufgrund des Schwierigkeitsgrads der Multiplikation, aber auch aufgrund der Tatsache, dass diese Operation in der 2. Primarschulklasse eingeführt und vertieft wird. Innerhalb des Forschungsprojekts MALKA (siehe Kapitel 4.1) stellte sich eine grosse Anzahl von Lehrperso-

nen und Kindern in 2. Primarschulklassen für diese Untersuchung freiwillig zur Verfügung. Diese Gelegenheit konnte für die vorliegende Arbeit und die Untersuchung des Operations- bzw. Multiplikationsverständnisses entsprechend genutzt werden.

Das Ziel dieser Arbeit war es, im Rahmen einer Interventionsstudie zu untersuchen, ob eine unterrichtsintegrierte Förderung des Operationsverständnisses zu einer Veränderung des Lernzuwachses bei den mathematischen Fähigkeiten bzw. beim Operationsverständnis der Kinder führen würde.

Die beiden sich daraus ergebenden Forschungsfragen lauteten:

- (FF1) Kann durch gezielte Interventionen zur Förderung des Operationsverständnisses der Lernzuwachs bei der Rechenleistung in Mathematik bzw. der Lernzuwachs beim Operationsverständnis beeinflusst werden?

- (FF2) Ist eine positive Wirkung der Interventionen besonders bei den Kindern mit schwacher Rechenleistung zu beobachten?

4.2.2 Design

Das Ziel der hier vorliegenden Arbeit bestand darin, den Lernzuwachs der Kinder in der 2. Klasse des Schuljahrs 2019/20 im Fach Mathematik zu erfassen und zwischen den verschiedenen Gruppen zu vergleichen.

Damit ist der blaue Teil des Forschungsdesigns aus Abbildung 14 für die hier vorliegende Arbeit zentral. Dieser Teil ist mit den Zeitpunkten für die Datenerhebungen in Abbildung 15 noch einmal wiedergegeben.

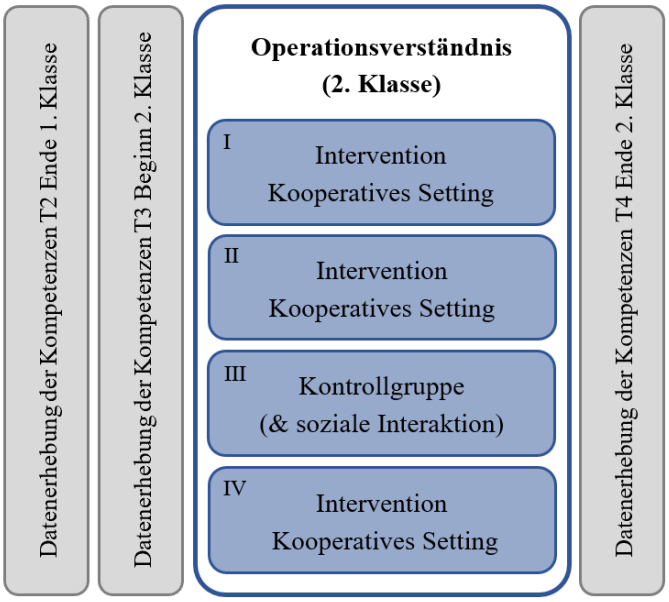


Abbildung 15: Für die vorliegende Arbeit zentraler Teil des Forschungsdesigns

Der Lernzuwachs der Kinder wurde bei zwei verschiedenen abhängigen Variablen (AV) untersucht: Einerseits wurde der Lernzuwachs bei den Rechenleistungen (AV1), andererseits beim Operationsverständnis (AV2) gemessen. Beide in 4.2.1 formulierten Forschungsfragen zielen auf beide Variablen (AV1 und AV2).

Für diese Untersuchung wird hauptsächlich die Unterrichtsgestaltung (Interventionen Ja/Nein) und der entsprechend definierte dichotome Prädiktor als unabhängige Variable (UV) verwendet. Die UV ist also definiert durch unterschiedliche unterrichtliche Bedingungen. Zusätzlich werden weitere Prädiktoren (zusätzliche UV) auf ihren Einfluss hin untersucht. Eine entsprechende Übersicht über die verwendeten Prädiktoren findet sich in Kapitel 4.5.6.

Die Auswahl der Klassen und deren Zuteilung zu den jeweiligen Gruppen erfolgte wie in 4.1.2 beschrieben nicht gänzlich durch Zufall. Eine vollständige Randomisierung wäre für diese Arbeit nicht möglich gewesen. Aus diesem Grunde handelt es sich um eine Studie mit quasi-experimentellem Design (Bortz & Döring, 2016, S. 199 ff.).

Auch zu erwähnen bleibt die Tatsache, dass die Interventionen im Rahmen des tatsächlichen Unterrichts von den Lehrpersonen und mehrheitlich unbeobachtet stattfanden. Aus diesem Grund konnten unerwünschte Einflüsse und Störvariable nur bedingt kontrolliert werden (Bortz & Döring, 2016, S. 196). Eine Zusammenstellung der hierfür getroffenen und für die Lehrpersonen zumutbaren Massnahmen erfolgt in Kapitel 4.2.5.

4.2.3 Stichprobe

Trotz den Bemühungen der verschiedenen Mitarbeitenden im Forschungsprojekt sind einige Lehrpersonen mit ihren Klassen im Laufe des ersten Schuljahrs oder beim Schuljahrswechsel zwischen dem ersten und zweiten untersuchten Schuljahr aus dem Projekt ausgestiegen. Die Gründe für die Projektausstiege waren unterschiedlich. Ein typischer Grund war ein unerwarteter Wechsel der Lehrperson der Klasse. Dies führte in denjenigen Fällen, in denen die neue Lehrperson kein Interesse daran hatte, im Projekt mitzuarbeiten, zu einem Projektausstieg der jeweiligen Klasse. Auch wenn ursprünglich Lehrpersonen gesucht wurden, die für voraussichtlich mindestens zwei Jahre mit der jeweiligen Klasse arbeiten würden, kam es in einigen Fällen zu dieser Situation.

Durch einzelne Abgänge von Lehrpersonen und ihren Klassen während des ersten Schuljahrs oder auch zwischen den beiden Schuljahren sank die Zahl der Klassen im Projekt von ursprünglich 80 auf 68. Zu Beginn des zweiten Schuljahrs sowie zum Testzeitpunkt T3 bestand die untersuchte Stichprobe also noch aus 68 Schulklassen, von denen alle bis zum Ende des zweiten Schuljahrs im Projekt dabeigeblieben sind. Im zweiten Schuljahr wurden somit keine weiteren Abgänge von Schulklassen verzeichnet. Bedingt durch die in der Einleitung bereits erwähnte Ausbreitung des Coronavirus und die damit verbundene Schulschliessung im Frühjahr 2020 wurde aber bei verschiedenen dieser 68 Klassen auf die letzte Erhebungsrunde zum Testzeitpunkt T4 verzichtet. Für diesen Verzicht gab es unterschiedliche Gründe. Bei einem Teil der betroffenen Klassen wurden die Interventionen noch kaum umgesetzt, was entweder zu einem Verzicht der Lehrperson oder zu einer Absage der Erhebungen durch das For-

schungsprojekt geführt hat. Einige Lehrpersonen machten auch die ausserordentlich hohe Arbeitslast für den Verzicht auf die Erhebungen geltend. In einigen weiteren Schulen wurden keine externen Besucher zugelassen, was auch die testleitenden Personen im Forschungsprojekt betraf. In einigen wenigen weiteren Fällen war die gesundheitliche Situation der Lehrpersonen der Grund für den Verzicht auf die Erhebungen. Die für diese Arbeit relevanten Erhebungen (BASIS-MATH-2+ und die Erhebung zum Operationsverständnis) konnten schliesslich mit 46 der 68 Klassen durchgeführt werden.

Tabelle 2: Stichprobe für das 2. Schuljahr (Anzahl Schulklassen)

Zeitpunkt	Beginn 2. Schuljahr (T3)	Ende 2. Schuljahr (T4)
Anzahl Schulklassen (Gesamtbestand)	68	46
davon in der Interventionsgruppe	52	34
davon in der Kontrollgruppe	16	12

Die vier Gruppen I, II, III und IV bestanden ursprünglich aus je 20 Schulklassen. Alle vier Gruppen hatten Abgänge zu verzeichnen. In den Gruppen I, II und IV (Interventionsklassen des 2. Schuljahrs) blieben bis Schuljahrsbeginn 2019/20 deren 51 (von ursprünglich 60) Klassen im Projekt, in der Gruppe III (Kontrollgruppe des 2. Schuljahrs) wurden mit 15 (von ursprünglich 20) Klassen die sozialen Interventionen vorgenommen. Zwei Klassen wurden zu Beginn des zweiten Jahrs eigenhändig der Interventions- bzw. der Kontrollgruppe zugeteilt. Diese zwei Klassen hatten keine klare Gruppenzugehörigkeit im ersten Schuljahr, da in ihnen die geplanten Interventionen des ersten Schuljahrs nicht wie vorgesehen hatten umgesetzt werden können.

Somit wurde in 52 Klassen die mathematische Intervention des zweiten Schuljahrs («unterrichtsintegrierte Förderung des Operationsverständnisses») durchgeführt. Die anderen 16 Schulklassen dienten für die mathematische Förderung als Kontrollgruppe (siehe Tabelle 2). In diesen 16 Schulklassen fand eine Förderung der sozialen Interaktion ohne offensichtlichen Bezug zur Mathematikleistung statt (Garrote et al., in Vorbereitung).

Der Mathematikunterricht in den 16 Klassen der Kontrollgruppe erfolgte wie üblich durch die Lehrperson, ohne dass die Lehrpersonen spezifische Instruktionen für ihren Unterricht bekamen. Einzige Ausnahme war die Bitte an die Lehrpersonen um Protokollierung der Lektionen, in denen die Multiplikation behandelt wurde. Sämtliche Lehrpersonen (sowohl in den Interventions- als auch in den Kontrollgruppenklassen) haben demgemäss die Menge und den Zeitpunkt ihres Mathematikunterrichts, welcher der Multiplikation gewidmet war, protokolliert. Dies diente der Vergleichbarkeit der verschiedenen Klassen.

Die Interventionen des 2. Schuljahrs unterschieden sich nicht in den 52 Klassen. Die 52 Klassen hatten aber im Laufe des Vorjahrs (im ersten Schuljahr im Forschungsprojekt MALKA) unterschiedliche Förderungen der flexiblen Rechenstrategien erhalten. Die Kontrollgruppe des 2. Schuljahrs (Gruppe III) hatte im 1. Schuljahr die intensivste Form der Förderung erhalten,

nämlich eine kooperative Förderung im Gesamtunterricht kombiniert mit der individuellen Förderung für einzelne Kinder (Leuenberger, in Vorbereitung).

Reduktion der Stichprobe durch die Covid-19-Situation im Frühjahr 2020

Der Verzicht auf die Erhebungen zum Testzeitpunkt T4 bei verschiedenen Klassen führte zu einer Reduktion auf einen Umfang von insgesamt 46 Klassen. 34 davon waren Interventionsklassen des 2. Schuljahrs, die anderen 12 Klassen waren Teil der Kontrollgruppe des 2. Schuljahrs (ursprüngliche Gruppe III). Die 34 Interventionsklassen waren gleichmässig auf die Interventionsgruppen des 1. Schuljahrs verteilt (Gruppe I: 11 Klassen, Gruppe II: 11 Klassen, Gruppe IV: 12 Klassen). Diese gleichmässige Verteilung war kein explizites Selektionskriterium für die letzte Erhebungsrunde und hat sich durch Zufall ergeben.

Für Untersuchungen des Längsschnitts im 2. Schuljahr und für die Validierung des Nachtests zum Multiplikationsverständnis reduziert sich somit die Stichprobe von 68 auf 46 Klassen bzw. von etwa 1'100 auf rund 700 Kinder.

Weitere Schwankungen der Stichprobengrösse

Je nach Testzeitpunkt und Auswertung unterscheidet sich die Anzahl erfasster Kinder, da der Bestand der Kinder in den 2. Schulklassen des Jahrs 2019/20 verschiedenen weiteren Schwankungen unterworfen war. Gründe für solche Schwankungen sind u.a.:

- Zuzüge in die MALKA-Klassen nach T1, T2 oder T3
- Wegzüge aus MALKA-Klassen nach T1, T2 oder T3
- Absenzen von Kindern an einzelnen Testzeitpunkten (T1, T2, T3 oder T4)

Obwohl bei Absenzen einzelner Kinder in den meisten Fällen Nachtestungen stattfanden, wurden bei einzelnen Erhebungen aus testökonomischen Gründen jeweils Konzessionen gemacht und nicht alle krankheitsbedingt abwesenden Kinder nachgetestet. Dies hat es erlaubt, Zeit- und Reisekosten der studentischen Hilfskräfte einzusparen.

Die meisten Absenzen von Kindern waren krankheitsbedingt. In sehr wenigen Fällen kam es auch zu absichtlichen Absenzen bzw. zum Verzicht auf die Erhebung für einzelne Kinder, weil die Lehrperson der Ansicht war, dass das Kind mit der Erhebung überfordert gewesen wäre. Noch seltener kam es vor, dass einzelne Kinder trotz ursprünglicher Zusage am Tag der Erhebung die Mitarbeit verweigerten oder während der laufenden Erhebung die Mitarbeit unterbrachen.

Stichprobengrösse für die Validierung des Erhebungsinstruments in dieser Arbeit

Zum Testzeitpunkt T3 wurden 1'084 Kinder im Vortest von der Erhebung zum Operationsverständnis erfasst. Zum Testzeitpunkt T4 wurden dann nach dem Verzicht von 22 der 68

Klassen insgesamt 738 Kinder im Rahmen des Nachtests von der Erhebung zu ebendiesen Fähigkeiten erfasst.

Die Querschnittanalysen und Validierungen der einzelnen Erhebungsinstrumente für das Multiplikationsverständnis erfolgten mit der Gesamtanzahl der Kinder, welche die jeweiligen Erhebungen abgelegt haben, also mit ebendiesen 1'084 Kindern für das Operationsverständnis zum Testzeitpunkt T3 (Vortest) sowie mit 738 Kindern für das Operationsverständnis zum Testzeitpunkt T4 (Nachtest).

Stichprobengrösse für die Längsschnittanalyse in dieser Arbeit

Zum Testzeitpunkt T4 nahmen aus den 46 getesteten Klassen insgesamt 751 verschiedene Kinder an einem der beiden Erhebungen mit dem BASIS-MATH 2+ und dem Instrument für das Multiplikationsverständnis teil. Davon nahmen 730 an beiden Erhebungen teil. 13 Kinder schrieben nur den BASIS-MATH 2+ Test, 8 Kinder nur den Test für das Multiplikationsverständnis. Die Bestände sind deshalb nicht identisch, weil die beiden Erhebungen in der Regel nicht am selben Tag durchgeführt wurden, um einen Ermüdungseffekt der Kinder zu vermeiden. Als Folge davon waren infolge der krankheitsbedingten Wechsel zum Teil verschiedene Kinder an den beiden Erhebungen beteiligt.

Tabelle 3: Stichprobe für das 2. Schuljahr (Anzahl Kinder)

Zeitpunkt	Beginn 2. Schuljahr (T2 bzw. T3)	Ende 2. Schuljahr (T4)
Anzahl Kinder (BASIS-MATH)	1'148	743
Anzahl Kinder (Operationsverständnis)	1'084	738
Anzahl Kinder (jeweils beide Erhebungen)	1'026	730
Anzahl Kinder (alle vier Erhebungen)	680	
Anzahl Kinder ohne Anpassung der Lernziele (alle vier Erhebungen)	669	

Für die Analyse des Längsschnitts in Kapitel 6 wurden nur die Daten derjenigen Kinder verwendet, von denen sowohl die Werte des BASIS-MATH-1+ zum Testzeitpunkt T2 als auch die Daten des Testzeitpunkts T3 und diejenigen des Testzeitpunkts T4 für beide Erhebungsinstrumente (BASIS-MATH 2+ und Operationsverständnis) vorlagen. Die Anzahl Kinder, für welche dies der Fall ist, betrug 680. Diese Kinder wurden in den 46 verschiedenen Klassen unterrichtet, die zum Zeitpunkt T4 noch von den Erhebungen erfasst wurden. Aus dieser Gruppe wurden weitere 11 Kinder mit angepassten Lernzielen (z.B. integrierte Sonderschülerinnen und Sonderschüler) aus der Stichprobe entfernt, so dass die Längsschnittanalysen letztlich mit 669 Kindern in 46 Klassen stattfanden.

Diese Kinderanzahl verringert sich weiter, wenn man die 669 Kinder in Gruppen unterschiedlicher Stärke einteilt. Der Umfang dieser Reduktion bei den einzelnen gerechneten Varianten ist im Ergebnisteil in Kapitel 6 festgehalten.

Ganz- bzw. Mischklassen, Lehrmittel und Verteilung über die Kantone

Die 52 Interventionsklassen des 2. Schuljahrs bestanden aus 37 Klassen, die ausschliesslich Kinder im 2. Schuljahr umfassten, 14 Mischklassen mit Kinder im ersten und zweiten Schuljahr sowie einer Mischklasse, die sowohl Kinder im ersten, im zweiten als auch im dritten Schuljahr umfasste. Insgesamt wurden 838 Kinder in diesen 52 Klassen unterrichtet (Stand: Februar 2020).

Diese 52 Klassen waren auf 13 Kantone verteilt. In 34 dieser Klassen kam das Lehrmittel «Schweizer Zahlenbuch 2» (Wittmann & Müller, 2008) zum Einsatz, während in 17 das Lehrmittel «Mathematik Primarstufe 2» (Brandenberg et al., 2013) verwendet wurde. In einer dieser Klassen ist im Sommer 2019 auf das Lehrmittel «MATHWELT» (Hess, 2018) umgestellt worden, nachdem im Jahr zuvor noch das «Schweizer Zahlenbuch 1» (Wittmann, 2011) eingesetzt worden war.

Unter den 16 Klassen der Kontrollgruppe hatte es 13 Klassen mit Kindern ausschliesslich in zweiten Klassen, 2 Mischklassen, die zusätzlich auch aus Kindern in der ersten Klasse bestanden, sowie einer Mischklasse mit Kindern in allen drei ersten Schuljahren. Insgesamt wurden 268 Kinder in diesen 16 Klassen unterrichtet (Stand Februar 2020).

Diese 16 Klassen verteilten sich auf 4 Kantone. In 6 dieser Klassen kam das Lehrmittel «Schweizer Zahlenbuch 2» zum Einsatz, in den anderen 10 Klassen das Lehrmittel «Mathematik Primarstufe 2».

4.2.4 Eingesetzte Instrumente

Sämtliche Instrumente, die in dieser Arbeit verwendet wurden, sind im Gesamtprojekt («im Forschungsverbund») eingesetzt worden. Von allen eingesetzten Instrumenten wurde bereits in Kapitel 4.1.3 berichtet.

Die für diese Arbeit relevanten Instrumente sind Teil der in Kapitel 4.1.3 aufgeführten und erläuterten Instrumente.

Eine zentrale Rolle für diese Arbeit spielt das Erhebungsinstrument für das Multiplikationsverständnis, welches eigens für die Untersuchung in dieser Arbeit entwickelt wurde. Von dieser Entwicklung wird in Kapitel 4.4 berichtet.

Zudem konnte vor allem die Erhebung der allgemeinen mathematischen Fähigkeiten mit dem Instrument BASIS-MATH für diese Arbeit verwendet werden. Mit Hilfe dieses Instruments wurde der Forschungsfrage nach der Zunahme der allgemeinen Rechenleistung nachgegangen.

Ein Teil der Daten aus den weiteren Erhebungsinstrumenten (z.B. Arbeitsgedächtnis und kognitive Fähigkeiten) sowie ein grosser Teil der Strukturdaten gemäss Anhang 8.1.1 wurden zudem als Prädiktoren verwendet.

4.2.5 Intervention

Im Forschungsprojekt MALKA wurden die in Kapitel 4.3 beschriebenen Fördereinheiten als Intervention in den Schulunterricht verwendet. Diese Fördereinheiten wurden im zweiten Schuljahr (2. Klassen des Schuljahrs 2019/20) zwischen den Erhebungszeitpunkten T3 und T4 in sämtlichen 52 Interventionsklassen eingesetzt.

In den Klassen der Kontrollgruppe hatten die Lehrpersonen hingegen keine Kenntnis von diesen Fördereinheiten. In diesen Klassen erfolgte der Unterricht gemäss Lehrmittel ohne spezifische Instruktion an die Lehrperson.

Aufgrund der Schulschliessung im Frühjahr 2020 konnten nicht in allen Interventionsklassen sämtliche Fördereinheiten eingesetzt werden. Bei der Auswahl der Klassen im Hinblick auf die Erhebung zum Testzeitpunkt T4 wurde auch darauf geachtet, wie weit die jeweilige Klasse mit den Fördereinheiten vor Schulschliessung am 16. März gekommen war. Es wurden dann lediglich Klassen getestet, die bereits vor Schulschliessung mindestens 10 der 16 Fördereinheiten bearbeitet hatten. Verschiedene Lehrpersonen haben nach Wiedereröffnung der Schulen im Mai 2020 die Arbeit an den Fördereinheiten fortgesetzt, so dass einige Klassen vor Einsatz der Erhebungsinstrumente trotz Schulschliessung das gesamte Förderpaket bearbeitet hatten. Einige Lehrpersonen der getesteten Klassen beschlossen jedoch, die Arbeit an den Fördereinheiten aufgrund anderer Prioritäten auszusetzen.

Einführungstreffen für die Lehrpersonen

Vor dem Einsatz wurden die Lehrpersonen aller 52 Klassen anlässlich einer Weiterbildungsveranstaltung (in den Monaten September und Oktober 2019) in die Arbeit mit den Fördereinheiten eingeführt. Diese halbtägige Schulung fand insgesamt fünfmal in mehreren Regionen der Schweiz statt und wurde von zwei Projektmitarbeitern geleitet.

Die Lehrpersonen wurden an diesen Einführungstreffen in die Bedeutung des Operationsverständnisses und in die fachdidaktische Diskussion zu diesem Konstrukt eingeführt.

Sie wurden im Weiteren im Hinblick auf das kooperative Setting instruiert und zur Verwendung der Fördereinheiten angeleitet. Der Ablauf der Unterrichtslektionen in den drei Phasen (Einstiegs-, Arbeits- und Reflexionsphase) wurde den Lehrpersonen dargelegt, die methodischen Formen der Wippe und der Weggabelung (siehe Kapitel 4.3) wurde erläutert. Es wurde besprochen, auf welche Art die Kinderpaare für die Paararbeitsphase gebildet werden sollte, und die Regeln für die Paararbeit wurden thematisiert. Zudem wurden die Ausgestaltungsfreiheiten, die den Lehrpersonen trotz Forschungsdesign blieben, mit ihnen besprochen.

Das Förderpaket mit den 16 Fördereinheiten wurde den Lehrpersonen sowohl in Form eines Ordners als auch in digitaler Form ausgehändigt. Der Aufbau des Förderpaketes wurde dargelegt und begründet.

Ein Zeitplan zur Behandlung der Fördereinheiten im Unterricht (siehe Kapitel 4.3.4) wurde den Lehrpersonen im Projekt vorgeschlagen. Diese waren aber frei, von diesem vorgeschlagenen Zeitplan auch abzuweichen. Sie wurden eingeladen, ihre Behandlung der Multiplikation und auch ihren Einsatz der Fördereinheiten im Laufe des Schuljahrs zu protokollieren. Für diesen Zweck wurden ihnen Protokollbögen ausgehändigt. Diese Protokollbögen waren so ausgestaltet, dass der Aufwand der Lehrpersonen für das Ausfüllen gering bleiben sollte, dass aber dennoch das Ausmass an Unterrichtszeit, die der Multiplikation gewidmet sein würde, daraus hervorgehen würde.

Bei den Einführungstreffen zeigte sich, dass viele Lehrpersonen mit dieser Art der Unterrichtsgestaltung vertraut waren und gemäss eigener Angabe auch sonst mit unterschiedlichen kooperativen Arbeitsformen arbeiteten.

Anlässlich dieser Einführungstreffen wurde eine Einschätzungsübung mit den Lehrpersonen durchgeführt, bei der Kinderarbeiten zum Darstellungswechsel eingeschätzt werden mussten. Die Übung diente als Einstieg in die Thematik und konnte von den Lehrpersonen für zusätzliche Untersuchungen innerhalb des Forschungsprojekts auch abgegeben werden. 64 von 65 Lehrpersonen haben ihre Arbeit dem Projektteam ausgehändigt.

Diese Arbeiten wurden später im Hinblick auf die Heterogenität der Einschätzungen gesichtet. Bei dieser Analyse bestätigte sich die Vermutung, dass die Lehrpersonen sehr unterschiedlich mit Kinderarbeiten und dem Begriff des Operationsverständnisses umgehen.

Tabelle 4: Vorgeschlagener Zeitplan zur Behandlung der 16 Fördereinheiten

Thema	Zeitpunkt der Behandlung	Anzahl Fördereinheiten
Operationsverständnis Addition und Subtraktion	Nov./Dez. 2019	4
Hinführung zur Multiplikation	Dez. 2019/Jan. 2020	2
Operationsverständnis Multiplikation	Jan. bis Apr. 2020	10

Alle Lehrpersonen der 52 Interventionsklassen wurden zu diesen Einführungstreffen eingeladen. Die Teilnahme an den Treffen war sehr hoch und deckte 50 der 52 Klassen ab. Verschiedene Lehrpersonen haben die Weiterbildungsveranstaltungen zusammen mit der schulischen Heilpädagogin/dem schulischen Heilpädagogen besucht. Zwei Lehrpersonen konnten aus terminlichen Gründen an den Einführungstreffen nicht teilnehmen und wurden einzeln besucht und auf dieselbe Art wie die anderen informiert. Die Veranstaltungen verliefen ohne Zwischenfälle und fanden bei den Lehrpersonen ein positives Echo.

Zusammenarbeit mit den Lehrpersonen

Die Fördereinheiten wurden den Lehrpersonen sowohl in elektronischer Form als auch in Papierform ausgehändigt und waren so ausgestaltet, dass sie die Einstiegs- und Reflexionsphase mit den jeweiligen Unterlagen (z.B. unter Einsatz eines Visualizers) gestalten konnten. Die Arbeitsunterlagen für die Paararbeit lagen ebenfalls in elektronischer Form und in Papierform vor. Damit waren die Lehrpersonen mit Kopiervorlagen für die Kinder bzw. die Kinderpaare ausgestattet.

Den Lehrpersonen wurde in Halbtages-Veranstaltungen eine ausführliche Einführung ins kooperative Setting sowie verschiedene präzise Instruktionen und Anweisungen zur Durchführung der einzelnen Fördereinheiten gegeben. Nach diesen Veranstaltungen wurden die Lehrpersonen weiterhin durch das Projektteam begleitet.

Das Projektteam stand den Lehrpersonen jederzeit für Rückfragen zur Verfügung. Die Umsetzung der Fördereinheiten gestaltete sich reibungslos, Rückfragen und Unklarheiten waren selten. Als Erklärung hierfür kann die hohe Zufriedenheit der Lehrpersonen mit dem für sie angefertigten Material angeführt werden. Diese Zufriedenheit kam insbesondere in den telefonischen Interviews zum Ausdruck, die im Verlauf der Monate November 2019 bis Januar 2020, also während der Umsetzung der Fördereinheiten 1 bis 11 mit den Lehrpersonen geführt wurden.

Einsatz der Fördereinheiten und Protokollierung durch die Lehrpersonen

Die 16 Fördereinheiten wurden in allen Schulklassen der Interventionsgruppe im Laufe des Schuljahrs 2019/20 zwischen Oktober 2019 und Juni 2020 eingesetzt.

Die Lehrpersonen haben den Einsatz dieser Fördereinheiten und auch sonstige Unterrichtsstunden zur Multiplikation protokolliert und übersichtsartig festgehalten. Diese Protokolle wurden vom Projektteam gesammelt und für den Zweck dieser Arbeit ausgewertet. Diese Auswertung ergab einen hohe Umsetzungstreue und auch hohe Zufriedenheit mit dem Material. Gelegentlich wurde darauf hingewiesen, dass es wünschenswert wäre, für stärkere Kinder zusätzliches weiterführendes Material zur Verfügung zu haben. Die einzige Einschränkung betreffend Umsetzungstreue betrifft die durch Covid-19 bedingte Situation im Frühjahr 2020, die zu einem Unterbruch bei der Durchführung der Fördereinheiten geführt hat. Nicht alle Lehrpersonen haben bei Wiedereröffnung der Schulen im Mai 2020 die Arbeit mit den Fördereinheiten fortgesetzt. Von den 34 Klassen, welche die Interventionen durchgeführt haben und im Frühjahr getestet wurden, haben aber alle mindestens 10 Fördereinheiten behandelt. In 12 dieser 34 Klassen wurden alle 16 Fördereinheiten behandelt. 25 Klassen kamen bis zur Fördereinheit 13 oder weiter. Die durchschnittliche Anzahl behandelter Fördereinheiten unter den getesteten Klassen lag bei 14. Alle Klassen haben die Fördereinheiten in der vorgegebenen Reihenfolge bearbeitet.

Telefonische Interviews

Zudem wurde mit jeder Lehrperson, die an der Umsetzung der Fördereinheiten in den 52 Klassen der Interventionsgruppe beteiligt war, im Laufe der Umsetzungsphase ein telefonisches Interview geführt. Die Rückmeldung aus diesen Interviews war erfreulich, die Zufriedenheit mit dem Material war ausgesprochen hoch und die Durchführung verlief in allen Klassen nach dem vorgesehenen Zeitplan. Kleinere Abweichungen (z.B. eine längere Abwesenheit eines Kindes und ähnliches) wurden notiert.

Audiographie einer Fördereinheit

Nebst der Protokollierung und dem telefonischen Interview wurden die Lehrpersonen eingeladen, von der Einstiegs- und Reflexionsphase einer Fördereinheit (Nr. 11) eine Tonaufnahme einzureichen. 27 Lehrpersonen kamen dieser Bitte nach, davon 25 vor Schulschliessung im März und zwei zusätzliche im Monat Juni nach Wiederaufnahme der Arbeit an den Fördereinheiten. Die Durchsicht der eingereichten Audiodateien brachte eine sehr enge Anlehnung der Lehrpersonen an das ihnen ausgehändigte Material zu Tage. Ausgehend von den gesichteten Dokumenten und Informationen kann davon ausgegangen werden, dass die Umsetzungstreue durch die Lehrpersonen hoch war. Viele Lehrpersonen waren angesichts ihrer Mitarbeit in einem derartigen Forschungsprojekt besonders bemüht, sich nahe an die Unterrichtsvorlage zu halten.

Rückmeldungen der Lehrpersonen (Erprobung und Projekt)

Die Rückmeldungen der Lehrpersonen zum Unterrichtsmaterial (Fördereinheiten) waren erfreulich und äusserst positiv. Der Neuigkeitswert des Materials wurde wahrgenommen, die Unterrichtsanregung geschätzt. Einige Lehrpersonen erwähnten die Fördereinheiten auch als Hilfestellung bei der Erfassung des Lernstands («Ich lerne die Kinder auf eine neue Art kennen»).

Auch eingegangen sind dabei einige konstruktive Anregungen zur weiteren Ausarbeitung und Verbesserung des Unterrichtsmaterials. Der Wunsch, weitere zusätzliche Arbeitsunterlagen für stärkere Kinder zur Verfügung zu haben, ist ein Beispiel einer solchen Anregung.

Einige Lehrpersonen merkten aber auch an, dass sie auch unabhängig von diesen Fördereinheiten verschiedene «kooperative Unterrichtsformen» mit den Kindern anwenden würden.

4.2.6 Hypothesen

Der Zusammenhang des Operationsverständnisses mit der Rechenleistung ist wie in Kapitel 3.1.2 dargelegt eine der Annahmen, die dieser Forschungsarbeit zugrunde liegt und die in dieser Arbeit anhand eines umfangreichen Datensatzes untersucht wurde. Es wird vermutet, dass eine Förderung des Operationsverständnisses sich auch positiv auf die Rechenleistung bzw. die allgemeinen mathematischen Fähigkeiten der Kinder auswirken kann. Kinder mit einem gefestigten und tragfähigen Operationsverständnis brauchen weniger auswendig zu lernen. Sie können Ergebnisse von Rechenaufgaben nicht nur abrufen, sondern gegebenenfalls auch aus benachbarten oder verwandten Rechenaufgaben selber herleiten. Solcherlei Fähigkeiten müssten sich auch bei den Rechenleistungen der Kinder bemerkbar machen. Diese Vermutung ist eingegangen in die Forschungsfragen in Kapitel 4.2.1 und in die Hypothesen dieser Arbeit, welche in diesem Kapitel festgehalten sind.

Die Hypothesen zu den beiden Forschungsfragen gemäss Kapitel 4.2.1 lauten:

- (H1) Durch gezielte Interventionen zur Förderung des Operationsverständnisses kann der Lernzuwachs bei der Rechenleistung in Mathematik bzw. der Lernzuwachs beim Operationsverständnis beeinflusst werden.

- (H2) Schwächere Kinder, die eine Förderung des Operationsverständnisses erhalten haben, zeigen bei der Rechenleistung in Mathematik sowie auch beim Operationsverständnis einen höheren Lernzuwachs als schwächere Kinder, die keine solche Förderung erhalten haben.

Die erste Hypothese (H1) ist ungerichtet. Es wird angenommen, dass die Intervention in Form einer Förderung des Operationsverständnisses Wirkung zeigt. Die zweite Hypothese (H2) ist gerichtet. Es wird angenommen, dass die Intervention sich gerade bei den schwächeren Kindern positiv auswirkt.

Den beiden Forschungsfragen sowie den dazugehörigen beiden Hypothesen wird nachgegangen, indem der Lernfortschritt der Kinder in der Interventions- und der Kontrollgruppe verglichen wird. Für die Untersuchung der schwächeren Kinder (Forschungsfrage FF2 bzw. Hypothese H2) wird der Bestand jeweils in zwei Stärkeklassen unterteilt. Es finden dann dieselben Vergleiche zwischen Interventions- und Kontrollgruppe sowohl für den Teilbestand der schwächeren als auch für den Teilbestand der stärkeren Kinder statt.

Die für die Längsschnittanalyse verwendete Methodik wird in Kapitel 4.5 beschrieben, die Ergebnisse werden in Kapitel 6 wiedergegeben.

4.3 Fördereinheiten

Im Projekt MALKA wurden im Schuljahr 2019/20 in den 2. Klassen Fördereinheiten eingesetzt, mit denen die Kinder zwischen den Herbstferien 2019 und den Testungen im Frühjahr 2020 von ihren Lehrpersonen im Hinblick auf das Operationsverständnis (Schwerpunkt Multiplikation) gefördert wurden. Diese Förderung fand unterrichtsintegriert im kooperativen Setting statt.

Die dafür notwendigen Fördereinheiten wurden für diese Arbeit eigens entwickelt, da das für eine solche Förderung notwendige Material noch nicht vorlag.

Diese Fördereinheiten bildeten im zweiten Beobachtungsjahr des Forschungsprojekts die Interventionen für den mathematischen Teil der Arbeit.

Unter einer Fördereinheit ist in dieser Arbeit eine vorbereitete Unterrichtslektion, also ein «Unterrichts-Baustein» für eine Unterrichts-Zeitdauer von etwa 30 bis 45 Minuten (eine Schulstunde, eine «Lektion») mit der gesamten Klasse zu verstehen. Die Fördereinheiten und alles Material, das zu diesen gehört, sind für das kooperative Setting ausgelegt. Sie gehen davon aus, dass die Unterrichtslektionen mit einem strukturierten Ablauf von Einstiegs-, Arbeits- und Reflexionsphase ausgestaltet sind.

Die Interventionsklassen des zweiten Schuljahrs im Projekt MALKA unterscheiden sich also von den Klassen aus der Kontrollgruppe insbesondere dadurch, dass bei Ihnen zwischen November 2019 und Juni 2020 mit diesen Fördereinheiten 16 Unterrichtslektionen durchgeführt wurden. Diese 16 Lektionen ersetzten Unterrichtslektionen im Fach Mathematik, welche in der Kontrollgruppe im üblichen Rahmen gemäss Lehrplan von der jeweiligen Lehrperson durchgeführt wurden. Die Interventionsklassen erhielten also keine höhere Anzahl Lektionen im Fach Mathematik. Wie bereits angemerkt gab es in der Interventionsgruppe einzelne Klassen, in denen aufgrund der Covid-19-Situation im Frühjahr 2020 nicht alle 16 Fördereinheiten durchgeführt wurden. In allen Interventionsklassen waren aber mindestens 10, in den meisten auch mehr als 12 Fördereinheiten durchgeführt worden.

Die Fördereinheiten unterschieden sich vom üblichen Unterricht sowohl inhaltlich als auch vom Ablauf her und wurden auch von den Kindern als Spezialunterricht wahrgenommen. Eine häufige Rückmeldung der Lehrpersonen war es, dass sich die Kinder jeweils freuen würden, wenn sie wüssten, dass sie heute wieder «MALKA machen» würden.

4.3.1 Grundlagen der Fördereinheiten

Bei dieser Förderung wurde auf bestehende Fördermaterialien zurück gegriffen, so zum Beispiel auf die Arbeit Häsel-Weide et al. (2017), in welcher zwanzig Bausteine für die Ablösung vom zählenden Rechnen in der zweiten Klasse im kooperativen Setting bereits vorlagen. Das Unterrichtssetting und der didaktische Hintergrund basierten auf dieser Vorarbeit. Insbesondere wurden die Ausführungen und Gedanken zur unterrichtsintegrierten Förderung für die hier vorliegende Arbeit übernommen. Die Absicht dieser Arbeit ist es, diese Ideen auf ein

neues Themengebiet anzuwenden und zugleich eine Aussage betreffend Wirksamkeit der Interventionen zu machen.

Fördermaterialien für das Operationsverständnis lagen zu Beginn der Studie bereits in mehrfacher Ausführung vor, doch verlangte das Setting in dieser Arbeit nach einer Neuentwicklung bzw. Umgestaltung des vorliegenden Materials. Die Kombination der Thematik Operationsverständnis mit der unterrichtsintegrierten Fördermethode im kooperativen Setting stellte eine Neuigkeit dar.

In einzelnen Fördereinheiten wurde aber trotz der Neuartigkeit des Themas auch inhaltlich an vorliegendes Material angeknüpft. So orientiert sich zum Beispiel die erste Fördereinheit stark an Baustein 5 aus Häsel-Weide et al. (2017), die sechste Fördereinheit greift Inhalte und einen Unterrichtsvorschlag aus Gaidoschik (2016a) auf. Auch die Arbeit mit Punktebildern und Punktefeldern, z.B. am Hunderterfeld, ist keine Neuigkeit an sich, sondern eine im Schulalltag weit verbreitete Aktivität. Erst die Einbettung dieses Fördermaterials ins kooperative Unterrichtssetting gibt dem Material Neuigkeitswert. Ebenfalls bemerkenswert ist auch die zeitlich relativ umfangreiche und umfassende und über einen längeren Zeitraum andauernde Behandlung dieser Förderung im Schulunterricht. Für die Wirksamkeit einer solchen Intervention wurde in früheren Studien festgehalten, dass die Zeitdauer nicht zu kurz sein darf. Der längere Zeitraum verhindert auch, dass Lücken unerkannt bleiben und die Lehrperson mit dem Stoff zu schnell weiterfährt, um allfällige starke Kinder nicht zu langweilen.

Im ersten Schuljahr des Projekts wurden mit rund 40 Schulklassen 21 Fördereinheiten zur Förderung flexibler Rechenstrategien behandelt. Die Fördersequenz erwies sich für ein Schuljahr als etwas umfangreich, weil es für Forschungszwecke notwendig war, vor und nach dem Einsatz der Fördereinheiten Zeit zu haben für die Erhebungen. Aus diesem Grund wurde die Fördersequenz von 21 Fördereinheiten auf das zweite Schuljahr hin bewusst auf eine Länge von 16 Fördereinheiten gekürzt.

Aufbau und Struktur der Fördereinheiten, Einsatz im Forschungsprojekt

Das Fördermaterial besteht aus 16 Fördereinheiten, in denen am Darstellungswechsel, an den Grundvorstellungen der Kinder und an ihren Strukturierungsfähigkeiten an Punktefeldern gearbeitet wird. Die Kinder erarbeiten sich ein solides Operationsverständnis durch das Üben von Übersetzungen zwischen Darstellungen in beide Richtungen «zum Term hin» und auch «vom Term weg». Dieses Üben soll sich über einen längeren Zeitraum hinweg erstrecken. Im Gegensatz zur Lernstanderhebung bei den Kindern mittels Vor- und Nachtest kommen in den Fördereinheiten somit ausdrücklich beide Übersetzungsrichtungen, also das «Mathematisieren» und auch das «Interpretieren» zur Anwendung.

In den ersten vier Fördereinheiten wird am Operationsverständnis für die Addition und für die Subtraktion gearbeitet. Für die Kinder kommen damit zuerst einmal vertraute Inhalte zur Sprache, so dass das neue und in verschiedenen Klassen noch ungewohnte kooperative Setting zuerst an bekannten Inhalten eingeübt werden kann. Die entsprechenden Verweise zu den Lehrmitteln führen in die Schulbücher der ersten Klasse (Brandenberg et al., 2015; Wittmann,

2011). Der Einsatz dieser ersten vier Fördereinheiten erfolgte in den meisten Interventionsklassen bis vor den Weihnachtsferien 2019/20.

Im Anschluss an diesen Einstieg anhand der Addition und Subtraktion kamen zwölf Fördereinheiten zum Multiplikationsverständnis zum Einsatz. Dieser Teil der Fördereinheiten setzt Kenntnis der Kinder für den Malpunkt voraus. Die Lehrpersonen des Projekts wurden daher angewiesen, eine kurze Einführung in die Bedeutung des Malpunktes vor oder spätestens parallel zum Einsatz der Fördereinheit 5 (FE05) vorzunehmen.

Fokus auf rechenschwache Kinder

Das Material ist so ausgestaltet, dass dem Operationsverständnis ausreichend Unterrichtszeit zur Verfügung gestellt werden sollte, sodass insbesondere schwächere Kinder vom Material und der Unterrichtsform profitieren können. Diese Ausgestaltung schlägt sich in den Fragestellungen dieser Forschungsarbeit entsprechend nieder.

Die Unterrichtsform verspricht es, die entsprechende Förderung mit Fokus auf die rechenschwachen Kinder vorzunehmen, ohne die stärkeren Kinder dabei zu langweilen. In den leistungsheterogen zusammengesetzten Kinderpaaren kommt den stärkeren Kindern eine Tutorenrolle zu, in der sie ebenfalls gefordert sind und entsprechend gefördert werden. Weitere Aspekte und Vorteile dieser Unterrichtsform finden sich z.B. in Häsel-Weide et al. (2017). Die Befundlage zu diesem Setting ist in dieser Arbeit in Kapitel 4.3.2 zu finden.

4.3.2 Das kooperative Setting als Unterrichtsform

Verwendung des kooperativen Settings in dieser Arbeit

Bei der geplanten Längsschnittstudie erfolgen die Interventionen in der 2. Klasse mittels einer bewährten und bereits mehrfach umgesetzten Unterrichtsform. Bei dieser werden vorbereitete Lektionen («Fördereinheiten») unterrichtsintegriert im Rahmen eines kooperativen Settings behandelt (Häsel-Weide et al., 2017; Wittich, 2017). Die Kinder arbeiten in diesem Setting während den Fördereinheiten zeitweise in kooperierenden («kooperativen») Kinderpaaren. Diese werden von der Lehrperson gebildet. Dabei wird auf eine leistungsheterogene Zusammensetzung der Kinderpaare geachtet. Hauptmerkmal dieses Settings ist die Zusammenarbeit zwischen den Kindern, welche zeitweise den frontalen Unterricht ersetzt. Die Phasen der Paararbeit werden umrahmt von durch die jeweilige Lehrperson moderierten Einstiegs- und Reflexionsphasen, in welchen die Kinder ebenfalls zur Kommunikation und zum Einbringen ihrer Gedanken inklusive Begründungen ermutigt werden. Förderung in ähnlicher Struktur wird u.a. beschrieben in Steenpass & Söbbeke (2009) und in Häsel-Weide et al. (2017).

Verschiedene Überlegungen aus Häsel-Weide & Nührenbörger (2018, S. 67) und einzelne Bestandteile bereits bestehender Fördermaterialien im kooperativen Setting können auch im

Lichte der Förderung des Operationsverständnisses betrachtet werden. Ein Beispiel hierfür bildet Baustein 5 aus Häsel-Weide et. al (2017).

Das kooperative Setting wurde im Themenfeld der Arithmetik der Schuleingangsstufe bereits angewandt und diskutiert (Häsel-Weide et al., 2017; Wittich, 2017; Häsel-Weide & Nührenbörger, 2018).

Zudem sei angemerkt, dass verschiedene Unterrichtsansätze Teile des hier dargelegten kooperativen Settings verwenden und umsetzen, auch ohne dass dies so bezeichnet bzw. als kooperative Lernform ausgegeben wird. In Venkat & Mathews (2019, S. 100) beispielsweise wird eine Intervention verwendet, welche ebenfalls mit einer Einführungsphase, einer Paararbeitsphase, Phasen der Einzelarbeit und einzelnen von der Lehrperson oder entsprechend ausgebildeten Studierenden moderierten Phasen arbeitet. Vor allem Paararbeitsphasen können in verschiedenen Unterrichtsformen auftreten, die nicht explizit dem «kooperativen Lernen» zugeordnet werden.

Neu ist die in dieser Arbeit vorliegende explizite Anwendung dieser Unterrichtsform auf die Thematik des Operationsverständnisses und auf dessen Förderung. Die geplante Förderung des Operationsverständnisses im kooperativen Setting mit wissenschaftlicher Erfassung der Auswirkungen dieser Unterrichtsform auf den Lernzuwachs der Kinder geschieht erstmalig. Auch eine eigentliche Serie von Fördereinheiten bzw. Bausteinen zur eingehenden Behandlung der Thematik Operationsverständnis in diesem Setting bestand vor Ausarbeitung des Materials für die Arbeit aber noch nicht. Für die Untersuchung, welche in dieser Arbeit beschrieben wird, wurden daher Fördereinheiten zur unterrichtsintegrierten Förderung des Operationsverständnisses im kooperativen Setting entwickelt.

Bei diesem Setting handelt es sich um eine Unterrichtsform, die wissenschaftlich bereits untersucht wurde und sich in verschiedenen Themengebieten schon bewährt hat. Es handelt sich aber auch um eine Unterrichtsform, mit der sichergestellt werden kann, dass die Lehrpersonen die Unterrichtslektionen einheitlich durchführen. Auf diese Art kann sichergestellt werden, dass tatsächlich auch in allen Interventionsklassen dieselben Interventionen in den normalen Unterricht eingebaut werden. Daher ist das kooperative Setting nicht nur eine verheissungsvolle Unterrichtsform, sondern auch eine, die sich für Interventionsstudien in besonderem Masse eignet.

Anspruch des kooperativen Settings

Das kooperative Setting kann betrachtet werden im Rahmen der Bemühungen um inklusiven Mathematikunterricht von hoher Qualität und hoher Güte (Scherer & Moser Opitz, 2010; Häsel-Weide, 2016; Korff, 2016; Peter-Koop et al., 2018; Krähenmann et al., 2018; Werner, 2019; Schnepel, 2019, S. 103 ff.). Bei gegebenem Anspruch, starke und schwache Kinder gemeinsam zu unterrichten, werden dabei Anforderungen an den Unterricht gestellt, welchen das kooperative Setting bei erfolgreicher Ausgestaltung genügt bzw. genügen sollte.

Das hohe Ausmass an Kommunikation und Kooperation im Klassenzimmer wird nebst dem Fokus auf fundamentale Lerninhalte, dem Zulassen und Ermöglichen von selbständigen und

individuellen Lernwegen im Rahmen von herausfordernden Lernanlässen in der Literatur explizit als zentraler Aspekt für gelungenen inklusiven Unterricht aufgeführt (Rohrbeck et al., 2003; Korff, 2016, S. 102 f.; Krähenmann et al., 2018; Häsel-Weide & Nührenbörger, 2018, S. 61). Das kooperative Setting stellt einen vielversprechenden Ansatz dar, diesen anspruchsvollen Anforderungen an erfolgreichen und inklusiven Unterricht zu genügen. Die Kinder öffnen und erweitern ihr eigenes Verständnis durch die kommunikative «Konfrontation mit anderen Sichtweisen». Aufgrund dieser Konfrontation kommt es zum Sprechen über verschiedene Deutungen und deren Begründungen (Steenpass & Söbbeke, 2009). Dies führt bei den Kindern zu Lernprozessen.

Das kooperative Setting ist eine mehrfach erprobte und empirisch untersuchte Unterrichtsform. Bei erfolgreicher Umsetzung genügt dieses Setting dem Anspruch, den Lernprozess aus Sicht der Lernenden aktiv, konstruktiv, sozial-interaktiv und individuell zu gestalten. Im kooperativen Setting kommen eigenaktive Konstruktionsprozesse der Lernenden unter Anknüpfung an die individuellen Vorerfahrungen im individuell adaptiven Rahmen zum Tragen (Gerlach, 2007, S. 7, 133).

Der Nutzen und didaktische Wert guter Schülergespräche, die bei erfolgreicher Umsetzung im kooperativen Setting entstehen, wurde ebenfalls bereits untersucht und dokumentiert (Götze, 2007; Gysin, 2013; Häsel-Weide & Nührenbörger, 2018, S. 67).

Kooperation zwischen Kindern ist häufig mit Kommunikation und dadurch mit dem Einsatz von Sprache verbunden. Das kooperative Setting weist auch in Bezug auf andere kritische Aspekte «guten» Unterrichts hohes Potential auf. Es kann bei erfolgreicher Ausgestaltung von positiver Wirkung auf die Motivation der Lernenden ausgegangen werden: die Gespräche der Kinder knüpfen an ihrem Vorwissen an (Hasemann & Stern, 2002, S. 223) und sind von individuell adaptiver Art (Pfister, 2016, S. 39 f.; Pfister et al., 2015, S. 55). Die Kinder lernen wie im Alltag so auch im Klassenzimmer miteinander und voneinander (Gaidoschik, 2016b, S. 73). Viele der Gütekriterien für inklusiven Unterricht bzw. inklusionsförderliche Merkmale des Mathematikunterrichts gemäss Sikora & Voß (2018) sind beim kooperativen Setting erfüllt bzw. vorhanden.

Auch ein solch vielversprechendes Setting kann aber – selbst bei optimaler Vorbereitung und Ausgestaltung − nur dann wirksam werden, wenn die Umsetzung mit der jeweiligen Schulklasse gelingt. Erfolgreiche Förderung ist unter anderem immer abhängig von den jeweiligen Gegebenheiten der Schulklasse bzw. von den einzelnen Kindern (Gaidoschik, 2016b, S. 10). Insofern ist auch das kooperative Setting kein einfaches Rezept für gelingende Förderung, jedoch ein aus den dargelegten Gründen vielversprechender Ansatz hierfür. Das kooperative Setting bleibt selbst bei guter Umsetzung eine aufwändige Unterrichtsform und für die Lehrpersonen eine besondere Herausforderung (Krähenmann et al., 2018, S. 54). Durch das Bereitstellen von ausgearbeitetem Unterrichtsmaterial – so wie es im Projekt MALKA erfolgt ist – kann aber den Lehrpersonen zumindest ein Teil des Vorbereitungsaufwands abgenommen werden.

Lernprozesse sind auch stark durch emotionale Aspekte geprägt (Steiner, 1997; Kretschmann, 2006, S. 46 f.). In Laakmann (2013) beispielsweise wird die Motivation als eine Grundvo-

raussetzung des Lernens beschrieben. Ohne Zweifel gilt diese Aussage für das Operationsverständnis genauso wie für alle anderen Themengebiete der Mathematik und weitere Fächer und Lebensbereiche. Für den Zweck dieser Arbeit sei erwähnt, dass das verwendete Unterrichtssetting auch im Hinblick auf die Motivation der Kinder und die emotionalen Aspekte des Lernens der einzelnen Kinder verheissungsvoll erscheint.

Forschungsstand zum kooperativen Setting

Die Wirksamkeit des kooperativen Settings wurde anhand anderer Themen der Primarschulmathematik bereits verschiedentlich untersucht (Wittich, 2017, S. 74 ff.). Zusammenfassend kann gesagt werden, dass das Setting im Hinblick auf seine Wirksamkeit vielversprechend ist und dass einige Studien wie z.B. Tarim & Akdeniz (2008) positive Effekte belegen. Auch in weiteren Untersuchungen wird das Setting als vielversprechend eingeschätzt (Nührenbörger, 2009; Durlak et al., 2011, S. 418; Häsel-Weide & Nührenbörger, 2016).

Es wird in der Literatur aber auch festgehalten, dass die Ausgestaltung des Settings mittels Setzens verschiedener Schwerpunkte (Souvignier, 2007, S. 138) wie z.B. Qualität der Interaktionen in den Kinderpaaren (Götze, 2007), dem Umgang mit der Heterogenität (Korff, 2016) sowie der Zeitdauer der Interventionen (Slavin, 2008; Wittich, 2017, S. 165 f.) hohe Anforderungen stellt. Sorgfältig ausgestalteten Einstiegs- und Reflexionsphasen sowie auch der Rolle der Lehrperson werden zusätzlich hohe Bedeutung beigemessen (Häsel-Weide et al., 2017; Krähenmann et al., 2018, S. 53).

Sowohl für die Lehrpersonen als auch für die Kinder ist das Setting unter Umständen gewöhnungsbedürftig. Für die Lehrpersonen kann es mit empfundenem Kontrollverlust einhergehen (Wittich, 2017, S. 63). Im Gegensatz zu anderen Unterrichtsformen sind von der Lehrperson eher Moderationskompetenzen gefordert (Krauthausen & Scherer, 2016). Es wurde in der Vergangenheit aber auch schon hohe Zufriedenheit der Lehrpersonen mit dem zur Verfügung gestellten Material beobachtet (Pfister et al., 2015, S. 65).

Es kann auch vorkommen, dass die gemessenen Effekte durch Förderung in dieser Unterrichtsform trotz professioneller Ausgestaltung und erfolgreicher Umsetzung gering ausfallen (Moser Opitz et al., 2018).

4.3.3 Entwicklung der Fördereinheiten

Die Einzelheiten zu den mehrfachen Erprobungen und den Abgleichen mit den gängigen Lehrmitteln der 2. Primarschulklassen sind in diesem Kapitel wiedergegeben.

Entwicklungsarbeiten und Erprobungen

Die Fördereinheiten wurden von November 2017 bis Juli 2019 unter Einbeziehung von Experteneinschätzungen und Rückmeldungen aus der Unterrichtserprobung entwickelt. Es wurden hierfür Expertinnen aus dem Projekt, aber auch drei externe mathematikdidaktische Expertinnen und Experten aus dem In- und Ausland einbezogen. Letzteren ist gemeinsam, dass sie sich in eigenen Entwicklungs- und Forschungsarbeiten bereits mit dem Thema Operationsverständnis auseinandergesetzt hatten. Eine erste Version mit 18 Fördereinheiten wurde im Laufe der Arbeiten auf insgesamt 16 Einheiten gekürzt. Inhalte, Reihenfolgen, Arbeitsmaterialien und didaktische Hinweise wurden mehrfach angepasst. Dabei wurde auch die Anschlussfähigkeit an die im ersten Projektjahr eingesetzten Fördereinheiten beachtet.

Eine Erprobung der Fördereinheiten erfolgte im Schuljahr 2018/19 in mehreren zweiten Klassen in der Deutschschweiz. Die beteiligten Lehrpersonen setzten jeweils eine Auswahl von Fördereinheiten in ihrem Unterricht ein. Die laufend eintreffenden Rückmeldungen der Lehrpersonen wurden im Projektteam ausgewertet und die Ergebnisse bei der Erstellung der Endversion der Fördereinheiten berücksichtigt.

Material und Lehrmittelabgleich

In den Fördereinheiten wurde vielfältiges Bildmaterial verwendet, welches entweder unter ausdrücklicher Angabe der entsprechenden Quellen verwendet oder aber von lizenzfreien Anbietern zur Verfügung gestellt wurde.

Die Fördereinheiten wurden abgeglichen mit zwei der häufigsten Lehrmittel, die in der Schweiz zum Einsatz kamen. Der Einsatz der Fördereinheiten erforderte aber in verschiedenen Klassen eine leichte Anpassung der Jahresplanung insofern, als dass die Multiplikation etwas früher zu behandeln war als sonst üblich. Die Lehrpersonen in der Studie wurden im Hinblick auf den Einsatz der Fördereinheiten entsprechend instruiert und geschult. Ziel dieser Anweisungen war, die Vergleichbarkeit zwischen den Interventionsgruppen und der Kontrollgruppe zu erreichen. Ein weiteres Ziel der Instruktionen bestand darin, dafür zu sorgen, dass die Behandlung der Multiplikation sich im Unterricht über einen längeren Zeitraum von mindestens 14 bis 16 Wochen erstrecken würde. Diese Zeitdauer wird in der Literatur als kritische Zeitdauer für die Wirksamkeit einer derartigen Intervention angesehen (siehe Kapitel 4.3.2).

4.3.4 Aufbau und Umsetzung des Förderpakets

In diesem Kapitel wird dargelegt, wie das gesamte Förderpaket aufgebaut und gegliedert ist. Der Aufbau der einzelnen Fördereinheiten ist dann Inhalt des folgenden Kapitels 4.3.5.

Aufbau des Förderpakets

Mit dem Begriff Förderpaket wird in dieser Arbeit die Gesamtheit der 16 Fördereinheiten zum Operationsverständnis beschrieben.

In dem Förderpaket, das für den Zweck dieser Arbeit entwickelt wurde (Florin et al., 2019), wird das Operationsverständnis so verstanden, wie es im theoretischen Teil dieser Arbeit in Kapitel 2.3.3 dargelegt wurde.

Wesentlich ist dabei die Fähigkeit der Kinder, Verknüpfungen zwischen verschiedenen Darstellungen (Bilder, Sachsituationen, Terme) der Grundoperationen vorzunehmen. An der Fähigkeit der Kinder, diesen **Darstellungswechsel** flexibel und auf sinnvolle Art und Weise vollziehen zu können, wird im Rahmen dieses Förderpaketes gearbeitet. Fördereinheiten, in denen besonders stark am Aufbau dieser Fähigkeit gearbeitet wird, sind in der folgenden Übersicht mit roter Farbe ( ) markiert.

Wie in Kapitel 2.2.5 dieser Arbeit beschrieben ist der Aufbau von **Grundvorstellungen** zu den verschiedenen Rechenoperationen ein damit eng verbundenes Ziel des Unterrichts. Fördereinheiten, die diesen Aspekt in den Vordergrund stellen, sind jeweils in grüner Farbe ( ) markiert.

Operationsverständnis umfasst − wie in Kapitel 2.2.6 beschrieben − auch **Strukturierungsfähigkeiten** der Kinder. Damit gemeint ist die Fähigkeit der Kinder, Beziehungen zwischen Aufgaben und Strukturen (z.B. in bildlichen Darstellungen) zu erkennen und für das flexible Rechnen zu nutzen. Diese Fähigkeiten ermöglichen flexible Darstellungswechsel und stehen somit mit den (oben) in roter Farbe markierten Fähigkeiten in Einklang. Fördereinheiten, welche den Aspekt der Strukturierungsfähigkeiten in besonderem Masse betonen, sind jeweils in blauer Farbe ( ) markiert. Typischerweise sind dies Fördereinheiten, in welchen die Kinder intensiv mit Punktebildern arbeiten.

Die drei Aspekte von Operationsverständnis lassen sich nicht trennscharf voneinander abgrenzen, sondern fliessen − wie in Kapitel 2.3.3 erläutert – ins selbe Konstrukt ein. Es könnte auch die rote Farbe in allen Fördereinheiten verwendet werden, denn auch die Arbeit an den Grundvorstellungen und die Arbeit an Strukturierungsfähigkeiten münden häufig in Übersetzungsprozessen und Darstellungswechseln. Trotzdem kommen bei den verschiedenen Tätigkeiten der Kinder manche Aspekte stärker zur Geltung als andere. In Tabelle 5 und in Tabelle 6 wird festgehalten, welcher der drei Aspekte von Operationsverständnis in der jeweiligen Fördereinheit hauptsächlich zum Tragen kommt.

Die Übersicht ist dabei aufgeteilt in einen ersten Teil (Addition, Subtraktion und Hinführung zur Multiplikation) von 6 Fördereinheiten und einen zweiten Teil (Vertiefung des Multiplikationsverständnisses) von weiteren 10 Fördereinheiten.

Tabelle 5: Operationsverständnis Addition und Subtraktion & Hinführung zur Multiplikation

Fördereinheiten Teil 1 (eingesetzt zwischen Herbst- und Weihnachtsferien 2019)

FE01		Operationsverständnis Addition: Zahldarstellungen am Zwanzigerfeld passende Rechenaufgaben zuordnen
FE02		Operationsverständnis Addition: Plusaufgaben unterschiedlich darstellen
FE03		Operationsverständnis Addition und Subtraktion: Darstellungswechsel zwischen Sachsituation und Term (1)
FE04		Operationsverständnis Addition und Subtraktion: Darstellungswechsel zwischen Sachsituation und Term (2)
FE05		Operationsverständnis Multiplikation: Die Multiplikation als wiederholte Addition
FE06		Operationsverständnis Multiplikation: Umgruppieren von Punktebildern zu neuen Malaufgaben

Die Trennung erfolgte auf diese Art, weil der erste Teil den Lehrpersonen für die Durchführung vor den Weihnachtsferien 2019/20, der zweite Teil für die Durchführung nach diesen Ferien empfohlen wurde. Die Lehrpersonen waren aber bei der detaillierten Zeitplanung frei, von dieser Empfehlung auch abzuweichen.

Didaktischer Hintergrund des Förderpakets

Die Förderung von Operationsverständnis erfolgte so wie in Kapitel 3.3.2 dargelegt. Die Fördereinheiten sollen die Kinder dabei unterstützen, multiplikative Strukturen in verschiedenen Darstellungen zu erkennen. Sie sollten lernen, einen Bezug zwischen einem Multiplikationsterm (einer «Malaufgabe») und einer nicht-symbolischen Darstellung herstellen zu können. Die Aufgabenstellungen fokussieren bewusst nicht auf das (richtige) Ergebnis der jeweiligen Rechenaufgaben, sondern auf den gelungenen Darstellungswechsel. Die Kinder diskutieren dabei die Frage, ob eine nicht-symbolische Darstellung oder eine Sachsituation zu einem angegebenen Term passt oder nicht zu diesem Term passt. Es kann beispielsweise ein Punktefeld von 3 mal 4 Punkten durch Umgruppieren der Punkte auch als 2 mal 6 als «passend» gedeutet werden. Das Bild mit 3 mal 4 Punkten passt aber keinesfalls zur Malaufgabe $4 \cdot 4$.
Auf diese Weise sollen die Kinder eine adäquate Vorstellungen der Multiplikation und somit «Operationsverständnis» aufbauen (Florin et al., 2019).

Bilder und Modelle (z.B. das Zwanzigerfeld inkl. Wendeplättchen) spielen in den Fördereinheiten eine besondere Rolle. Punktebilder eignen sich dabei in besonderem Masse, Beziehungen und Strukturen zwischen Aufgaben zu veranschaulichen. Mit dem Einsatz von Punktebildern kann zum Aufbau der visuellen Strukturierungsfähigkeit beigetragen werden. Diese Fä-

higkeit wiederum trägt zum Operationsverständnis und damit auch zur Ablösung vom zählenden Rechnen bei.

Als Sachsituationen werden Rechengeschichten und insbesondere auch Bild-Text-Geschichten eingesetzt. Während Rechengeschichten ausschliesslich Textelemente umfassen, enthalten die Bild-Text-Geschichten nebst Text auch bildliche Komponenten. Diese bildlichen Komponenten dienen vor allem der sprachlichen Entlastung. Sie liefern Informationen zum Kontext in bildlicher Form oder stellen einen der beiden Faktoren der Malaufgaben bildlich dar. In einem Bild können z.B. drei Äpfel in einem Netz gezeigt werden, während sprachlich die Information geliefert wird, dass Julia fünf solche Netze kauft.

Tabelle 6: Multiplikationsverständnis

Fördereinheiten Teil 2 (eingesetzt im ersten Halbjahr 2020)

FE07	■ ■	Operationsverständnis Multiplikation: Multiplikative Strukturen in Alltags- und Punktebildern
FE08	■ ■	Operationsverständnis Multiplikation: Darstellungswechsel vom (Alltags-) Bild zum Term
FE09	■	Operationsverständnis Multiplikation: Unterschiedliche Malaufgaben im Punktebild finden
FE10	■	Operationsverständnis Multiplikation: Malaufgaben im Hunderterfeld finden
FE11	■ ■	Operationsverständnis Multiplikation: Darstellungswechsel vom Term zum Bild
FE12	■ ■	Operationsverständnis Multiplikation: Darstellungswechsel zwischen Sachsituation und Term
FE13	■	Operationsverständnis Multiplikation: Verschiedene Darstellungswechsel
FE14	■	Operationsverständnis Multiplikation: Malaufgaben verändern
FE15	■	Operationsverständnis Multiplikation: Multiplikative Strukturen im Hunderterfeld erkennen (1)
FE16	■	Operationsverständnis Multiplikation: Multiplikative Strukturen im Hunderterfeld erkennen (2)

Hohe Bedeutung wird der Sprache beigemessen und dies auch dann, wenn keine sprachbasierte Darstellung (wie z.B. eine Rechengeschichte) vorliegt. Die sprachlichen Aspekte kommen während der Arbeitsphasen in den Gesprächen zwischen Kindern (z.B. wenn die Kinder ihr Vorgehen oder ihre Bilder beschreiben) zur Geltung. Sie haben aber auch hohe Bedeutung bei der Begründung der eigenen Überlegungen durch die Kinder in den Einstiegs- und in den Reflexionsphasen (Florin et al., 2019).

Umsetzung der Fördereinheiten

Die Fördereinheiten wurden in einem kooperativen Setting in Anlehnung an Häsel-Weide et. al (2017) umgesetzt. Jede Fördereinheit benötigte etwa 30 bis 45 Minuten Unterrichtszeit und wurde in drei Phasen durchgeführt: Einstiegs-, Arbeits- und Reflexionsphase. Die Einstiegs- und die Reflexionsphase wurde von der Lehrperson im Plenum moderiert. Alle Materialien, die für den Einstieg und die Reflexion (z.B. Aufgaben, Bildmaterial, Punktefelder usw.) notwendig waren, standen als Kopiervorlage zur Verfügung. Die Materialien konnten vergrössert und entweder im Stuhlkreis, an der Wandtafel oder am Visualizer verwendet werden. In der Arbeitsphase arbeiteten leistungsheterogene Kinderpaare im kooperativen Setting (siehe Kapitel 4.3.2) zusammen. Der Unterschied in der Leistungsstärke in Mathematik sollte dabei zwischen den Kindern eines Kinderpaares nicht zu gross sein. Wenn möglich wurde die Zusammensetzung der Kinderpaare für sämtliche Fördereinheiten unverändert beibehalten. Zu Wechseln in der Zusammensetzung kam es nur in begründeten Fällen, z.B. bei disziplinarischen Schwierigkeiten innerhalb eines Kinderpaars. Die Wechsel blieben aber die Ausnahme. Wenn es in der Klasse eine ungerade Anzahl Kinder hatte, so wurde eine Dreiergruppe gebildet.

Das kooperative Setting umfasste zwei verschiedene methodische Formen:

- Weggabelung: Die Kinder arbeiteten zunächst individuell an einer Aufgabe, bevor sie dann ihre Produkte verglichen und gemeinsam weiterarbeiteten.

- Wippe: Die Kinder arbeiteten im Wechsel gemeinsam an einer Aufgabe und übernahmen dabei unterschiedliche Aktivitäten (z.B. Kind A legte eine Aufgabe mit Material, Kind B schrieb die Rechnung dazu). Dabei übernahmen sie unterschiedliche, aber aufeinander bezogene Tätigkeiten.

Die Arbeitsunterlagen für die Arbeitsphase lagen als Kopiervorlage vor. Nur in der Fördereinheit 13 mussten die Kärtchen vorgängig noch durch die Kinder zugeschnitten werden.

Die Lehrpersonen gaben die notwendigen Anweisungen für die Arbeitsphasen jeweils zu Beginn der Arbeitsphase. In diesen Arbeitsphasen sollten die Kinder dann ohne weitere Anweisung und ohne Eingreifen der Lehrperson (zusammen) arbeiten. Die Lehrperson war angehalten, nur dann einzugreifen, wenn das weitere Vorgehen unklar war, nicht aber zur Korrektur «falscher» Gedanken oder Vorschläge.

4.3.5 Aufbau einer einzelnen Fördereinheit

Die einzelnen Fördereinheiten liegen im Moment in Form eines nicht publizierten Dokuments an der PH FHNW vor (Florin et al., 2019).

Anhand einer Fördereinheit (Fördereinheit Nr. 6, genannt «FE06») wird in diesem Kapitel Einblick in den Aufbau dieser Fördereinheiten gegeben. Bei Fördereinheit 6 handelt es sich um eine eher schlichte und kurze Fördereinheit. Sie dient hier als exemplarisches Beispiel für den Aufbau einer Fördereinheit.

Direkte Rede sowohl der Kinder als auch der Lehrperson ist in kursiver Schrift beschrieben. Aussagen der Lehrperson wurden zusätzlich mit dem Zeichen ⇔ markiert. Besonders zentrale Aussagen oder Fragen sind in roter Farbe festgehalten.

Im Anhang dieser Arbeit findet sich ergänzend die Fördereinheit FE06 mit sämtlichen Arbeitsunterlagen (siehe 8.3).

Fördereinheit FE06

Die Struktur und der Aufbau sowie auch der didaktische Hintergrund der Fördereinheit FE06 ist wie auch bei den anderen Fördereinheiten dem Werk Häsel-Weide et al. (2017) entnommen.

Inhaltlich ist die Fördereinheit FE06 stark an Gaidoschik (2016a, S. 48 ff., 65 ff.) angelehnt. In dieser Fördereinheit werden mit und von den Kindern Punktebilder zu neuen Malaufgaben (Multiplikationsaufgaben) umgruppiert. Die Kinder sollen dabei zu einem vorgegebenen Punktebild mindestens eine neue Malaufgabe finden. Sie lernen dabei auch, neue Malaufgaben mit demselben Ergebnis zu entwickeln, indem sie die Punkte des Punktebilds umgruppieren. Handelnd erfahren die Kinder dabei, dass das Ganze (das Produkt bzw. das Ergebnis) sich nicht ändert, auch wenn die Punkte verschoben werden und die bisherigen «gleich grossen Teile» zu neuen «gleich grossen Teilen» umstrukturiert werden. Diese Erkenntnis ist für die Kinder wichtig im Hinblick auf die Entwicklung ihrer Fähigkeit, flexibel zu rechnen.

Der Unterrichtsverlauf wurde in den Vorlagen an die Lehrpersonen sehr detailliert vorgegeben.

In der Fördereinheit FE06 beginnt die Lehrperson den Unterricht an der Wandtafel mit Magneten oder am Visualizer mit Plättchen. Das Tafelbild zu Beginn und am Ende wurde den Lehrpersonen dabei vorgegeben (siehe Abbildung 16 und Abbildung 17).

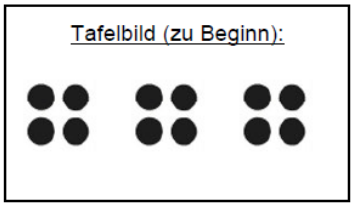


Abbildung 16: Tafelbild zu Beginn der Einstiegsphase der Fördereinheit 6 (FE06)

Für die Gestaltung der Einstiegsphase standen Vorlagen für eine etwaige Verwendung des Visualizers zur Verfügung (siehe Anhang 8.3.2).

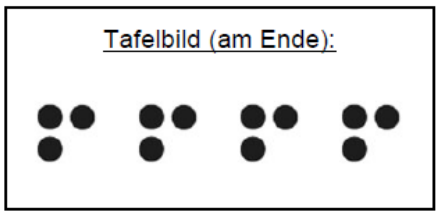


Abbildung 17: Tafelbild am Ende der Einstiegsphase der Fördereinheit 6 (FE06)

Der Unterricht verläuft gemäss einem Skript, das den Lehrpersonen überreicht wurde. Dieses Skript ist für die Fördereinheit 6 in Abbildung 18 zu finden.

In Analogie zur Einstiegsphase wurden auch die Arbeits- und die Reflexionsphasen für die Lehrperson entsprechend dargelegt. Für die Arbeitsphase ist diese Instruktion sowohl im Anhang als auch in Abbildung 19 wiedergegeben.

Einstiegsphase:
Die Lehrperson präsentiert ein Punktebild in Form eines Würfelbilds (mit Magneten an der Wandtafel oder mit Plättchen am Visualizer.

⇔ *Welche Plusaufgabe passt zu diesem Punktebild? Weshalb passt diese Plusaufgabe zum Bild?*
⇔ *Gibt es weitere Plusaufgaben, die zum Bild passen?*
⇔ *Welche Malaufgabe passt zu diesem Punktebild? Kannst du beschreiben und uns zeigen, was du siehst?*
Mögliche Kinderantwort: *Ich sehe 3-mal* (zeigt nacheinander auf die drei Vierergruppen) *4 Plättchen* (umkreist die Vierergruppen mit dem Finger).
Die Lehrperson schreibt die Malaufgabe auf.
⇔ *Wir legen einzelne Plättchen so um, dass eine neue Malaufgabe entsteht. Wir legen dabei möglichst wenige Plättchen um.*
Die Lehrperson beginnt von jeder Viergruppe ein Plättchen (die jeweils rot markierten) zu verschieben:

Das Ergebnis der Verschiebung ist das Folgende:

⇔ *Welche Malaufgabe passt nun?*
Mögliche Kinderantwort: *Ich sehe 4 mal 3 Plättchen, also 4 mal 3.*
⇔ *Können wir noch andere Malaufgaben bilden?*
Die Lehrperson ergänzt:
⇔ *Versuche immer möglichst wenige Plättchen zu bewegen.*
⇔ *Versuche dir vorzustellen, welches Bild du haben möchtest, bevor du die Plättchen bewegst.*
Die Kinder beschreiben das Umgruppieren einzelner Punkte. Die Lehrperson führt die beschriebene Umgruppierung aus.
Die Kinder nennen die entstehenden neuen Malaufgaben:

Mögliche Kinderantwort: *Hier kann man die Malaufgabe 2 · 6 sehen.*

Mögliche Kinderantwort: *Hier kann man die Malaufgabe 6 · 2 sehen.*
Die Malaufgaben 12 · 1 oder 1 · 12 werden nur thematisiert, wenn der Vorschlag von den Kindern kommt.
Die Lehrperson schliesst mit der Beobachtung:
⇔ *Wir haben ganz unterschiedliche Malaufgaben gefunden. Die Anzahl der Plättchen, nämlich 12, bleibt immer gleich.*

Abbildung 18: Skript zur Einstiegsphase der Fördereinheit 6 (FE06)

Arbeitsphase:
Beide Kinder bekommen dieselbe Arbeitsunterlage mit einem Punktebild in Form eines Würfelbilds, das zur Malaufgabe 4 · 5 passt.
Kind A legt dieses Punktebild mit Plättchen nach. Kind B nennt die dazu passende Malaufgabe.
Kind B bildet durch Umgruppieren der Plättchen eine neue Malaufgabe. Kind A nennt die Malaufgabe. Beide Kinder zeichnen das neue Punktebild auf ihrer Arbeitsunterlage ein.
Nun sucht Kind A durch weiteres Umlegen nach einer weiteren Malaufgabe usw.
Zusatzaufgabe:
Kind A denkt sich eine neue Malaufgabe aus und legt die Plättchen entsprechend. Kind B fährt mit den Plättchen fort wie oben und sucht eine neue Malaufgabe usw.
Methodische Form: Wippe

Abbildung 19: Beschreibung der Arbeitsphase der Fördereinheit 6 (FE06)

Material, das für diese einzelnen Phasen notwendig war, wurde genannt bzw. den Lehrpersonen übergeben. In dieser Fördereinheit bestand das notwendige Material aus 20 Plättchen für jedes einzelne Kind gemäss Arbeitsunterlage (siehe Abbildung 20).

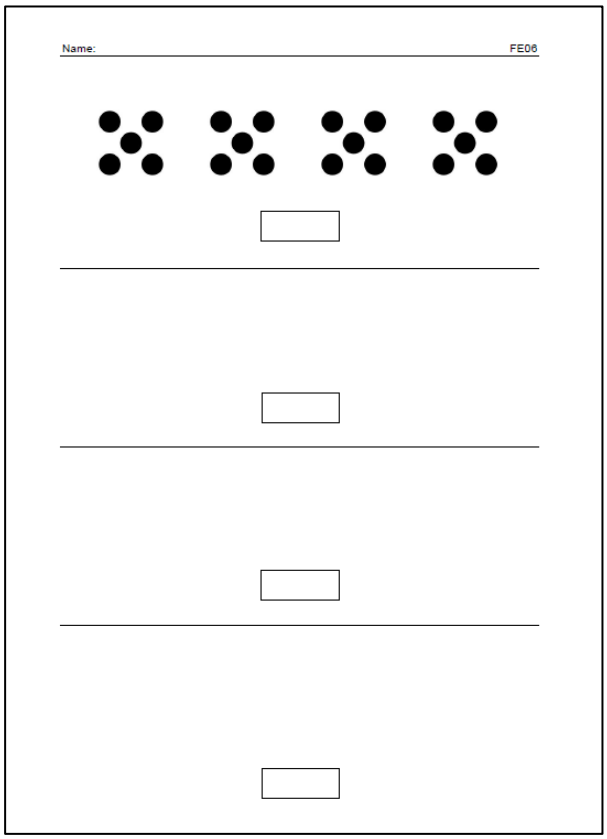


Abbildung 20: Arbeitsunterlage der FE06 für die Kinder (im Format A4)

Die Reflexionsphase wurde analog zur Einstiegsphase ausführlich und unter Vorgabe der direkten Rede ausformuliert. Die entsprechenden Anleitungen finden sich in Anhang 8.3.1.

4.4 Das Erhebungsinstrument für das Multiplikationsverständnis

Nebst dem standardisierten BASIS-MATH-Instrument, das für die Erfassung der Rechenleistung der Kinder eingesetzt wurde, kam im Forschungsprojekt MALKA nebst verschiedenen anderen Tests auch ein spezifischer Test zur Erfassung des Operationsverständnisses zum Einsatz. Von den Entwicklungsarbeiten für diesen spezifischen Test wird in diesem Kapitel berichtet.

4.4.1 Vorarbeiten für einen Test des Multiplikationsverständnisses

Aufgrund der Tatsache, dass Mängel im Operationsverständnis meist im Einzelsetting erfasst und diagnostiziert werden, sind Testinstrumente des Operationsverständnisses für die Anwendung im Klassenverband bisher kaum zu finden. In einem dieser Arbeit vorangehendem Forschungsprojekt der PH FHNW (Projektname: DORE) kam aber bereits ein Test zum Einsatz, mit dem das Operationsverständnis erfasst worden war (Royar et al., 2014). In der Literatur und auf Tagungen wurde davon bereits berichtet, das Erhebungsinstrument wurde aber bisher nicht veröffentlicht. Es handelt sich bei dieser Vorarbeit um einen Test für das Multiplikationsverständnis, der in dritten Klassen eingesetzt wurde. Die Kinder wurden bei der Testdurchführung mit diesem Instrument jeweils eingeladen anzugeben, welche Malaufgabe zu einem vorgegebenen Bild oder einer vorgegebenen Geschichte passe.

Auch in Royar (2013) wird vom Einsatz eines Tests für das Operationsverständnis berichtet. Dieses Instrument umfasst alle vier Grundrechenarten. Die Publikation enthält alle Items inklusive Antworten der getesteten Kinder.

Beim Test für das Operationsverständnis im Rahmen des Projekts MALKA konnte auf diese Vorarbeiten zurückgegriffen werden, sie waren aber aus unterschiedlichen Gründen nur in eingeschränktem Masse für das hier beschriebene Forschungsvorhaben verwertbar. Einer der Hauptgründe hierfür war die Tatsache, dass die vorliegenden Testversionen das Operationsverständnis in der dritten Schulklasse und damit zu einem späteren Unterrichtszeitpunkt – nach Einführung und Behandlung der jeweiligen Operationen im Unterricht – erfassten. Im Rahmen des hier beschriebenen Projekts sollte aber eine erste Erfassung des Lernstands zu Beginn des zweiten Schuljahrs im Rahmen eines Vortests vorgenommen werden. Zu diesem Zeitpunkt war die Multiplikation in den jeweiligen Schulklassen noch nicht eingeführt und behandelt worden. Die Kinder waren zu diesem Zeitpunkt mit dem Wort «mal» noch nicht vertraut. Das Erhebungsinstrument musste also stark überarbeitet bzw. neu entwickelt werden. Ein weiterer Unterschied zu den Vorarbeiten bestand in der sehr hohen Anzahl Kinder, die im Projekt MALKA erfasst wurden. Die Testökonomie und die Einbettung ins Gesamtprojekt erforderten daher einen kurzen Test und eine Durchführung mit Testheften und Instruktionen, ohne dass elektronische Hilfsmittel in Form von PowerPoint-Präsentationen verwendet wurden. Dynamische Items (bei denen wie in den Vorarbeiten mit Hilfe von animierten PowerPoint-Folien Punkte bewegt wurden) konnten im Projekt MALKA nicht eingesetzt werden.

Der Nachtest, der am Ende des zweiten Schuljahrs zum Einsatz kam, sollte das Operationsverständnis für die Multiplikation nach Einführung und Behandlung dieser Grundoperation

erfassen. Dieser Nachtest wiederum musste ausreichend Brückenitems zum Vortest und trotzdem einen ausreichend hohen Schwierigkeitsgrad aufweisen. Somit musste auch der Nachtest auf Grundlage des Vortests neu entwickelt werden (Florin et al., 2020; Strasser, 2020).

Trotz diesen Einschränkungen bei der Verwendbarkeit der Vorarbeiten konnte in hohem Masse von den Vorarbeiten und der damit verbundenen Erfahrung profitiert werden. So beispielsweise bei der Erkenntnis, dass sich die Übersetzungsrichtung des «Interpretierens» («vom Term weg») kaum und wenn dann nur mit hohem Aufwand bzw. mit Interviews testen lässt. Mit dieser Übersetzungsrichtung bzw. mit mündlichen Interviews wurde daher nur in der Erprobungsphase gearbeitet, im Forschungsprojekt wurde dann aber – wie im folgenden Kapitel dargelegt wird – bewusst darauf verzichtet.

Im Weiteren wurde bei der Erprobung einzelner Items auf vorbestehende Literatur zurückgegriffen. Es wurde mit Bildmaterial und mit Ideen gearbeitet, die sich in der Diagnose von Mängeln beim Operationsverständnis und bei der Förderung im Einzelsetting bewährt hatten, so z.B. aus Scherer (2016). Zu den Erfahrungen mit diesem Material wird im folgenden Kapitel eingehend berichtet.

Einige Items blieben – trotz den erwähnten Schwierigkeiten bei der Übertragbarkeit auf das vorliegende Projekt − bis zuletzt inhaltlich und auch mit Blick auf den Testablauf von den vorliegenden Testversionen inspiriert. Es hatte sich insbesondere vom Testablauf her in früheren Tests bewährt, den Kindern die Aufgabenstellung mehrmals deutlich vorzulesen. Dieser Aspekt des Testablaufs konnte für einen Teil des Vor- als auch des Nachtests beibehalten werden. Einige Anregungen für Items wurden auch aus dem Rechenleistungstest «BASIS-MATH» übernommen. Eine dieser Ideen war die Anregung, dass nicht nur Fragen nach dem Term, sondern auch Fragen nach dem Ergebnis einer Malaufgabe dasselbe Konstrukt «Operationsverständnis» messen würden, auch wenn dies so in der mathematikdidaktischen Theorie nicht ausdrücklich formuliert wird. Die späteren quantitativen Analysen stützen diese Hypothese. Durch Einbezug dieses Aufgabentyps konnte das Erhebungsinstrument durch einige zusätzliche Items ergänzt werden.

4.4.2 Entwicklung und Erprobungen des Testinstruments

Die auf Basis der in 4.4.1 erwähnten Vorarbeiten entwickelten Items wurden nach den notwendigen Anpassungen zu einem Erhebungsinstrument zusammengeschnürt. Dieses sollte sich in einer ersten Erprobung bewähren. Verschiedene projekt-interne und auch externe Experten hatten diese frühen Testversionen kommentiert und Einzeltests mit Kindern in der entsprechenden Altersklasse gaben weitere Anregungen zur Verbesserung des Instruments.

Sodann wurden die erste Vortest-Version einer ersten und später einer zweiten Erprobungsrunde unterzogen, deren Ablauf und deren Erkenntnisgewinn in diesem Kapitel wiedergegeben wird.

Die einzelnen Items wurden dann im Laufe der Entwicklungsarbeiten aufgrund der Erkenntnisse aus diesen Erprobungen entweder beibehalten, verbessert oder aber verworfen.

Einschränkung auf eine Übersetzungsrichtung

In beiden Versionen (Vor- und Nachtest) wurden aus Gründen der Testökonomie nur die Übersetzungsrichtung des «Mathematisierens» («zum Term hin») erfasst. Zu den Nachteilen der Rückübersetzung («vom Term weg»), die auch in der Literatur bereits von anderen Autoren genannt werden (Burkhardt, 2008, S. 81; Schulz et al., 2019, S. 16) gehört insbesondere der hohe Aufwand, den es mit sich bringen würde, Zeichnungen und Textteile von Kindern zu einer vorgegebenen Malaufgabe zu erfassen und nach der Testdurchführung dann auf ihre «Passung» hin zu prüfen und zu bewerten. Der hohe Aufwand würde sich bei der Testdurchführung zeigen, da Kinderzeichnungen und das Schreiben von Geschichten in Textform Zeit benötigen und die Testdauer deutlich verlängert hätten. Der hohe Aufwand würde aber zusätzlich auch bei der Korrektur und Codierung zu Buche schlagen. Für jede Kinderzeichnung wäre eine Diskussion und Entscheidung durch mehrere Personen, z.B. durch eine Jury mit mindestens drei Experten, notwendig. Frühere Untersuchungen mit Interviews haben auch gezeigt, dass in vielen Kinderzeichnungen überraschend viele gute und korrekte Gedanken des Kindes eingegangen sind. Für die erwachsene Person waren diese aber ohne weitere Informationen und Erläuterungen des Kindes nicht immer auf Anhieb und zum Teil auch gar nicht als korrekte Gedanken zu erkennen. Folglich wäre für die Beurteilung einer Kinderzeichnung nicht nur eine Entscheidung durch eine Expertenjury, sondern letztlich auch ein Gespräch mit dem Kind, also ein Einzelinterview mit jedem einzelnen Kind notwendig. Dieser Aufwand für die Erhebung und Analyse von Kinderzeichnungen (möglicherweise begleitet von Interviews mit den getesteten Kindern) wäre für eine tiefe Anzahl Kinder zu bewältigen, nicht aber für eine so hohe Anzahl Kinder, wie sie im Projekt MALKA untersucht wurde. Aus diesen Gründen wurde in beiden Erhebungsinstrumenten lediglich die eine der beiden Übersetzungsrichtungen erfasst.

Der Zusammenhang zwischen den beiden Übersetzungsrichtungen wurde anhand von kleineren Stichproben bereits untersucht (Royar et al., 2014, 2015; Strasser, 2020; Florin et al., 2020). In diesen Arbeiten wurden Korrelationen zwischen den beiden Übersetzungsrichtungen gemessen. Die Ergebnisse aus diesen Arbeiten sind unterschiedlich und die Befundlage daher nicht eindeutig. In der Literatur finden sich zudem auch Aussagen dazu, dass die beiden Übersetzungsrichtungen inhaltlich nicht gleichzusetzen sind. Sie sind insbesondere auch nicht gleich schwierig, so dass bei den Kinderleistungen nicht notwendigerweise dieselben Ergebnisse zu erwarten sind. In dieser Arbeit wird die Erhebung des Lernstands wie oben beschrieben hauptsächlich aus testökonomischen Gründen auf die eine Übersetzungsrichtung eingeschränkt. Eine Diskussion der Validität des so entwickelten Instruments wird in Kapitel 4.4.3 vorgenommen. Die Thematik wird auch bei der Diskussion der Ergebnisse in Kapitel 7.1 noch einmal aufgenommen.

Erste Vortest-Version

Der Vortest war zur Lernstands-Erfassung der Kinder zu Beginn der zweiten Klasse vorgesehen. Zu diesem Zeitpunkt sind die meisten Kinder mit dem Malpunkt und der Malrechnung noch nicht vertraut, da die Multiplikation im Unterricht in der Regel bis dahin noch nicht be-

handelt wurde. Diese Tatsache wurde in der Testentwicklung berücksichtigt. Der Vortest enthielt aus diesem Grund auch einige Aufgaben zum Operationsverständnis der beiden Operationen 1. Stufe (Addition und Subtraktion). Diejenigen Aufgaben, in denen multiplikative Strukturen vorkamen, konnten auch unter Verwendung der Addition bearbeitet und gelöst werden. Diese erste erprobte Version des Vortests hatte 31 Items.

Die Kinder wurden im ersten Teil des Tests von der Testleitung durch die Items geführt. Dabei hat die Testleitung mit Hilfe optischer Unterstützung die Aufmerksamkeit der Kinder jeweils auf das entsprechende Item gelenkt und die Textteile und die Aufgabe jeweils zweimal vorgelesen. Dies stellt eine Praxis dar, die auch bei anderen Testdurchführungen in dieser Form angewandt wurde, so z.B. in Venkat & Mathews (2019, S. 100).

Den zweiten Teil des Tests konnten die Kinder selbständig bearbeiten. Die Items in diesem zweiten Teil waren so ausgestaltet, dass die Kinder sie ohne Unterstützung der Testleitung selber verstehen konnten. Dies wurde erreicht durch den Einsatz von Items, die durch rein optische Mittel vermittelt werden konnten (z.B. Punktefelder) oder aber durch Items, bei denen nur minimale Lesefähigkeiten notwendig waren.

Der gesamte Test war so ausgestaltet, dass er in einer Schullektion innerhalb einer Unterrichts-Zeitdauer von etwa 30 Minuten von allen Kindern der Klasse bzw. der Halbklasse bewältigt werden konnte.

Bei einzelnen Items wurde nach einer passenden Rechnung gefragt, bei andern nach einer Zahl, bei manchen Items nach beiden Teilen (Rechnung und Zahl). Die Frage nach der Zahl war bei einigen Items notwendig um den Fokus der Kinder so zu lenkten, dass auch die Frage nach der Rechnung wie gewünscht interpretiert wurde.

Erste Vortest-Erprobung und Auswertung

Die erste Version des Vortests wurde im November 2018 mit Kindern der 2. Klasse des Schuljahrs 2018/2019 erprobt. Diese Erprobung fand mit insgesamt 106 Kindern aus sieben verschiedenen Schulklassen statt. Die sieben Schulklassen waren regional verteilt auf vier verschiedene Kantone der deutschsprachigen Schweiz.

Die verwendete Testversion umfasste 31 Items. Die Durchführung fand in Form eines «paper and pencil»-Tests in Halbklassen zu maximal 12 Kindern statt.

Bei den meisten Items wurde ausgehend von einer Geschichte (mit optischer Unterstützung) oder eines Bilds nach einem Term (einer Rechnung), einem Ergebnis (einer Zahl) oder aber nach beidem gefragt.

Die erste Erprobung des Vortests wurde sowohl mit klassischer Testtheorie als auch mit Item Response Theorie (IRT) ausgewertet. Innerhalb der IRT wurde mit dem Rasch-Modell gearbeitet. Die Items wurden daraufhin untersucht, ob und wie gut sie für die spätere Testauswertung ins Rasch-Modell passen. Es wurden die einzelnen Item-Schwierigkeiten, die Trennschärfen der einzelnen Items sowie die Passung der Items ins Rasch-Modell unter Betrach-

tung der MSQ-Infit-Werte (berechnet mit den R-Packages eRm und TAM) analysiert. Zudem wurde eine Faktoranalyse gerechnet und verschiedene Reliabilitätsmasse berechnet.

Geprüft wurde zudem auch die sprachliche Verständlichkeit der einzelnen Items.

In dieser ersten Erprobung wurde festgestellt, dass einige Items für die Kinder nicht leicht zu verstehen waren. Einige Items wurden daraufhin mit einer leichten sprachlichen Überarbeitung versehen. Abbildung 21 zeigt ein solches Beispiel eines Items. Bei diesem Item erwies es sich als notwendig, die sprachliche Verfeinerung «Vier *davon* schenkt sie ihrem Bruder. » vorzunehmen. Ohne diese Ergänzung war es für einige Kinder nicht klar, ob die vier Stifte eine Teilmenge der zehn ursprünglichen Stifte waren oder nicht. Zu diesem Item haben die Kinder sowohl die Rechnungen «10-4» als auch «10+4» als passend angegeben.

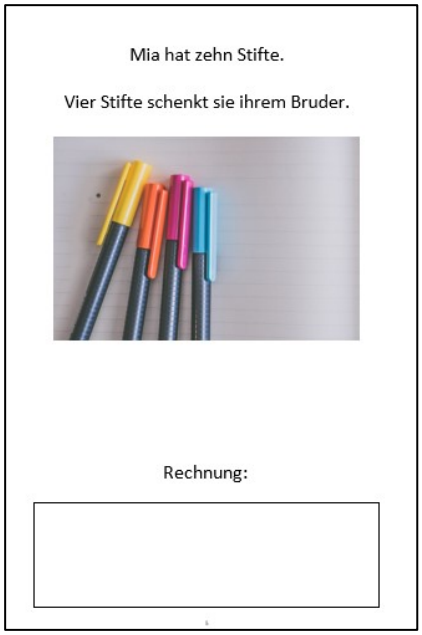


Abbildung 21: Beispiel eines Items aus der Erprobung, welches sprachlich verbessert werden musste; Quelle des Bilds: www.unsplash.com

Einige Items erwiesen sich als sehr leicht und wurden deshalb ersetzt.

Bei der Testerprobung zeigte sich u.a. auch, dass die Kinder die Aufgaben aufgrund des offenen Fragenformats auf sehr vielfältige Art und Weise bearbeiteten. Einzelne Items wurden von den Kindern unterschiedlich interpretiert. Abbildung 22 zeigt ein bewährtes Item aus der Einzelförderung aus Scherer (2016), das sich für die Erfassung des Lernstands in der Gruppe als nicht geeignet erwies, weil die Kinder es nach dem Arbeitsauftrag «Finde eine passende Malrechnung» teilweise mit einer Multiplikation ($3 \cdot 5$), teilweise mit einer Subtraktion ($15 - 4$) in Verbindung brachten. Diese Reaktion der Kinder hatte eine sinnvolle Codierung in 0 und 1 stark erschwert.

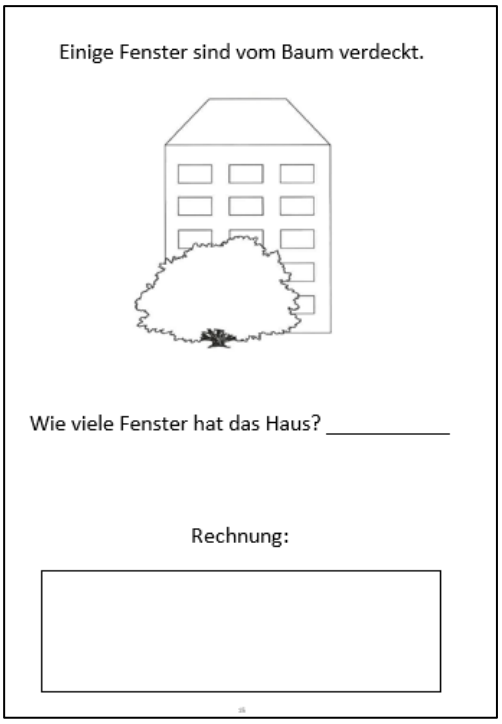


Abbildung 22: Beispiel eines Items aus der Erprobung, welches unterschiedlich interpretiert wurde; Quelle des Bilds: (Scherer, 2016)

Bei einigen weiteren Items erwies es sich als notwendig, während der Erhebung auch nach der Anzahl und nicht nur nach dem Term zu fragen. Dies war beispielsweise beim Item in Abbildung 23 der Fall (Royar et al., 2015). Bei diesem Item zeigte es sich, dass erst die zusätzliche Frage nach der Anzahl Eier die Aufmerksamkeit der Kinder auf die Eier lenkte. Ohne diese Frage wurden als passende Rechnungen auch Terme angegeben, bei welchen der Fokus auf den Anzahl Schachteln lag. So kamen z.B. die Kinderantworten 1+3 (Interpretation: «Eine offene Schachtel, drei geschlossene Schachteln. ») oder 3-1 (Interpretation: «Drei Schachteln. Eine davon wird geöffnet und verwendet. ») angegeben. Diese Schwierigkeit hatte bei der ursprünglichen Verwendung des Items mit Kindern in der 3. Klasse nicht bestanden. Die Frage lautete damals «Welche Malrechnung passt? » und mit dieser Frage war die Zweideutigkeit des Items natürlich nicht vorhanden. Im Projekt MALKA wurde das Item hingegen vor Einführung der Multiplikation mit dem offenen Arbeitsauftrag «Finde eine passende Rechnung. » bearbeitet, was die dargelegte Ergänzung notwendig machte.

Die nach der Erprobung beibehaltenen Items wurden zum Teil sprachlich angepasst und verfeinert, z. B. durch die zusätzliche Frage nach der «Anzahl» (siehe Beispiel in Abbildung 23). Einige zusätzliche Items wurden für die zweite Erprobung auch neu entwickelt.

Der Codierleitfaden wurde so ausgestaltet, dass nicht nur das Ergebnis der angegebenen Terme Relevanz hatte, sondern auch die Struktur der von den Kindern als passend angegebenen Terme berücksichtigt wurde. Dies führte zum Beispiel beim Eier-Bild in Abbildung 23 dazu, dass nur Terme mit 1 codiert wurden, die auch die im Bild dargestellte Struktur aufnahmen (z.B. ein Term wie «6+6+6+6», nicht aber ein Term wie z.B. «17+7» oder «30 – 6» auch

wenn das Ergebnis dieser Rechnungen ebenfalls 24 lautet). In einzelnen Fällen kam es vor, dass beim Entwurf des Codierleitfadens ein Term übersehen wurde, der ebenfalls mit 1 hätte codiert werden sollen. In diesen Fällen wurde der Codierleitfaden im Hinblick auf den späteren Einsatz des Instruments im Projekt ergänzt.

Abbildung 23: Beispiel eines Items aus der Erprobung, welches später ergänzt wurde; Quelle des Bilds: (Royar et al., 2014)

Zweite Vortest-Version: erneute Erprobung, Auswertung und Anpassungen

Basierend auf allen Erkenntnissen aus der ersten Erprobung des Vortests wurde eine zweite Version des Vortests entworfen, die sich in einer zweiten Erprobung bewähren sollte. Diese neue Vortest-Version umfasste 30 Items. Das Material wurde wiederum ergänzt durch eine Anleitung zur Testdurchführung und durch einen Codierleitfaden.

Die zweite Version des Vortests wurde im Mai 2019 mit Kindern am Ende der 1. Klasse des Schuljahrs 2018/2019 erprobt. Auch bei dieser Durchführung kamen Testhefte zum Einsatz, welche die Kinder dann mit Bleistift bearbeiteten («paper and pencil»-Test). Die meisten Erprobungen fanden in Halbklassen zu maximal 12 Kindern statt. In einzelnen Fällen wurde die Erprobung in der Gesamtklasse und in grösseren Kindergruppen durchgeführt, um die Machbarkeit auch dieser ökonomischeren Form der Testdurchführung und den Einfluss auf den Testverlauf zu erproben.

Die eingesetzte zweite Version des Vortests entsprach vom Ablauf her der ersten Version, allerdings waren verschiedene Items aufgrund der Erfahrungen in der ersten Erprobung ersetzt oder verfeinert worden. Einige Items kamen neu dazu und auch die zweite Version des Tests

war vom Umfang her etwas länger als die angestrebte finale Version des Vortests. Neu waren zwei Items auf derselben Seite des in Form einer Broschüre im Format A5 gedruckten Testhefts abgebildet. Dieses Format hat sich bewährt. Die Verwendung von zwei Items auf jeder Seite führte zu keinerlei Schwierigkeiten. Mit dieser Änderung konnte die Seitenzahl des Testhefts von 28 auf 12 reduziert werden. Das Format wurde dann im Projekt so beibehalten.

Die Stichprobe der zweiten Vortest-Erprobung umfasste 160 Kinder. Diese Kinder besuchten 13 verschiedene Schulklassen. Die Schulklassen waren auf sieben verschiedene Schulhäuser und auf vier verschiedene Kantone aus der deutschsprachigen Schweiz verteilt.

Aufgrund der Auswertung der Stichprobe wurden vier weitere Items aus dem Test entfernt. Zwei hatten sich als zu leicht erwiesen, ein weiteres hatte einen schlechten Model-Fit-Wert im Rasch-Modell und wurde zudem nicht in allen Klassen einwandfrei verstanden, da die Kinder den Begriff «Malaufgabe» mehrheitlich noch nicht kannten. Ein viertes Item wurde gestrichen, um einen stärkeren Fokus auf Items mit multiplikativen Strukturen zu erreichen.

Bei einem weiteren Item wurde die Codierung verfeinert, da wiederum eine gute Kinderlösung nicht vorausgesehen worden war. Diese Verfeinerung führte dazu, dass das Item besser ins Rasch-Modell passte.

Auf Basis dieser zweiten Erprobung wurde der Vortest für den Einsatz im Forschungsprojekt in den Monaten August und September 2019 erstellt. Das Instrument umfasste somit 26 Items (Fragen nach Zahl und Term bei gleicher Ausgangssituation zweifach gezählt). Insgesamt wurden den Kindern 19 Ausgangssituationen vorgelegt.

Vortest – finale Version

In der finalen Version des Vortests kam ein Einstiegs-Item zum Operationsverständnis der Addition vor. Der Test umfasst weitere vier Items zum Operationsverständnis der Subtraktion sowie 21 Items zum Multiplikationsverständnis. Die 21 Items mit multiplikativer Struktur konnten alle auch durch das Nennen von passenden Additionstermen («Plusaufgaben») gelöst werden.

Die einzelnen Items der finalen Version des Vortests sind im Anhang 8.2.1 zu finden.

Erste Nachtest-Version: Erprobung und Auswertung

Im Rahmen einer studentischen Masterarbeit wurde basierend auf der finalen Vortest-Version ein Nachtest entworfen (Strasser, 2020). Dieser war vorgesehen für das Ende der zweiten Klasse (Testzeitpunkt T4 im Projekt MALKA) und kam zum Einsatz, nachdem die Multiplikation mit allen Kindern des 2. Schuljahrs behandelt worden war. Von der Entwicklung dieses Nachtests wurde bereits berichtet in Florin et al. (2020).

Als Brückenitems wurden unter den Items des Vortests vor allem solche ausgesucht, die sich im Vortest als besonders schwierig erwiesen hatten. Diese Wahl erfolgte in der Absicht, Items

zu verwenden, die dann im Nachtest – der rund acht Monate später zum Einsatz kommen sollte – nicht zu einfach sein würden.

Von der Nachtest-Erprobung im Juni 2019 mit Kinder der 2. Klasse des Schuljahrs 2018/19 wird berichtet in Strasser (2020). Diese Arbeit umfasst auch die Auswertung dieser Erprobung mit klassischer Testtheorie. Die Auswertung wurde im Rahmen der hier vorliegenden Arbeit und im Hinblick auf die Ausgestaltung des Nachtests ergänzt durch eine Überprüfung der Items im Hinblick auf ihre Passung ins Rasch-Modell der IRT.

Anpassung des Nachtests und Einsatz im Forschungsprojekt

Die erprobte Version des Nachtests wurde vor dem Einsatz im Forschungsprojekt MALKA in verschiedenen Hinsichten angepasst. Einzelne Items wurden gestrichen, insbesondere solche, die sich für die Kinder als sehr leicht erwiesen hatten. Das neue Item-Format mit den Multiple-Choice Items wurde überarbeitet und angepasst. Der erprobte Nachtest wurde so stark gekürzt, dass die finale Version des Nachtests denselben Umfang wie der finale Vortest hatte.

Der Nachtest war für Kinder am Ende der zweiten Klasse vorgesehen und wurde im Projekt MALKA im Frühjahr 2020 mit den Kindern am Ende der 2. Klasse eingesetzt.

Zu diesem Zeitpunkt war die Multiplikation mit den Kindern in allen Schulklassen der Studie im Unterricht eingeführt, geübt und behandelt worden. Die Kinder sind am Ende der 2. Klasse mit dem Malpunkt vertraut und setzen ihn entsprechend auch in ihren Bearbeitungen der Aufgaben ein. Neu wird gegen Ende des Nachtests auch ausdrücklich nach «Malrechnungen» gefragt («Finde eine passende Malrechnung»). Diese Frage kam aber erst gegen Ende des Instruments zur Anwendung. So wurde sichergestellt, dass für die Brückenitems keine anderen Instruktionen als im Vortest gegeben wurden. Es wurde darauf geachtet, dass die Aufmerksamkeit der Kinder nicht frühzeitig auf «Malrechnungen» gelenkt wurde.

Der im Projekt MALKA eingesetzte Nachtest umfasste 26 Items. Bei dieser Zählung wurde die Frage nach der Anzahl und nach dem Term bei derselben Ausgangssituation wiederum doppelt gezählt. Von diesen 26 Items waren 12 Brückenitems zum eingesetzten Vortest. Diese 12 Brückenitems waren auf 7 Ausgangssituationen verteilt. Die Struktur des Tests und der Ablauf orientierten sich am Vortest. Die ersten 16 Items wurden den Kindern vorgelesen, weitere 10 Items wurden von den Kindern selbständig bearbeitet. Im Nachtest kamen keine Items zum Operationsverständnis der Addition, lediglich zwei Items (eine Ausgangssituation) zum Operationsverständnis der Subtraktion sowie 24 Items zum Multiplikationsverständnis zum Einsatz. Die Items zur Multiplikation konnten weiterhin auch unter Nennung eines Additionsterms gelöst wurden. Auch im Nachtest wurden die Kinder bei den ersten 16 Items mit der Aufforderung «Finde eine passende Rechnung. » beauftragt.

4.4.3 Beurteilung und Diskussion der Testgütekriterien

In vielen Werken der Psychologie und der Sozialwissenschaften werden die Gütekriterien für einen empirischen Test beschrieben. Einig ist sich die Literatur dahingehend, dass es drei Hauptgütekriterien (die Objektivität, die Reliabilität und die Validität des Testinstruments) zu prüfen gilt und dazu verschiedene Nebengütekriterien (die Normierung, die Skalierung, die Ökonomie, die Nützlichkeit, die Zumutbarkeit, die Unverfälschbarkeit, die Fairness, die Vergleichbarkeit und die Transparenz) ebenfalls zu beachten bleiben (Pospeschill, 2010; Moosbrugger & Kelava, 2012; Eid & Schmidt, 2014; Bortz & Döring, 2016, S. 81 ff.). Die Literatur ist sich einig hinsichtlich der drei Hauptgütekriterien, bei der Anzahl und der Beschreibung der Nebengütekriterien gibt es leichte Unterschiede zwischen den verschiedenen Werken.

Im Folgenden werden die drei Hauptgütekriterien der Testgestaltung inhaltlich kurz beschrieben. Sodann wird das neu entwickelte Testinstrument im Hinblick auf diese drei Hauptkriterien diskutiert.

Objektivität

Das Testgütekriterium der Objektivität stellt sicher, dass die Testleistungen verschiedener Probanden vergleichbar sind. Um Objektivität eines Erhebungsinstruments zu erreichen, ist es notwendig, dass das Ergebnis der Messung (des gemessenen Merkmals) unabhängig ist von der Person, welche den Test leitet, und auch unabhängig ist von der Person, welche den Test auswertet (Moosbrugger & Kelava, 2012, S. 8 ff.). Es werden in der Fachliteratur typischerweise drei Aspekte der Objektivität unterschieden: die Durchführungsobjektivität, die Auswertungsobjektivität und die Interpretationsobjektivität.

Die Objektivität wurde in der Testentwicklung und in der Anwendung des Erhebungsinstruments im Projekt sichergestellt, indem eine präzise und detaillierte Testanleitung an die testleitenden Personen abgegeben wurde (Durchführungsobjektivität). Diese Testanleitung wurde an den Testleitungsschulungen eingehend besprochen. Sämtliche testleitenden Personen haben diese Schulungsveranstaltungen besucht. Im Falle von Absenzen fanden zusätzliche Treffen statt, bei denen diese Schulung im Einzelgespräch nachgeholt wurde. Die Gesamttestdauer wurde vorgegeben und pro Item wurde eine zeitliche Vorgabe gemacht. Klassengrössen waren im Voraus festgelegt. Die Auswertungsobjektivität wurde sichergestellt, indem den testleitenden Personen eine präzise und detaillierte Codieranleitung ausgehändigt wurde. Auch diese wurde an den Schulungen vorgängig mit sämtlichen Personen, welche Testleitungen vornahmen, besprochen. Zudem wurden alle korrigierten Testhefte von einer Zweitperson begutachtet und die Erstkorrektur gegebenenfalls angepasst. Die Testleitung war angehalten, allfällige Unklarheiten zu markieren, um eine einheitliche Handhabe in allen Fällen sicherzustellen. Viele der eingesetzten Hilfskräfte kamen zudem zu verschiedenen Testzeitpunkten zum Einsatz, so dass sie sich an das Material und die Art der Korrektur und des Arbeitens gewöhnen konnten. Dies stellte auch eine Auswertungsobjektivität über die Zeit sicher. Einige der eingesetzten Hilfskräfte waren auch Projektmitarbeitende und viele der eingesetzten Personen kannten sich persönlich, was eine Absprache und einen Abgleich vereinfachte und dazu

führte, dass die offenen Fragen und etwaige Unklarheiten schnell ihren Weg zu den zuständigen Projektmitarbeitenden fanden und dort geklärt werden konnten.

Die Interpretation der Ergebnisse erfolgt theoriegeleitet (Anlehnung an die Fähigkeit der Kinder, den Darstellungswechsel vorzunehmen). Innerhalb des vorliegenden Testinstruments wurden sämtliche Leistungen der Kinder einheitlich interpretiert.

Im Rahmen des Forschungsprojekts wurde also mit präzisen Vorgaben (Testanleitung, Codieranleitung), Testleitungsschulungen und optimierten Abläufen die Objektivität des Testablaufs sichergestellt.

Reliabilität

Ein Testinstrument wird dann als reliabel bezeichnet, wenn der Messfehler bei den Messungen durch das Testinstrument gering ist. Wissenschaftliche Instrumente zur Erfassung der Reliabilität gibt es verschiedene. Es kann z.B. die Retest-Reliabilität durch mehrmalige Testdurchführungen bestimmt werden. Weiter kann mit mehreren parallelen Testungen die Paralleltest-Reliabilität bestimmt werden. Testhalbierungs-Reliabilität und interne Konsistenz sind weitere Ansätze zur Erfassung der Reliabilität eines Instruments.

Aufgrund der Grösse des Forschungsprojekts und des Designs der Studie kamen verschiedene Verfahren nicht in Frage. Ein Retest war nicht denkbar aufgrund von Erinnerungseffekten bei den Kindern. Auch die Herstellung eines Paralleltests wäre aufgrund des Aufwands, den bereits die Herstellung einer einzelnen Testversion mit sich brachte, nicht möglich gewesen. Auch wäre es bei der vorliegenden Thematik äusserst schwierig, ausgehend von einem Testinstrument exakt gleich schwierige Items zu entwerfen.

Im Rahmen des vorliegenden Projekts wurde die Reliabilität mit Hilfe der internen Konsistenz (Cronbachs α) gemessen. Ergänzt wurde diese Grösse durch Reliabilitätsmasse aus der Item Response Theorie, so z.B. durch die WLE-Reliabilität aus der Theorie rund um das Rasch-Modell.

Diese Reliabilitätsmasse erwiesen sich sowohl beim Vor- als auch beim Nachtest als hoch. In den Erprobungen wurden hohe Werte erreicht und auch in der Erhebung vom September 2019 wurde für den Vortest eine interne Konsistent (Cronbachs α) von α=.89 erzielt. Die WLE-Reliabilität des Vortests (berechnet mit dem R-package TAM) betrug .85.

Die Erprobungen des Nachtests wiesen bereits einen Wert von α=.89 auf und bei der Hauptdurchführung des Nachtests wurde wiederum für Cronbachs α ein Wert von α=.89 gemessen. Die WLE-Reliabilität des Nachtests (berechnet mit dem R-package TAM) betrug .82.

Validität

Das wichtigste Testgütekriterium ist jenes der Validität. Die Validität stellt sich der Frage, ob ein Testinstrument auch wirklich dasjenige misst, was es zu messen behauptet. Die Validität ist aber häufig – und dies auch im hier vorliegenden Fall – wenn auch das wichtigste so auch

dasjenige Gütekriterium, das am schwierigsten zu belegen ist. Zur wissenschaftlichen Erfassung der Validität wird dieser Begriff unterteilt und näher beschrieben mit den Begrifflichkeiten der Inhaltsvalidität, Augenscheinvalidität, Konstruktvalidität und Kriteriumsvalidität. Innerhalb der Konstruktvalidität kann unterschieden werden zwischen den Facetten konvergenter und diskriminanter Validität (Moosbrugger & Kelava, 2012; Eid & Schmidt, 2014).

Die Inhaltsvalidität muss häufig theoriebasiert und von Experten geprüft und beurteilt werden. Die Augenscheinvalidität umfasst das Ausmass, in welchem die Validität auch von Laien auf einen ersten Augenschein hin für gegeben, d.h. die Messungen des Testinstruments für plausibel gehalten werden.

Die konvergente und die diskriminante Validität können quantitativ, z.B. mit Hilfe von Korrelationen der Testergebnisse mit anderen Tests erfasst werden. Diese anderen Tests können etwas ähnliches («konvergente Validität») oder etwas ganz anderes («diskriminante Validität») messen.

Im vorliegenden Fall wurde die Validität des Testinstruments dadurch geprüft, dass verschiedene Expertinnen und Experten sich viele der erprobten Testversionen immer wieder angesehen und kommentiert haben. Das Testinstrument wurde dabei als ein Test des Operationsverständnisses im Sinne der Literatur eingeschätzt.

Inhaltliche Aspekte, die dies hätten gefährden können, wurden eliminiert oder dort, wo dies nicht möglich war, so weit wie möglich reduziert. Beispielsweise wurde darauf geachtet, dass keine allzu hohen Zahlenwerte in den Testaufgaben vorkommen, damit die Bearbeitung durch die Kinder nicht am Zahlen- oder am Stellenwertverständnis scheitern würde. So sollte sichergestellt werden, dass es wirklich bei einer Messung des Operationsverständnisses bleibt. Auch wurden der sprachliche Teil und die verwendeten Sätze möglichst kurz und klar formuliert und das Vokabular möglichst vereinfacht, damit das Testergebnis nicht (oder zumindest nicht allzu sehr) vom Sprachverständnis der Kinder beeinflusst wird. Die Items wurden so stark wie möglich mit optischer Unterstützung ausgestattet, auch wenn diese zum Teil redundante (d.h. auch im Textteil vorkommende) Informationen enthielten. Die Aufgaben wurden den Kindern im ersten Teil des Tests zweimal vorgelesen, um zu verhindern, dass die auditiven Speicherfähigkeiten einen zu grossen Einfluss auf den Test haben könnten. Die Aufgaben im zweiten Teil des Tests waren so gestaltet, dass minimale Lesekompetenzen notwendig waren.

Immer wieder wurden über einen Verlauf von rund 18 Monaten Rückmeldungen von Expertinnen und Experten (Professorinnen, Professoren, Dozierenden und Lehrpersonen) aufgenommen und in den Testinstrumenten berücksichtigt.

Die Augenscheinvalidität war ebenfalls gegeben, denn die Hilfskräfte, welche als testleitende Personen zum Einsatz kamen, haben den Test nach kurzer Einführung zum Konstrukt «Operationsverständnis» für plausibel und vernünftig gehalten. Verschiedene Gespräche mit studentischen Hilfskräften und Mitarbeitenden des Projektteams führten zu diesem Fazit.

Eine einschränkende Klammerbemerkung zum Testinstrument soll noch einmal gemacht werden: Das vorliegende Testinstrument erfasst − aus den bereits dargelegten Gründen der

Testökonomie und der Machbarkeit – lediglich die Übersetzungsrichtung des «Mathematisierens» (Übersetzung «zum Term hin»). Das Konstrukt Operationsverständnis wird aber − wie bereits dargelegt − in der Literatur beschrieben als Fähigkeit zum Darstellungswechsel in beide Richtungen. Inwieweit die gemessene Richtung des Mathematisierens repräsentativ für das Gesamtkonstrukt ist, bleibt offen. Die Untersuchungen dazu sind gering an der Zahl und liefern zum Teil auch Widersprüchliches.

Zwecks Belegung der Validität des Erhebungsinstruments wurden Zusammenhangsmasse erfasst. Die Ergebnisse dieser Berechnungen sind in Kapitel 5.6 wiedergegeben.

Im Hinblick auf die faktorielle Validität wurden die Ergebnisse der Erhebungen auch einer exploratorischen Faktoranalyse unterzogen. Auf diese Art lässt es sich prüfen, ob bei der Erhebung ein eindimensionales oder ein mehrdimensionales Konstrukt erfasst wurde (Moosbrugger & Kelava, 2012, S. 162). Die Ergebnisse dieser Analyse sind in Kapitel 5.3 festgehalten.

4.4.4 Einschränkungen des Geltungsbereichs

Im vorhergehenden Kapitel wurde dargelegt, dass ein Erhebungsinstrument von ausreichender Testgüte entwickelt und eingesetzt wurde. Trotz diesen Ausführungen unterliegt das Instrument einigen Einschränkungen, welche im Folgenden ausgeführt werden und deren man sich bei einer allfälligen Replikation bewusst sein sollte. Die Diskussion des neuen Erhebungsinstruments für das Multiplikationsverständnis wird auch im Lichte der vorliegenden Ergebnisse in Kapitel 7.1 noch einmal aufgenommen.

Kinder in der zweiten Primarschulklasse

Beide Versionen (Vor- und Nachtest) des Erhebungsinstruments zur Erfassung des Multiplikationsverständnisses sind eigens für das Forschungsprojekt MALKA entwickelt worden. Von Beginn weg war klar, dass die Multiplikation untersucht werden sollte.

Eine Eigenheit dieses Projekts war, dass der Vortest vor Behandlung der Multiplikation im Unterricht eingesetzt werden sollte. Der Nachtest kam dann nach Behandlung der Multiplikation im Unterricht zum Einsatz und sollte eine ausreichende Menge an Brückenitems aufweisen.

Dieses Forschungsdesign hat sich auf die Ausgestaltung der Erhebungsinstrumente stark niedergeschlagen. Das vorliegende Erhebungsinstrument mit den beiden Versionen Vor- und Nachtest eignet sich also vor allem für eine Längsschnittuntersuchung von Kindern in der zweiten Primarschulklasse.

Für Kinder in späteren Schulklassen wären andere, anspruchsvollere Items notwendig. Auch für Querschnittuntersuchungen am Ende der zweiten Klasse könnten die Items anders gewählt werden, denn die Restriktion von Brückenitems mit einer früheren Version würde entfallen.

Wird das vorliegende Erhebungsinstrument für weitere Untersuchungen verwendet, so sollten diese Aspekte des Instruments bedacht und berücksichtigt werden.

Nur eine Übersetzungsrichtung

Aus Gründen der Testökonomie wurde durch das vorliegende Erhebungsinstrument nur die Übersetzungsrichtung «zum Term hin» geprüft. Dies stellt eine Vereinfachung des Konstrukts Operations- bzw. Multiplikationsverständnis dar. In dieser Arbeit wurde aus den dargelegten Gründen bewusst darauf verzichtet, die andere Übersetzungsrichtung «vom Term weg» (das «Interpretieren» bzw. das «Entmathematisieren») zu erheben.

Wollte man auch diese zweite Übersetzungsrichtung erfassen, so wären vermutlich Interviews mit einzelnen Kindern notwendig. In Studien mit einer tieferen Kinderanzahl könnte dies ins Auge gefasst werden. Zu beachten bleibt dann auch der zeitliche Aufwand, der mit einer solchen Studie einherginge.

Vereinfachungen bei der Codierung

Die finale Codierung der Items in diesem Erhebungsinstrument war eine dichotome Codierung. Auch wenn einige Sondercodes verwendet wurden, so wurden die Bearbeitungen der Kinder letztlich mit 0 oder mit 1 codiert. Mit einer solchen Codierung sind Vereinfachungen verbunden, die eine quantitative Auswertung erleichtern bzw. überhaupt erst ermöglichen.

Bewusst sein sollte man sich, dass eine solche Vereinfachung in einigen Fällen der Kinderleistung nicht ganz gerecht werden kann. Diese Einschränkung gilt auch dann, wenn sehr viel Arbeit und Energie in einen ausgeklügelten Codierleitfaden investiert wird. Eine der besonders schönen Erfahrungen in diesem Forschungsprojekt war es, zu erleben, wie man als Forscher immer wieder mit schlichten und guten Antworten von den Kindern überrascht – um nicht zu sagen «überlistet» − wurde.

Abbildung 24: Beispiel eines Items, das überraschende Kinderlösungen hervorbrachte.

Ohne Zweifel sind auch in unserem Projekt einige reichhaltige und gut begründete Kinderbeiträge unentdeckt geblieben. Stellvertretend für viele überraschende Antworten von unbeschwerten Kindern wird hier eine Antwort von zwei Kindern ausgeführt. Die beiden Kinder haben auf das Schokoladen-Item 5a aus dem Vortest auf die Frage «Wie viele Stücke Schokolade hat die Tafel? » mit der Zahl «1» geantwortet. Schliesslich hat die Schokolade ja keine Bruchstellen und Risse und besteht aus einem einzigen Stück. Bei der Antwort der Kinder handelt es sich um eine nachvollziehbare und korrekte Antwort, die aber nicht der ursprünglichen Absicht des Erhebungsinstruments entsprach. Auch wenn diese beiden Beantwortungen bei 2 von 1’084 Kindern bei der quantitativen Auswertung nicht ins Gewicht fällt, so illustriert sie doch eine Problematik, die bei der Ausgestaltung dieses Erhebungsinstruments immer wieder aufkam. Kinder können mit überraschenden, unbelasteten und durchaus korrekten Reaktionen die Forscher «überlisten». Der nachträgliche Verbesserungsvorschlag für dieses Item würde nun lauten «Wie viele *kleine* Stücke Schokolade hat die Tafel? ».

Einschränkung auf den deutschsprachigen Raum

Natürlich können die Items auch nicht ohne Weiteres auf andere Sprachräume übertragen werden. Bereits die Anweisungen in hochdeutscher Sprache mit Kindern aus der Schweiz, welche mit einem Teil des Vokabulars (z.B. Worte wie «Murmeln») nicht vertraut sind, erfordert Feingefühl. Die Vermeidung von Zungenbrechern und/oder von Mehrdeutigkeiten würde in einer anderen Sprache von Neuem beginnen.

Bemerkung zur Sensibilität des Erhebungsinstruments

Wie sich in Kapitel 6 zeigen wird, wies das neue Erhebungsinstrument im Gegensatz zum standardisierten Instrument «BASIS-MATH» eine geringere Sensibilität auf. Mit dem Instrument für das Multiplikationsverständnis war es nicht im selben Masse wie mit dem Instrument «BASIS-MATH» möglich, Unterschiede zwischen Stärkegruppen bei den untersuchten Kindern zu orten. Diese «Schwäche» des Erhebungsinstruments wird in der Diskussion in Kapitel 7.1 noch einmal aufgenommen. Zu erwähnen bleibt, dass sich dieses Phänomen bisher erst in dieser ersten Studie auf diese Art gezeigt hat.

Ein zweites kritisches Element des neuen Erhebungsinstruments ist der Deckeneffekt, der sich für die Brückenitems im Nachtest gezeigt hat. Dass sich dieser auch im Schuljahr 2019/20 gezeigt hat, lässt vermuten, dass dieser Deckeneffekt in einem regulären Schuljahr (ohne Schulschliessung) noch stärker auftreten würde. Für eine zukünftige Verwendung des Erhebungsinstruments müsste also diesen Brückenitems noch einmal Aufmerksamkeit geschenkt werden, allenfalls müsste – für den Einsatz in einer Längsschnittstudie − auf Brückenitems auch verzichtet werden.

4.4.5 Vor- und Nachtest «Multiplikationsverständnis»

Mit einer operationalen Definition eines Konstrukts wird diesem Konstrukt eine Bedeutung zugemessen, die eine Erfassung im Rahmen eines Erhebungsinstruments erlaubt (D. H. Rost, 2007, S. 55). Sowohl im Vor- als auch im Nachtest wurde das Operationsverständnis operationalisiert basierend auf der Theorie, die im theoretischen Teil dieser Arbeit in Kapitel 2 dargelegt wurde.

Es sollte erfasst werden, ob die Kinder die jeweiligen Übersetzungsleistungen zwischen unterschiedlichen Darstellungen erbringen können. Dabei sollten bewusst auch Aufgaben zu unterschiedlichen inhaltlichen Grundvorstellungen verwendet werden sowie Strukturierungsfähigkeiten bzw. Aufgaben mit Punktefeldern einfliessen. Die Kinder sollten ihre Fähigkeiten zu Übersetzungsleitungen und Darstellungswechseln an einer möglichst vielfältigen und breit abgestützten Aufgabensammlung unter Beweis stellen können. Es sollten die verschiedenen Aspekte des Konstrukts Operationsverständnis einfliessen, die in Kapitel 2.2 beschrieben und auch im Fördermaterial (siehe Kapitel 4.3.4) Eingang gefunden hatten.

Dabei galt es aber immer auch zu beachten, dass dieses Erhebungsinstrument mit der Aufforderung «Finde eine passende Rechnung. » einen neuartigen und zumindest im Vortest vielleicht auch überraschenden Aufgabentyp enthielt. Instruktionen an die Kinder durften nicht zu komplex werden und die einzelnen Items sollten so ausgestaltet sein, dass im Laufe der Testdurchführung nicht zu viele Formatwechsel auftreten würden. Zudem war es ein Anspruch an den Testablauf, dass die Kinder einen Teil des Hefts auch selbständig in ihrem eigenen Tempo bearbeiten konnten. Dies wurde den Kindern sowohl im Vor- als auch im Nachtest in der zweiten Testhälfte ermöglicht, erforderte aber auch eine entsprechende Ausgestaltung der Items. Die schriftliche Anweisung bei diesen Items musste beispielsweise kurz und klar sein, damit das Lösen der Items nicht an den Lesefähigkeiten der Kinder scheitern würde.

Auch sollte die Bearbeitung der Items nicht an den auditiven Speicherfähigkeiten der Kinder scheitern, was v.a. bei der zeitlich-sukzessiven Grundvorstellung zu Schwierigkeiten führte. Diese Schwierigkeit führte letztlich dazu, dass die verschiedenen inhaltlichen Grundvorstellungen bei der Multiplikation nicht sehr ausgewogen im Erhebungsinstrument abgebildet wurden. Das Item «Korb» war letztlich das einzige Item der Multiplikation, welches die zeitlich-sukzessive Grundvorstellung abdeckte. Bei den (wenigen) Items zur Addition und Subtraktion war es einfacher, verschiedene Grundvorstellungen zu berücksichtigen.

In diesem Kapitel werden nun die verwendeten Versionen des Vor- und des Nachtests beschrieben. Der Inhalt der Testhefte ist im Anhang 8.2 zu finden. Zu einer kompletten Dokumentation der Testversionen gehören aber auch die jeweiligen Anleitungen für die Testleitung (auf Anfrage an der PH FHNW verfügbar) sowie die jeweiligen Codieranleitungen, welche ebenfalls im Anhang 8.2 dieser Arbeit enthalten sind.

Vortest

Der Vortest war so ausgestaltet, dass er mit etwas zeitlicher Reserve von allen Kindern in einer Lektion bearbeitet werden konnte. Er wurde ausschliesslich mit den Kindern in den 2.

Klassen des Schuljahrs 2019/20 im Klassenverband (als «Gruppentest») durchgeführt, sofern die Kinderzahl der jeweiligen Klasse unter 20 lag. Grössere Klassen wurden für die Testdurchführung in zwei Halbklassen aufgeteilt.

Der Test begann mit genauen Anweisungen an die Kinder für den Testablauf. Diese Anweisung umfasste auch zwei einführende Beispiele von Items, die dazu dienten, die Kinder mit dem Itemformat vertraut zu machen. Diese Beispiele wurden nicht korrigiert und bewertet.

Nach den zwei einführenden Beispielen wurden die 10 Ausgangssituationen (für die ersten 17 Items) von der Testleitung zweimal langsam und deutlich vorgelesen. Die Kinder erhielten für jedes Item ausreichend Bearbeitungszeit. Diese Bearbeitungszeit lag bei etwa 30 Sekunden pro Item bzw. bei 45 Sekunden für Items, bei denen sowohl eine Zahl als auch eine Rechnung angegeben werden musste. Leitend war dabei der Auftrag «Finde eine passende Rechnung. » Sowie die Anweisung «Schreibe eine passende Rechnung ins Kästchen. Auf die Linie kommt eine Zahl. »

Die letzten 9 Ausgangssituationen (9 Items) wurden von den Kindern selbständig gelöst.

Weitere Testverarbeitung und Codierung

Der Vortest wurde von der Testleitung gemäss dem abgegebenen und in der Testleitungsschulung besprochenen Codierleitfaden korrigiert. Diese Korrektur wurde von einer Zweitperson gegengelesen und die finalen Werte dann elektronisch für die weitere Analyse erfasst. Dafür wurde eine FileMaker-Eingabemaske für dieses Instrument entworfen und eingesetzt. Sodann wurde die FileMaker-Datei in andere elektronische Formate übertragen und in Excel, SPSS und RStudio bearbeitet.

Sowohl im Vor- als auch im Nachtest sollte das Operationsverständnis der Kinder und nicht deren Rechenfähigkeiten geprüft werden. Erfasst wurde bei der Mehrzahl der Items die angemessene Übersetzungsleistung einer Geschichte bzw. eines Bilds in einen Term. Passende Terme wurden mit 1, unpassende Terme 0 codiert. Notwendig für eine Codierung mit 1 war nicht nur ein Term mit einem richtigen Ergebnis, sondern einer, der auch die Struktur der vorgegebenen Geschichte in wünschenswerter Weise erfasste.

Einzelne frühe Sondercodes für Verwechslung der 6 mit der 9 oder für das Vertauschen von Ziffern, wie z.B. 42 statt 24 wurden nach einer Expertendiskussion umcodiert in 0 und 1 (siehe Kapitel 5.1.1).

Bei den Fragen nach einer Zahl wurde die richtige Zahl mit 1, andere Zahlen mit 0 codiert.

Nachtest

Der Ablauf des Nachtests entsprach demjenigen des Vortests. Die Anweisungen an die Kinder waren dieselben. Der Test begann ebenfalls mit zwei einleitenden Beispielen. Der erste Teil des Tests wurde zusammen mit den Kindern bearbeitet, d.h. die testleitende Person hat die Items zweimal langsam vorgelesen und den Kindern bei jedem Item jeweils ausreichend Zeit

für die Bearbeitung zugestanden. Der letzte Teil des Tests wurde von den Kindern eigenständig bearbeitet.

Ein neu eingeführtes Format waren die beiden Multiple-Choice-Aufgaben auf der Rückseite des Testhefts (Items 25 und 26, siehe Anhang 0).

Das Instrument wurde wiederum mit 0 und 1 codiert und auf dieselbe Art korrigiert, gegenkorrigiert und in eine elektronische Version umgewandelt wie der Vortest.

Sondercodes und deren Umwandlung werden in Kapitel 5.1.1 beschrieben.

Brückenitems

Teil der beiden beschriebenen Versionen (Vor- und Nachtest) sind die folgenden 12 Brückenitems, welche für die Messungen des Lernzuwachses im Längsschnitt (siehe Kapitel 6) verwendet werden. Bei den Brückenitems («Ankeritems») handelt es sich um Items, die in beiden Testversionen in identischer Form vorkommen.

Tabelle 7: Liste der Brückenitems

Brückenitem	Item-Nr. im Vortest	Item-Nr. im Nachtest
Stifte	2	7
Autos (Zahl)	4a	1a
Autos (Rechnung)	4b	1b
Bücher (Zahl)	6a	3a
Bücher (Rechnung)	6b	3b
Stühle (Zahl)	8a	8a
Stühle (Rechnung)	8b	8b
Korb (Zahl)	10a	5a
Korb (Rechnung)	10b	5b
Füsse	13	13
Beeren (3 Teller)	14	11
Beeren (7 Teller)	15	12

4.5 Auswertungsmethoden

Die dieser Arbeit zugrunde liegenden testtheoretischen Modelle entstammen sowohl der klassischen Testtheorie als auch der Item Response Theorie. Aus der Item Response Theorie wurde insbesondere das eindimensionale Rasch-Modell verwendet. Eine Einführung in die Testtheorie und die jeweiligen Modelle findet sich in verschiedenen Lehrbüchern (J. Rost, 2004; Kubinger et al., 2011; Moosbrugger & Kelava, 2012; Eid & Schmidt, 2014; Strobl, 2015; Bühner & Ziegler, 2017; Eid et al., 2017). Im folgenden Kapitel 4.5.1 erfolgt daher eine Beschreibung der klassischen Testtheorie in kürzester Form.

Etwas eingehender wird das Rasch-Modell aus der Item Response Theorie in Kapitel 4.5.2 beschrieben. Für die Einordnung des Rasch-Modells erfolgt auch ein Vergleich mit dem Birnbaum-Modell, das für diese Arbeit aber nicht verwendet wurde.

Nebst den üblichen Selektionskriterien aus der klassischen Testtheorie wurden die Items des Erhebungsinstruments für das Multiplikationsverständnis vor allem auch im Hinblick auf unterschiedliche Wirkungsweise («Differential Item Functioning») untersucht (Koller et al., 2012; Strobl, 2015; Trendtel, Schwabe, et al., 2016).

Für die Auswertung im Längsschnitt wurden auch Mehrebenenmodelle bzw. Mehrebenenanalysen herangezogen. Einführung und Beschreibung zu dieser Thematik findet sich u.a. in Ditton (1998), Engel (1998), Raudenbush & Bryk (2002), Langer (2009), Snijders & Bosker (2012) und Finch et al. (2019). Diese Art der Analyse wird in Kapitel 4.5.5 beschrieben. Die Ergebnisse und die unter Verwendung dieser Theorien erzielten Auswertungen werden dann in Kapitel 6 ausgeführt.

Die Auswertungen erfolgten mehrheitlich mit der Software R (Luhmann, 2015; Breit & Schreiner, 2016; Baker, 2017) und verschiedenen Packages für diese Software, so z.B. (Mair et al., 2020; Revelle, 2020; Robitzsch et al., 2020). Ergänzend wurde an einigen Stellen auch mit SPSS (Brosius, 2013; Field, 2013) und mit Microsoft Excel gearbeitet.

4.5.1 Klassische Testtheorie

Bei der klassischen Testtheorie (KTT) handelt es sich um eine Theorie basierend auf einem Modell für die gemachten Beobachtungen. In diesem Modell wird davon ausgegangen, dass die für eine Person v beim Item i gemachten Beobachtungen sich aus zwei Summanden zusammensetzen. Die Summanden sind der Erwartungswert der Person v beim Lösen des Items i sowie der Messfehler dieser Beobachtung. Die klassische Testtheorie wird deshalb auch eine Messfehler-Theorie genannt (Moosbrugger & Kelava, 2012, S. 104 ff.). Für die Beobachtung x_{vi} des Ergebnisses der Person v am Item i gilt also gemäss dieser Theorie:

$$x_{vi} = \tau_{vi} + \varepsilon_{vi}$$

wobei τ_{vi} gleich dem Erwartungswert des Ergebnisses der Person v für das Item i ist und ε_{vi} für den Messfehler dieser Person für dasselbe Item i steht.

Die Axiome der klassischen Testtheorie besagen, dass für jede Person und jedes Item der Erwartungswert x_{vi} existiert (Existenzaxiom) und dass sich der Erwartungswert und der Fehlerterm in der obigen additiven Form kombinieren lassen (Verknüpfungsaxiom). Aus diesen beiden Axiomen kann gefolgert werden, dass der Fehlerterm den Erwartungswert 0 hat. Zudem wird bei diesem Modell davon ausgegangen, dass die beiden Grössen τ_{vi} und ε_{vi} nicht korreliert sind (Unabhängigkeitsaxiom).

Auf diesen Grundannahmen wurde die klassische Testtheorie entwickelt. Sie fand lange Zeit und nach wie vor weite Verbreitung und ist in vielen Anwendungsfeldern präsent. Die Theorie wird darum «klassisch» genannt, weil sie schon seit langer Zeit in Gebrauch ist (Pospeschill, 2010; Moosbrugger & Kelava, 2012).

Im Rahmen der klassischen Testtheorie werden die Grössen Item-Schwierigkeit, Itemvarianz und Itemtrennschärfe sowie auch das Reliabilitätsmass Cronbachs α berechnet. Die Item-Schwierigkeit ergibt sich bei dichotom codierten Items − wie sie für diese Arbeit vorliegen – im Wesentlichen aus dem Verhältnis der mit 1 codierten Antworten zur Summe aller mit 0 oder 1 codierten Antworten. Je höher die Item-Schwierigkeit ist, desto leichter ist das Item. Die Itemtrennschärfe für ein Item gibt die Korrelation des Items mit dem restlichen Test an. Dieser Wert sollte nicht zu hoch und nicht zu klein sein, sonst würde das Item entweder keine neue zusätzliche Information liefern oder aber etwas anderes messen als der restliche Test. Cronbachs α misst die interne Konsistenz des Testinstruments. Dieser Wert sollte möglichst hoch, d.h. möglichst nahe bei 1 zu liegen kommen.

Für die Analyse des neuen Erhebungsinstruments kommt eine exploratorische Faktoranalyse mit Hauptkomponentenanalyse zur Anwendung (Eid et al., 2017, S. 920 ff., 932 ff.; Schweinberger, 2018).

Die Grössen der KTT lassen sich mit der Software SPSS der Firma IBM oder auch mit der Open-Source-Software R bestimmen. Auswertungen und weitere Einzelheiten hierzu folgen im Ergebnisteil in Kapitel 6.

4.5.2 Item Response Theorie (IRT)

In jüngerer Zeit wurde die klassische Testtheorie in den Sozialwissenschaften immer häufiger ergänzt durch unterschiedliche Modelle der Item Response Theorie (IRT). Diese Theorie wird auch Probabilistische Testtheorie (PTT) genannt, weil sie von verschiedenen Wahrscheinlichkeitsverteilungen für die Lösung bzw. Beantwortung der einzelnen Items ausgeht. Die Item Response Theorie ergänzt die bewährte klassische Testtheorie, soll diese aber nicht gänzlich ersetzen oder ablösen. Einführungen zu dieser Theorie finden sich in vielen Lehrmitteln, u.a. auch in Pospeschill (2010, S. 114 ff.) und in Moosbrugger & Kelava (2012, S. 227 ff.). Im Gegensatz zur klassischen Testtheorie geht die IRT von latenten Variablen aus, die mittels Beobachtungen geschätzt werden können. Zentral sind Annahmen über die Wahrscheinlichkeitsverteilungen für die Lösung bzw. Beantwortung der Items durch die verschiedenen Personen. Modelle der IRT gibt es verschiedene. Die Anzahl und die Bedeutung der verwendeten

latenten Variablen hängen vom jeweiligen Modell ab. Sie können von Modell zu Modell sehr unterschiedlich sein.

Als ein weit verbreitetes und beliebtes Modell aus der IRT hat sich das dichotome Rasch-Modell etabliert (Pospeschill, 2010, S. 122 ff.; Moosbrugger & Kelava, 2012, S. 236 ff.; Eid & Schmidt, 2014, S. 144 ff.; Strobl, 2015). Dieses wird auch 1PL oder Einparameter-Logistisches Modell genannt und weist für die Auswertungen aufgrund seiner Einfachheit besonders angenehme Eigenschaften auf. Im Rahmen dieser Arbeit wird aus der IRT ausschliesslich dieses Modell verwendet, weshalb dieses im Folgenden etwas eingehender beleuchtet werden soll.

Das dichotome Rasch Modell

Das dichotome Rasch-Modell geht von den beiden latenten Variablen «Personenfähigkeit» und «Item-Schwierigkeit» aus. Diese beiden Grössen werden aufgrund von gemachten Beobachtungen für die verschiedenen Personen und für alle Items geschätzt. Die Personenfähigkeit ist eine Eigenschaft der Person und hängt nicht vom jeweiligen Item ab. Sie beeinflusst aber das Ergebnis der Person bei der Bearbeitung der einzelnen Items.

Typischerweise kommen in der Praxis für die Schätzung der Personenfähigkeit und der Item-Schwierigkeit Schätzverfahren aus der Gruppe der Maximum-Likelihood-Methoden zur Anwendung (Long, 2012, S. 191 ff.; Trendtel, Pham, et al., 2016, S. 191 ff.). Bei diesen Methoden wird unter den möglichen Werten für die Personen- und die Item-Schwierigkeit diejenige Kombination gesucht, unter welchen – bei gegebenen gemachten Beobachtungen – eben diesen Beobachtungen die grösstmögliche Wahrscheinlichkeit zukommt. Es kommen hierfür Computerprogramme mit iterativen Suchalgorithmen unterschiedlicher Art zur Anwendung. Die Ergebnisse der verschiedenen Programme können sich dabei unterscheiden. Grund dafür kann z.B. sein, dass die gesuchte Parameterkombination nicht eindeutig ist oder dass der Algorithmus aus strukturellen Gründen bei den gemachten Beobachtungen nicht optimal konvergiert.

In dieser Arbeit werden für die Suchprozesse die frei erhältlichen und vielfach verwendeten R-packages (z.B. eRm, TAM, difR) verwendet, in welchen unterschiedliche Algorithmen implementiert sind.

In dieser Arbeit kommt als Schätzmethode auch eine verbesserte Methode mit zusätzlichen wünschenswerten Eigenschaften zur Anwendung, bei der das Schätzverfahren eine andere Likelihood-Funktion maximiert und einen gewichteten («weighted») Schätzer entwickelt (Warm, 1989; Trendtel, Pham, et al., 2016). Die auf diese Art geschätzten Werte werden mit WLE für «weighted likelihood estimates» abgekürzt.

Denkbar sind auch andere Verfahren der Parameterschätzung, so z.B. die «Expected A Posteriori»-Schätzmethode als auch «Plausible Values», für deren Vor- und Nachteile auf die Literatur verwiesen sei (Hartig & Kühnbach, 2006; Trendtel, Pham, et al., 2016, S. 203 f.).

Beim Rasch-Modell wird die Personenfähigkeit als nicht auf Anhieb oder ohne entsprechende Erhebung ersichtliche (latente), aber stabile Eigenschaft einer Person angesehen. Sie bildet also einen «Trait», im Gegensatz zu einem «State», der sich situationsabhängig oder zeitlich instabil äussert.

Das Rasch-Modell ist eines der einfachsten und daher auch angenehmsten Modelle aus der Gruppe der Latent-Trait-Modelle. Angenehm ist es insbesondere darum, weil es verhältnismässig wenige Variablen und somit Parameter aufweist.

Nimmt man an, dass jede Person v eine Personenfähigkeit p_v aufweist und jedes Item mit einer Item-Schwierigkeit β_i beschrieben werden kann, dann hat unter der Annahme der Gültigkeit des Rasch-Modells die Person v für das Item i eine Lösungswahrscheinlichkeit von

$$\mathrm{P}(x_{vi}) = \frac{e^{x_{vi} \cdot (p_v - \beta \mathrm{i})}}{1 + e^{(p_v - \beta \mathrm{i})}}$$

Damit ergeben sich für jede Person und jedes Item Lösungswahrscheinlichkeitskurven (sogenannte ICC «Item Characteristic Curves»), deren Form in Abbildung 25 dargestellt ist. Diese Kurven haben alle dieselbe Form, liegen aber abhängig von der Item-Schwierigkeit an unterschiedlichen Stellen.

Auf der vertikalen Achse der graphischen Darstellung kann für jede ICC die Lösungswahrscheinlichkeit einer Person mit entsprechender Personenfähigkeit abgelesen werden.

Die Kurven von schwierigeren Items liegen weiter rechts. Die rote Kurve in der Abbildung 25 beispielsweise gehört zu einem Item mit Item-Schwierigkeit $\beta_3 = 3$. Eine Person mit einer Fähigkeit, welche der Item-Schwierigkeit entspricht (abgelesen auf derselben horizontalen Achse) hat jeweils eine Lösungswahrscheinlichkeit für dieses Item von genau 50%. Fähigere Personen sind rechts davon abzulesen und werden mit Lösungswahrscheinlichkeiten über 50% modelliert.

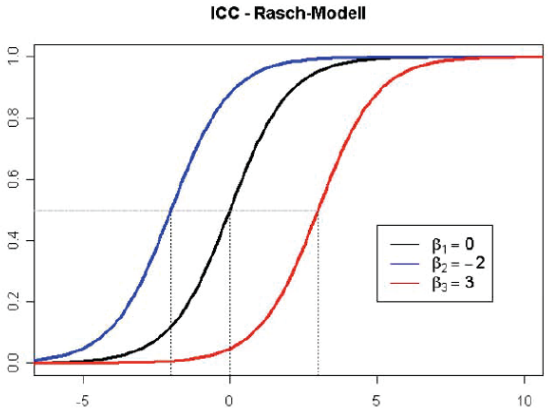


Abbildung 25: Drei ICCs aus dem Rasch-Modell für verschiedene Item-Schwierigkeiten aus Lechner & Niehaus (2010)

Eine der Grundannahmen des Rasch-Modells ist die Kongruenz der einzelnen ICC, welche auch Rasch-Homogenität genannt wird (Eid & Schmidt, 2014, S. 146).

Mittels eines grossen Datensatzes können nun die Fähigkeitsparameter für die einzelnen Personen, vor allem aber auch die einzelnen Item-Schwierigkeiten geschätzt werden. Bei der Testentwicklung und auch bei der Testauswertung kann geprüft werden, ob einzelne Items ins Rasch-Modell passen, indem rechnerisch eruiert wird, ob die ICCs für diese Items die obige Form annehmen.

Einzelne geschätzte Kennwerte, wie z.B. die Mean-Square-Infit-Werte («MSQ-Infit-Werte») der einzelnen Items zeigen auf, wie sehr die Daten zum Rasch-Modell passen, d.h. inwiefern die einzelnen ICCs zumindest annäherungsweise parallel verlaufen. Dieser Wert, der als Mass für die Steigung der geschätzten Kurve bei ihrem Wendepunkt verstanden werden kann, sollte für die einzelnen Items zwischen 0.7 und 1.3, wenn möglich aber zwischen 0.8 und 1.2 liegen, auch wenn in der Literatur zum Teil auch grosszügigere Bandbreiten von 0.5 bis 1.5 genannt werden (Ames & Penfield, 2015; Trendtel, Pham, et al., 2016). Ein Wert von 1.0 bedeutet, dass die Kurve für dieses Item die ideale Modellform hat und optimal ins Rasch-Modell passt. Bei höheren Werten ist die Steigung der Kurve zu tief, bei tieferen Werten verläuft die für dieses Item geschätzte Kurve zu steil. Bei tiefen Werten verhält sich das Item also «trennschärfer» als es diejenigen Items tun, die optimal ins Rasch-Modell passen. Ein Item mit tiefem Wert passt also gewissermassen «zu gut» ins Modell. Hohe MSQ-Infit-Werte (d.h. flache ICC-Kurven) deuten darauf hin, dass die ICC-Kurve im Wendepunkt zu wenig steil verläuft und das Item somit zu wenig stark zwischen fähigen und weniger fähigen Personen trennt. Hohe Werte sind daher als etwas problematischer zu betrachten als tiefe Werte (Grob, 2018).

Im Verlauf der Testentwicklung in dieser Arbeit wurde diese Art der Rasch-Homogenität als eines der Selektionskriterien für die erprobten Items verwendet. So wurde eine bessere Passung der für das Forschungsprojekt beibehaltenen Items ins Rasch-Modell erreicht.

Die Personenfähigkeitsparameter des Rasch-Modells wurden in Kapitel 6 für die vorgenommenen Längsschnittanalysen herangezogen. Dieses Vorgehen hat Vorteile, auch wenn der Testscore aus einer Erhebung und die jeweiligen Personenfähigkeiten im Rasch-Modell in einem monotonen und (beinahe) linearen Zusammenhang zueinander stehen und die Unterschiede deshalb auf den ersten Blick gering erscheinen mögen. In Abbildung 26 wird der Zusammenhang zwischen Testscore im Vortest für das Operationsverständnis (vor Entfernung der Items 5a und 5b, dargestellt auf der horizontalen Achse) mit der im Rasch-Modell berechneten Personenfähigkeit β (auf der vertikalen Achse) optisch dargestellt. Der lineare Zusammenhang der beiden Grössen ist mit einer Korrelation von .98 sehr hoch.

Die Wahl dieser Parameter anstelle des Testscores aus den jeweiligen Erhebungen hat den Vorteil, dass Kinder im sehr hohen und im sehr tiefen Leistungsbereich modellmässig besser abgebildet werden (Zhao et al., 2017, S. 606). In Abbildung 26 ist dies am linken und rechten Ende in den Bereichen mit etwas höherer Krümmung des Funktionsgraphen zu erkennen.

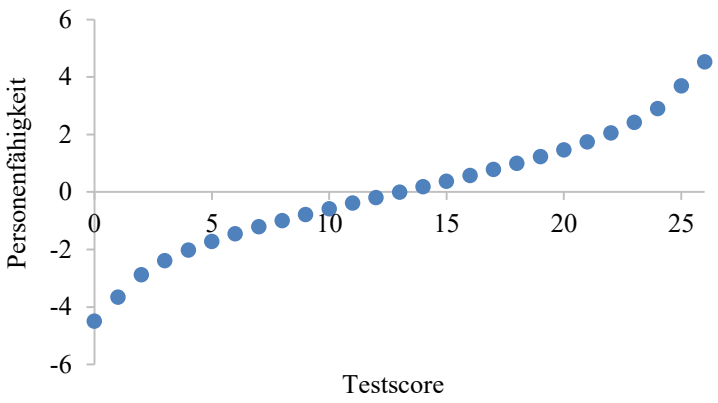


Abbildung 26: Personenfähigkeit und Testscore; gerechnet mit den Vortest-Daten

Für die hier vorliegende Arbeit wäre auch die Verwendung des Testscores vertretbar, vor allem auch darum, weil lediglich zwei Messzeitpunkte für die Auswertung ausgewertet werden. Die Verwendung der Personenfähigkeit aus dem Rasch-Modell anstelle des Testscores bringt für diese Arbeit aber keinerlei Nachteile mit sich (Hartig & Frey, 2013; Zhao et al., 2017).

Das dichotome Birnbaum-Modell

Eine Alternative zur Verwendung des Rasch-Modells wäre die Wahl eines komplexeren Modells, z.B. eines Modells mit mehreren freien Parametern. Ein denkbares Modell mit einem zusätzlichen Parameter ist das Birnbaum-Modell (Pospeschill, 2010, S. 132 f.; Eid & Schmidt, 2014, S. 204 ff.), welches einen Parameter mehr umfasst und darum auch «2PL» (Zweiparameter-Logistisches Modell) genannt wird. Der zusätzliche Parameter ermöglicht es, ICCs mit unterschiedlicher Steilheit im Modell zu erfassen. In Abbildung 27 sind beispielhaft drei ICCs aus diesem Modell abgebildet.

Bei der Verwendung dieses Modells müssten Items nicht mehr aufgrund ihrer unerwünschten Steigungen eliminiert werden, denn das Modell erlaubt es, die Steigung mit einem passenden Modellparameter abzubilden.

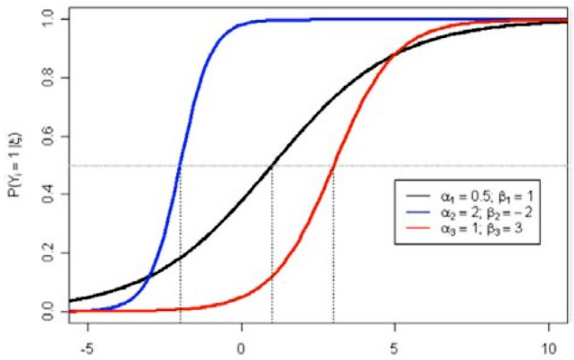


Abbildung 27: Drei ICCs aus dem Birnbaum-Modell für verschiedene Item-Schwierigkeiten aus Lechner & Niehaus (2010)

Die Nachteile, die mit der Verwendung des Birnbaum-Modells in Kauf genommen werden müssten, sind die höhere Komplexität dieses Modells und die weniger angenehmen Modelleigenschaften. Insbesondere ginge mit der Verwendung des Birnbaum-Modells die spezifische Objektivität des Modells verloren. Diese Eigenschaft des Rasch-Modells umschreibt u.a. die Tatsache, dass der Unterschied in der Item-Schwierigkeit von zwei Items nicht davon abhängt, welche Personengruppe die Items bearbeitet hat. Diese angenehme Modelleigenschaft ist beim Birnbaum-Modell nicht erfüllt, was sich der Darstellung in Abbildung 27 leicht entnehmen lässt. Das in schwarzer Farbe dargestellte Item 1 (mit den Parametern $\alpha_1 = 0.5$ und $\beta_1 = 1$) ist für Personen mit tieferer Fähigkeit leichter zu lösen als das in roter Farbe dargestellte Item (mit den Parametern $\alpha_3 = 1$ und $\beta_3 = 3$). Bei Personen mit sehr hoher Fähigkeit verhält es sich umgekehrt.

Die Verwendung des Birnbaum-Modells für diese Arbeit wäre möglich gewesen, hätte aber zu Komplikationen und zusätzlichen Parametern geführt. In dieser Arbeit wurde daher mit dem Rasch-Modell gearbeitet. Vergleiche der Modellgüte für die beiden Modelle bestätigten, dass diese Entscheidung angemessen war.

Erwähnung finden soll das Birnbaum-Modell aber deshalb, weil es aufzeigt, dass die Auswahl der Items mit Hinblick darauf, ob sie ins Rasch-Modell passen oder nicht, kein eigentliches Gütekriterium für das Item selbst darstellt. Diese Art der Auswahl von Items erleichtert lediglich die spätere Auswertung mit einem Modell, das angenehme Modelleigenschaften aufweist.

4.5.3 Überprüfung von Differential Item Functioning

Ein Bestandteil der Theorie zum Rasch-Modell sind auch Werkzeuge zur Überprüfung des Modells und der Modellwahl. Mit solchen Werkzeugen wird der Frage nachgegangen, ob und wie gut die tatsächlich gesammelten Beobachtungen und Messwerte zum gewählten Modell passen.

Die Überprüfung der Fit-Werte der einzelnen Items im Hinblick auf die Form ihrer ICC im Rasch-Modell – so wie im vorhergehenden Kapitel beschrieben – stellt ein solches Vorgehen dar.

Analysen des «Differential Item Functioning» (DIF), mit denen überprüft wird, ob die Items in derselben Art und Weise wirken, sind ein weiteres derartiges Werkzeug. Dabei kommen unterschiedliche Methoden zum Einsatz, mit denen jeweils auf verschiedene Art und Weise geprüft wird, ob einzelne Items sich für einen Teil des beobachteten Bestands anders verhalten als für einen andern. Unterschieden werden kann dabei zwischen «uniform DIF» und «non-uniform DIF» sowie DIF, welches beide Facetten «uniform DIF» und «nonuniform DIF» aufweist. Bei Items mit «uniform DIF» unterscheiden sich die ICC der beiden Teilbestände lediglich durch eine unterschiedliche Item-Schwierigkeit. Bei Items, welche «nonuniform DIF» aufweisen, unterscheiden sich die ICC der Teilbestände auch durch ihre Form, z.B. durch die maximale Steigung, welche die Graphen der ICC in ihrem Wendepunkt aufweisen (Trendtel, Schwabe, et al., 2016). Einige der Methoden zur Bestimmung von DIF eignen sich besonders zur Ortung von «uniform DIF», andere eher für «nonuniform DIF».

Mit solchen Methoden zur Bestimmung von DIF kann beispielsweise untersucht werden, ob es Unterschiede im Lösungsverhalten zwischen den Geschlechtern gibt. Wird die Item-Schwierigkeit β eines Items deutlich anders geschätzt, wenn nur die männlichen Personen für die Schätzung von β verwendet werden, so spricht man von DIF für dieses Item für das «Split-Kriterium» Geschlecht. Dieses Vorgehen kann für beliebige andere Eigenschaften der Personengruppe (z.B. Alter, Sprachverständnis, sprachliche Ausdrucksfähigkeit usw.) angewendet werden.

Derartige Analysen werden in der vorliegenden Arbeit mit unterschiedlichen Methoden insbesondere mit Blick auf die Sprachkompetenzen der Kinder vorgenommen (siehe Kapitel 5.4.2 und 5.5.2).

Unterschieden werden muss innerhalb der Methoden zur Überprüfung der Modellgültigkeit zwischen

a) Methoden, welche die globale Modellgültigkeit prüfen
b) Methoden, welche die Gleichheit der geschätzten Itemparameter für einzelne Subpopulationen prüfen (eigentliche DIF-Prüfung wie oben beschrieben) und
c) Methoden, welche die Gleichheit der geschätzten Personenparameter bei Aufteilung der Items in Subgruppen prüfen
d) Weiteren Methoden

Mit dem Anderson Likelihood-Ratio-Test wird im Folgenden eine Methode des Typs a) beschrieben. Es folgt mit dem Wald-Test, der graphischen Modellüberprüfung und dem Mantel-Haenszel-Verfahren für die DIF-Analyse eine kurze Beschreibung von drei Methoden des Typs b). Mit dem Martin-Loef-Test wird auch eine Methode des Typs c) dargelegt. Weitere Möglichkeiten der Modellprüfung finden sich in (Eid & Schmidt, 2014, S. 182 ff.; Trendtel, Schwabe, et al., 2016; Magis et al., 2018).

Andersen Likelihood-Ratio-Test

Der Likelihood-Ratio-Test von Andersen (1973) stellt einen globalen Model-Fit-Test dar. Er untersucht, ob sich die Items als Ganzes im Hinblick auf ein Trennungskriterium des Personenbestands (wie z.B. männlich/weiblich) in den beiden Subgruppen gleich verhalten. Bei diesem Test werden zwei verschiedene Likelihood-Funktionen, welche zur Schätzung von Personenfähigkeit und Item-Schwierigkeit verwendet werden können, zueinander ins Verhältnis gesetzt. Eine Likelihoodfunktion ist dabei diejenige des Rasch-Modells, als zweite Likelihoodfunktion wird eine Funktion verwendet, welche keinerlei Restriktionen bezüglich Modellwahl auferlegt werden. Das Verhältnis der beiden Funktionen folgt dabei – nach einer arithmetischen Transformation – einer χ^{2-}Verteilung (Pospeschill, 2010, S. 140 ff.; Long, 2012, S. 291; Eid & Schmidt, 2014, S. 185 ff., 195 ff.).

Verwirft der Test, so wird dies so interpretiert, dass die Daten nicht optimal zum Modell passen bzw. dass das Antwortverhalten in den beiden Subgruppen unterschiedlich ist. Der Likelihood-Ratio-Test macht keine Angabe dazu, an welchen Items es liegt, falls der Test verwirft.

Wald-Test

Mit Hilfe des Wald-Tests kann der Fit einzelner Items in den Gesamttest untersucht werden. Dieser Test erlaubt es also, ein andersartiges Verhalten einzelner Items, also tatsächliches «Differential Item Functioning» zu orten.

Bei diesem Test wird der Personenbestand in zwei Gruppen (z.B. in männlich/weiblich) unterteilt. Für jedes einzelne Item werden die Item-Schwierigkeitsparameter im Rasch-Modell für die einzelnen Personengruppen einzeln geschätzt. Sodann können für jedes Item die Schätzwerte der beiden Personengruppen verglichen werden. Sind die beiden Werte für ein Item ausreichend unterschiedlich, dann weist dieses Item für das gewählte Split-Kriterium «Differential Item Functioning» (DIF) auf. Zur Modellprüfung werden dafür die beiden geschätzten Item-Schwierigkeiten zu einer Teststatistik verrechnet, die im Wesentlichen mit der Differenz der beiden geschätzten Item-Schwierigkeiten arbeitet und einer Standardnormalverteilung folgt (Koller et al., 2012; Eid & Schmidt, 2014, S. 187 ff.; Strobl, 2015).

Mit diesem Test wird aufgrund des gewählten Vorgehens vor allem «uniform DIF» entdeckt.

Graphische Überprüfung des Item-Fits

Eine weitere Möglichkeit, DIF zu orten, besteht darin, die Item-Schwierigkeiten der beiden Personengruppen graphisch darzustellen. Dabei wird die Item-Schwierigkeit der einen Personengruppe auf der x-Achse, diejenige der anderen Personengruppe (für dasselbe Item) auf der y-Achse dargestellt. Für jedes Item wird auf diese Weise ein einzelner Punkt dargestellt. Bei optimalem Modell-Fit würden alle so dargestellten Punkte auf der Geraden mit der Gleichung y=x zu liegen kommen. Die Abweichung der einzelnen Punkte von dieser Geraden illustriert das Mass an DIF für das jeweilige Item.

Typischerweise werden die Punkte mit 95%-Konfidenzellipsen versehen, die illustrieren, ob die Abweichung von der Geraden y=x durch Zufall erklärt oder aber (auf dem 5%-Niveau) mit nur sehr geringer Wahrscheinlichkeit durch Zufall erklärt werden kann. Schneidet also die Konfidenzellipse die Gerade y=x nicht, so geht man von DIF für das jeweilige Item aus (Eid & Schmidt, 2014, S. 183 ff.).

Entsprechende graphische Illustrationen zu den Ergebnissen in den Kapiteln 5.4.2 und 5.5.2 finden sich im Anhang 8.4.4.

Mantel-Haenszel Verfahren

Ein bereits etwas älteres Verfahren zur Ortung von DIF ist das Mantel-Haenszel-Verfahren (Mantel & Haenszel, 1959; Zwick et al., 1999; Paek & Wilson, 2011). Mit diesem Verfahren kann nicht nur erhoben werden, ob DIF vorliegt, sondern auch eine Aussage über das Ausmass des DIF-Effekts gemacht werden. Der DIF-Effekt wird dabei in vorhanden, aber gering bzw. vernachlässigbar («Negligible», Typ A), leicht bis mittel («Slight to Moderate», Typ B) oder erheblich («Moderate to Large», Typ C) eingeteilt.

Verwendet wird dafür die Mantel-Haenszel Teststatistik, welcher Chancenverhältnisse («odd ratios») zugrunde liegen. Die Chance («odds») einer Person, ein Item richtig zu lösen, ist dabei definiert als Verhältnis zwischen ihrer Wahrscheinlichkeit p und ihrer Gegenwahrscheinlichkeit 1-p, das Item korrekt zu lösen.

Chance bei Lösungswahrscheinlichkeit p: $\frac{p}{1-p}$

Beim Chancenverhältnis werden die Chancen von Personen der einen Personengruppe ins Verhältnis gesetzt zur Chance der Personen der anderen Personengruppe. Das -2.35fache des natürlichen Logarithmus dieses Werts wird als Teststatistik MH-DIF verwendet. Abhängig von der Höhe dieser Teststatistik und der Signifikanz eines Werts grösser als 1 bzw. grösser als 0 erfolgt die Einteilung in die obigen Kategorien für das Ausmass des DIFs (Zwick et al., 1999, S. 3 f.).

Umgesetzt wird dieser Test mit Hilfe des R-packages difR. Die Ergebnisse der Analysen sind in den Kapiteln 5.4.2, 5.5.2 sowie 6.3.1 beschrieben.

Martin Loef-Test

In den bisher beschriebenen Testverfahren wurde zuerst der Personenbestand in zwei Gruppen eingeteilt und dann die Parameter für die Item-Schwierigkeit getrennt geschätzt. Die Absicht dabei war es, zu prüfen, ob dieselben oder ähnliche Item-Schwierigkeitswerte erhalten werden. Analog zu diesem Vorgehen ist es auch möglich, die Items in Gruppen einzuteilen und sodann zu prüfen, ob die Paramater für die Personenfähigkeit in den beiden Item-Subgruppen gleich oder zumindest nicht allzu stark unterschiedlich geschätzt werden.

Damit kann beispielsweise geprüft werden, ob die früheren Items im Test besser gelöst werden als die späteren (Stichwort: Ermüdungseffekt) oder ob Items mit höheren Anforderungen an die Lesekompetenzen anders gelöst werden als andere.

Der Martin-Loef-Test leistet genau dieses und prüft auf diese Art die Modellgültigkeit des Rasch-Modells auf globaler Stufe (Koller et al., 2012; Eid & Schmidt, 2014, S. 197 f.). Ähnlich wie beim Likelihood-Ratio-Test von Andersen werden dabei Verhältnisse verschiedener Likelihood-Funktionen als Teststatistik verwendet.

Signifikanz des Tests bedeutet, dass die vorgängig vorgenommene Einteilung in die beiden Item-Subgruppen bedeutenden Einfluss hat, erklärt aber nicht die Gründe für diesen Einfluss.

Verwendung der Ergebnisse der DIF-Untersuchungen in dieser Arbeit

In dieser Arbeit kommen sämtliche oben beschriebenen Verfahren zum Einsatz. In der Literatur werden diese Verfahren im Hinblick auf ihre Vor- und Nachteile eingehend diskutiert (Koller et al., 2012; Eid & Schmidt, 2014; Strobl, 2015; Trendtel, Schwabe, et al., 2016; Magis et al., 2018).

Es wurden sowohl der Vor- als auch der Nachtest mit den verschiedenen Verfahren auf DIF hinsichtlich der Kriterien Geschlecht und Sprachverständnis der Kinder untersucht. Die Längsschnittuntersuchung für das Multiplikationsverständnis mit Hilfe von Brückenitems erlaubte weitere DIF-Untersuchungen, welche unterschiedliches Lösungsverhalten «über die Zeit» orteten.

Die Ergebnisse dieser DIF-Untersuchungen werden detailliert berichtet. Ob diese Angaben aber Konsequenzen haben, ist immer auch im Zusammenhang mit der theoriegeleiteten Entwicklung der einzelnen Items zu entscheiden. Eine denkbare Konsequenz wäre z.B. die nachträgliche Entfernung eines oder mehrerer Items aus dem Erhebungsinstrument (sowohl für die Analyse als auch – als Empfehlung − für zukünftige Erhebungen). Die Ergebnisse der DIF-Analyse für das neue Erhebungsinstrument sollten auch in Zusammenhang mit den Ergebnissen der Faktoranalysen für dieses Instrument gelesen werden. So erstaunt es weniger, wenn Items, die auf einen separaten Faktor laden, auch bei DIF-Analysen auffällig werden. Die DIF-Analysen können z.B. auch das Laden auf einen anderen Faktor erklären, wenn z.B. klar wird, dass ein Item von Kindern mit unterschiedlichen Sprachfähigkeiten anders bearbeitet wird.

Für die Elimination von Items aus dem Erhebungsinstrument werden mehrere Kriterien herangezogen. Es werden beispielsweise unterschiedliche Verfahren zur Ortung von DIF eingesetzt. Wird das Item bei vielen verschiedenen Verfahren auffällig oder zeigt sich DIF in höherem Ausmass (z.B. im Ausmass «Moderate to Large» im Mantel-Haenszel-Verfahren), so ist dieses Item eingehender zu untersuchen als andere. Trotzdem kann auch bei deutlichem DIF aufgrund theoriebasierter Überlegungen entschieden werden, ein Item beizubehalten. Als Beispiel für eine solche Entscheidung können die beiden Brückenitems zur Ausgangssituation «Korb» herangezogen werden. Diese zwei Items wurden letztlich trotz Auffälligkeiten beibehalten, weil sie als einzige die zeitlich-sukzessive Grundvorstellung der Multiplikation bedienen und deshalb eine wertvolle Komponente des Erhebungsinstruments bilden. Dass das Item DIF aufgrund des Sprachverständnisses aufweist, wurde dabei mit Bedauern, aber bewusst in Kauf genommen. Zwei Items, welche hingegen aus dem Test entfernt wurden, sind die beiden Items zur Ausgangssituation «Schokolade» aus dem Vortest. Bei diesen beiden Items war die Beobachtung, dass sie von den Kindern häufig zählend gelöst wurden, mitentscheidend. Die Beobachtung, dass die zwei Items auch auf einen anderen Faktor laden und DIF bezüglich Sprachverständnis aufwiesen, trug zu dieser Entscheidung bei und erleichterte sie. Als alleiniges Kriterium für den Ausschluss dieser zwei Items wäre DIF bezüglich Sprachverständnis nicht ausreichend gewesen.

DIF bezüglich Geschlecht wurde bei einigen Items ebenfalls in Kauf genommen, ohne dass dafür eine klare Begründung gefunden worden wäre. Zumindest gab es keine inhaltlichen Hinweise darauf, weshalb diese Beobachtungen gemacht wurden.

Die einzelnen Items des neuen Erhebungsinstruments wurden alle im Lichte der Ergebnisse, die in den Kapiteln 5.4.2 und 5.5.2 berichtet werden, beleuchtet. Die Konsequenzen aus diesen Betrachtungen werden in den jeweiligen Kapiteln festgehalten. Die Erhebungsinstrumente wurden für die folgenden Analysen entsprechend angepasst.

4.5.4 Modellprüfung und Kriterien zur Modellbewertung

Unter verschiedenen möglichen Massen für die Messung der Modellgüte wurden in dieser Arbeit die beiden Masse «Akaike Information Criterion» (AIC) und «Bayes Information Criterion» (BIC) verwendet. Diese beiden informationstheoretischen Masse und ihr Nutzen für diese Arbeit werden in diesem Abschnitt erläutert.

Information Criteria AIC und BIC als Modellgüte-Masse

In die beiden Masse AIC und BIC gehen die Likelihood L des jeweiligen Modells sowie die Anzahl der zu schätzenden Parameter n_{par} ein. Die beiden Masse liefern für das jeweilige Modell und die vorliegenden Daten jeweils einen Zahlenwert, der zwischen den verschiedenen zu prüfenden Modellen verglichen werden kann. Auf diese Weise kann die Modellwahl evaluiert und so z.B. das Birnbaum-Modell mit dem Rasch-Modell verglichen werden.

Die Verwendung der zu schätzenden Parameter erfolgt in der Absicht, «sparsamen» bzw. einfachen Modellen mit verhältnismässig wenigen Parametern auch rechnerisch einen Vorzug zu geben. Tiefere Werte für AIC und BIC sind vorzuziehen, daher kommen die Summanden, welche die Anzahl Parameter abbilden, in additiver Form hinzu. Eine höhere Likelihood L hingegen ist ein Indiz für einen guten Model-Fit und wird darum mit einem negativen Vorzeichen versehen. In den Wert BIC fliesst im Gegensatz zum Wert von AIC noch die Stichprobengrösse n mit ein (Eid & Schmidt, 2014, S. 190 f.; Eid et al., 2017, S. 750).

Die beiden Werte werden berechnet mit den Zusammenhängen:

$$\mathrm{AIC} = -2 \cdot \ln \mathrm{L} + 2 \cdot n_{par} \text{ und}$$

$$\mathrm{BIC} = -2 \cdot \ln \mathrm{L} + (\ln n) \cdot n_{par}$$

Varianzaufklärung durch die einzelnen Modelle

Als Mass für die Güte bzw. die Passung eines Modells zu den vorhandenen Daten kann auch der Anteil der Gesamtvarianz betrachtet werden, der durch die jeweilige Modellausgestaltung erklärt werden kann. Dafür kommen die Masse R^2_m sowie R^2_c in Frage (Nakagawa & Schielzeth, 2013; Nakagawa et al., 2017; Hox et al., 2018).

- R^2_m bezeichnet den Anteil an der Gesamtvarianz, welcher durch die fixen Faktoren des Modells erklärt wird (m für «marginal»).
- R^2_c bezeichnet den Anteil an der Gesamtvarianz, welcher durch die fixen und die zufälligen Faktoren erklärt wird (c für «conditional»).

Die Berechnung dieser Masse erfolgt mit der Funktion r.squaredGLMM() aus dem R-Package «MuMIn» (Barton, 2020).

4.5.5 Mehrebenenanalyse

Dieses Kapitel beginnt mit der Begründung der Notwendigkeit einer Mehrebenenanalyse für die vorliegende Arbeit. Im zweiten Teil werden dann einige wesentliche Aspekte der Arbeit mit solchen Modellen beleuchtet.

Notwendigkeit des Einsatzes einer Mehrebenenanalyse

Für die Auswertung einer Studie mit dem in Kapitel 4.1 beschriebenen Design ist der Einsatz von Mehrebenenanalysen angezeigt. Mehrebenenanalysen werden eingesetzt bei Daten mit einer hierarchischen Struktur, in denen Grössen auf «verschiedenen Ebenen» einen Einfluss auf die Beobachtungen haben können. Die Gruppierung von Kindern in Schulklassen ist ein typisches solches Beispiel einer hierarchischen Struktur, das in dieser Form häufig auch in der Literatur zur Illustration solcher Modelle verwendet wird (Ditton, 1998; Langer, 2009; Snijders & Bosker, 2012; Rasbash, 2019; Finch et al., 2019). Die Beobachtungen werden dabei in verschiedenen Ebenen strukturiert dargestellt und entsprechend interpretiert. Die Anwendungsmöglichkeiten für solche Strukturen sind vielfältig und kommen in vielen Forschungsfeldern vor. Typische Beispiele sind Kinder (auf der ersten Ebene) in Schulklassen (auf der zweiten Ebene), Kinder (erste Ebene) in Schulen bzw. Schulhäusern (zweite Ebene), Arbeitnehmende (erste Ebene) in Firmen (zweite Ebene), Patientinnen und Patienten bei Ärztinnen und Ärzten oder z.B. auch Befragte und ihre jeweiligen Interviewenden (Langer, 2009).

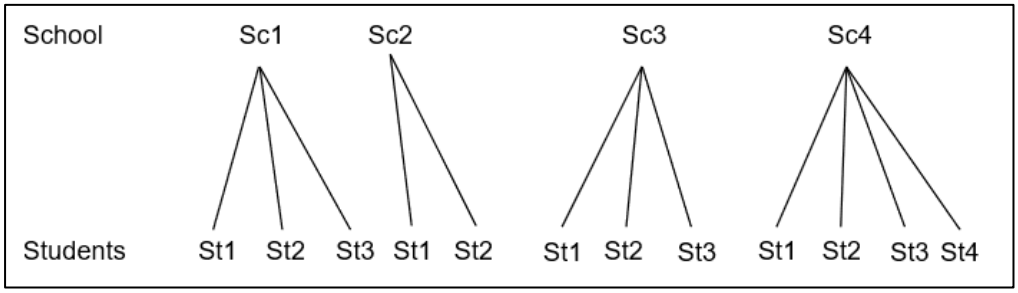


Abbildung 28: Schülerinnen und Schüler (erste Ebene) in einzelnen Schulen (zweite Ebene), www.cmm.bris.ac.uk/lemma

Die Eigenschaften der einzelnen Individuen (auf der ersten Ebene) hängen sowohl von ihren individuellen Eigenschaften und Voraussetzungen als auch von kollektiven Eigenschaften der zweiten Ebene ab. So können die Leistungen von Kindern in der Schule beispielsweise geprägt sein von ihrem individuellen Vorwissen, ihrem Alter, ihrem Geschlecht, ihrem Intelligenzquotienten usw., aber auch von Eigenschaften der Schule (Grösse des Schulhauses, Lage bzw. Region oder Kanton) oder der Schulklasse (Grösse der Klasse, Lehrperson, Lehrmittel usw.).

Möglich sind auch Interaktionseffekte zwischen individuellen Eigenschaften und kollektiven Eigenschaften. Denkbar wären solche Effekte dann, wenn z.B. eine Lehrperson die Kinder eines bestimmten Geschlechts besonders fördert oder wenn sich das Lehrmittel bei Kindern

mit hohem Vorwissen im Gegensatz zu den Kindern mit tieferem Vorwissen als besonders wirkungsvoll herausstellt.

Mehrebenenanalysen ermöglichen es, individuelle und kollektive Effekte getrennt zu orten und Interaktionseffekte zu entdecken. Die Probleme und Defizite, die bei der Untersuchung mit traditionellen Strategien gemacht werden, wenn solche hierarchische Strukturen nicht berücksichtigt werden, sind weitreichend (Ditton, 1998, S. 21 ff.; Finch et al., 2019, S. 29).

In der Mehrebenenanalyse sind auch Modelle mit mehr als zwei Ebenen denkbar. Kinder (erste Ebene) können beispielsweise Klassen (zweite Ebene) zugeordnet sein und diese könnten entweder Schulen oder Regionen (Kantonen) als dritter Ebene zugeordnet werden. Arbeitnehmende (erste Ebene) könnten in Firmen (zweite Ebene) und die Firmen einer bestimmten Branche oder einem Land (dritte Ebene) zugeordnet werden.

Auch Längsschnittuntersuchungen werden häufig in unterschiedlichen Ebenen analysiert, wobei die Ergebnisse der verschiedenen Messzeitpunkte die unterste (erste) Ebene darstellen und die jeweiligen Individuen sich dann auf der zweiten Ebene vorfinden. Sind die Individuen dann wieder in Gruppen geordnet, so entstehen auch hier Modelle mit mehr als zwei Ebenen.

In der hier vorliegenden Arbeit wird ein Modell mit zwei Ebenen gewählt. Die Kinder, die im zweiten Schuljahr des Forschungsprojekts MALKA untersucht wurden, bilden die erste Ebene, und diese Kinder sind den insgesamt 46 Schulklassen (zweite Ebene) zugeordnet.

Für die Auswertung der Wirkung der Förderung werden zwei Messzeitpunkte herangezogen (vor bzw. nach der Intervention bzw. dem Unterricht in der Kontrollgruppe). Die Messzeitpunkte werden typischerweise nur dann als eine zusätzliche (unterste) Ebene ins Modell integriert, wenn eine Serie von Messungen über mehr als zwei Zeitpunkte vorliegt. In der vorliegenden Arbeit werden für die Auswertung der Interventionen im zweiten Schuljahr (sowohl im Hinblick auf die allgemeinen mathematischen Fähigkeiten als auch im Hinblick auf das Operationsverständnis) jedoch lediglich die Messungen von zwei Messzeitpunkten verwendet. Es handelt sich dabei um die beiden Messungen vor bzw. nach den Interventionen. Die zusätzlichen Messungen zum Zeitpunkt T1 werden für die Auswertung des zweiten Schuljahrs im Längsschnitt nicht herangezogen.

Eine weitere dritte Ebene, in welcher die 46 Klassen wiederum gruppiert dargestellt werden, erscheint für den hier vorliegenden Forschungszweck ebenfalls nicht sinnvoll. Denkbar wäre beispielsweise eine Gruppierung der 46 Klassen zu Kantonen. Allerdings würden dabei einzelne Gruppen sehr klein. Zudem muss bedacht werden, dass verschiedene Eigenschaften der Schulklassen durch die Wahl der Klasse auf Ebene 2 bereits erfasst sind. Die Wahl des Lehrmittels beispielsweise steht bereits auf Ebene 2 zur Verfügung, so dass die zusätzliche Verwendung des Kantons auf Ebene 3 in dieser Hinsicht wenig zusätzliche Information liefern würde.

Die wesentlichen Eigenschaften sind also bereits auf Ebene 2 als Prädiktoren erfasst. Es handelt sich dabei vor allem um die Zuteilung zur Interventions- oder Kontrollgruppe. Denkbare weitere Eigenschaften für eine Analyse auf Ebene 2 wären z.B. das verwendete Lehrmittel,

allenfalls die Klassengrösse oder auch die Ausgestaltung der Klasse als Mischklasse (Ja/Nein).

Denkbar wäre auch noch eine Gruppierung der Klassen in einzelne Schulhäuser, doch erscheint dies angesichts der grossen regionalen Streuung der Klassen als wenig verheissungsvoll. Bei den 68 im zweiten Schuljahr beteiligten Schulklassen kam es nur einmal vor, dass mehr als zwei im selben Schulhaus unterrichtet werden. In diesem einen Schulhaus wurden vier Mischklassen von vier verschiedenen Lehrpersonen unterrichtet. Da es sich um vier Mischklassen handelte, hatten sie alle lediglich die Grösse (das «Gewicht») von Halbklassen. Bei diesen vier Mischklassen wurde zudem zum Zeitpunkt T4 bedingt durch die Covid-19-Situation im Frühjahr 2020 auf die Erhebungen verzichtet.

In sechs weiteren Schulhäusern kam es vor, dass zwei Schulklassen im Forschungsprojekt mitarbeiteten. Diese sechs Schulhäuser waren wiederum auf fünf verschiedene Kantone verteilt. Die meisten der 52 Klassen waren in ihrem jeweiligen Schulhaus alleine am Projekt beteiligt. In einem einzigen Fall kam es vor, dass eine Schulische Heilpädagogin im selben Schulhaus die MALKA-Interventionen mit zwei Klassen durchführte. Auch diese beiden Klassen hatten keine Erhebungen zum Zeitpunkt T4. In den anderen 66 Klassen waren unterschiedliche Ansprechpersonen für das Forschungsprojekt verantwortlich.

Angesichts dieser breiten Streuung der Schulklassen auf verschiedene Lehrpersonen, sehr viele verschiedene Schulhäuser und insgesamt 15 Kantone der deutschsprachigen Schweiz wurde auf die Arbeit mit einer dritten (höheren) Ebene verzichtet. Aufgrund der wenigen Messzeitpunkte wurde auch auf diese (denkbare) zusätzliche Ebene verzichtet, so dass mit einem Zwei-Ebenen-Modell gearbeitet wurde.

Auch aufgrund der Grösse der Stichprobe empfiehlt sich ein Modell mit lediglich zwei Ebenen. Für den Einsatz von drei oder mehr Ebenen wären grössere Stichproben erforderlich (D. H. Rost, 2007).

Arbeit mit Mehrebenenmodellen

Mehrebenenmodelle werden meist schrittweise eingeführt und aufgebaut. Dabei wird durch die Hinzunahme weiterer Freiheitsgrade jeweils zusätzliche Flexibilität ins Modell eingeführt. Mit der Hinzunahme neuer Parameter ergibt sich aber auch eine zusätzliche Komplexität des bisherigen Modells, so dass immer auch zu prüfen ist, ob sich der jeweilige Ausbau des Modells gelohnt hat und das Modell auch neue Informationen ans Licht bringt.

Der Bestand wird (auf der zweiten Ebene) zerlegt in einzelne Cluster oder Gruppen, die mit dem Index j versehen werden. In jeder Gruppe befinden sich einzelne Individuen oder Beobachtungen, für die der Index i eingesetzt wird. In dieser Arbeit bedeutet dies, dass der Schulklasse j verschiedene Kinder i zugeteilt sind und für die unabhängige Variable jedes Kindes X_{ij} (z.B. die kognitiven Fähigkeiten) eine abhängige Variable Y_{ij} (z.B. die Leistung im Vortest zum Operationsverständnis oder die Leistung im BASIS-MATH Erhebungsinstrument) erfasst wurde.

Der schrittweise Aufbau einer Mehrebenenanalyse beginnt mit einem reinen Nullmodell («leeres Modell» oder «intercept only»-Modell), in welchem mit einem festen Wert βo lediglich der Mittelwert für alle Individuen auf der untersten Ebene geschätzt wird (Eid et al., 2017, S. 732 ff.; Votta, 2017; Hox et al., 2018; Finch et al., 2019; Centre for Multilevel Modelling, 2019; Stöckli, 2019, S. 135 ff.). Die entsprechende Modellgleichung lautet

$$Y_i = \beta_0 + \varepsilon_i$$

Der Summand ε_i bezeichnet hierbei die Abweichung der Beobachtung beim Individuum i vom Mittelwert βo sämtlicher Individuen. Eine unabhängige Variable X kommt bei diesem Modell noch nicht zum Einsatz.

Dieses Modell wird nun stufenweise ausgebaut. Der erste Schritt umfasst die Einführung unterschiedlicher Mittelwerte für die einzelnen Gruppen auf Ebene 2. In dieser Arbeit werden diese Gruppen durch die einzelnen Klassen gebildet werden.

Die entsprechende Modellgleichung für das Nullmodell mit «random intercept» lautet

$$Y_{ij} = \beta_0 + u_j + \varepsilon_{ij}$$

In dieser Modellgleichung beschreibt der neue Parameter u_j die Abweichung des Mittelwerts der Gruppe (Klasse) j vom Gesamtmittelwert βo. Auch dieses Modell kommt ohne unabhängige Variable X aus.

Durch Verwendung einer unabhängigen Variable X kann das Modell mit einem zusätzlichen Freiheitsgrad ausgestattet werden. Die Modellgleichung für das «random intercept, fixed slope»-Modell lautet

$$Y_{ij} = \beta_0 + \beta_1 \cdot X + u_j + \varepsilon_{ij}$$

und beschreibt eine klassische lineare Regression mit einem (vorerst) festen Parameter β_1.

In einem weiteren Schritt kann auch für die Steigung der Regressionsgeraden im Modell eine gruppenabhängige Schwankung zugelassen werden. Das derart ausgebaute «random intercept, random slope»-Modell wird durch die Modellgleichung

$$Y_{ij} = \beta_0 + (\beta_1 + u_{1j}) \cdot X + u_j + \varepsilon_{ij}$$

beschrieben. Neu eingeführt wurde die Abweichung u_{1j} der Gruppe (Klasse) j vom Durchschnittswert β_1 über alle Klassen.

Selbstverständlich können auch in Mehrebenenmodellen mehrere unabhängige Variablen verwendet werden. Die entsprechenden Modelle der multiplen Regression enthalten dann mehrere Regressionskoeffizienten, so z.B. β_1 und β_2 im Falle einer Modellgleichung mit zwei unabhängigen Variablen X_1 und X_2:

$$Y = \beta_0 + \beta_1 \cdot X_1 + \beta_2 \cdot X_2 + \varepsilon$$

Die Modellgleichungen für mehr als zwei unabhängige Variablen werden auf die sinngemäss gleiche Art und Weise erweitert.

Wahl der Zentrierung bei Mehrebenenanalysen

Die Art der Zentrierung der Prädiktor-Variablen bei Mehrebenenanalysen kann unterschiedlich gewählt werden. Als Hauptmethoden sind die Zentrierung am jeweiligen Klassenmittelwert («group mean centering») oder auch die Zentrierung am Gesamtmittelwert der Stichprobe («grand mean centering») denkbar. Die beiden Methoden werden für Lernfortschrittsmessungen bei Kindern verwendet (Enders & Tofighi, 2007; Lüdtke et al., 2009). In dieser Arbeit wurde für alle Analysen die Zentrierung am Gesamtmittelwert vorgenommen (Eid et al., 2017, S. 743). Lediglich bei dichotomen Variablen (0/1-Variablen) wurde diese Art der Zentrierung nicht gewählt, da bei diesen der Nullpunkt sinnvoll interpretiert werden kann. Die dichotomen Prädiktor-Variablen wurden deshalb nicht zentriert.

Voraussetzungen für Mehrebenenanalysen

Für die Arbeit mit Mehrebenenmodellen sind die Voraussetzungen der Daten für die Anwendung der jeweiligen Modelle vorgängig zu prüfen (Raudenbush & Bryk, 2002, S. 255). Notwendig sind intervallskalierte abhängige Variablen, entweder intervallskalierte oder dichotome unabhängige Variablen sowie lineare Zusammenhänge zwischen unabhängigen und abhängigen Variablen. Es sind dies im Wesentlichen dieselben Voraussetzungen, die auch für die «gewöhnliche» multiple lineare Regression ohne die Verwendung mehrerer Ebenen zu prüfen sind. Weitere Voraussetzungen für die Mehrebenenanalyse sind normalverteilte und voneinander unabhängige Fehler auf Ebene 1 und auf Ebene 2, Unabhängigkeit der Prädiktoren auf Ebene 1 (keine Multikollinearität) sowie homogene Varianzen der Fehlerterme auf Ebene 1 (Homoskedastizität). Die Prüfung dieser Voraussetzungen ist mit Hilfe der Funktionen hist(resid()), plot() und vif() in der Software R vorgenommen worden. Die Analyse sowohl mit den Daten aus dem BASIS-MATH als auch mit den Daten aus den Erhebungen für das Multiplikationsverständnis führte zu zufriedenstellenden Ergebnissen. Somit kann das vorgesehene Mehrebenen-Modell für die Analyse des Lernfortschritts mit diesen beiden Datensätzen eingesetzt werden.

4.5.6 Auswertungsmethoden für die Längsschnittanalyse

Dieses Kapitel umfasst verschiedene Erläuterungen zur Wahl der methodischen Mittel, mit denen später die Längsschnittanalyse erfolgt.

Verwendung der Personenfähigkeit aus dem Rasch-Modell

Als Mass für die Leistung der einzelnen Kinder wird in allen drei Kapiteln 7.1, 7.2 und 7.3 die für das Rasch-Modell geschätzte Personenfähigkeit der Kinder anstelle der jeweiligen Testscores verwendet.

Wahl der Erhebungsinstrumente und Anzahl Items

Das BASIS-MATH-Erhebungsinstrument wurde unverändert übernommen. Es umfasst 28 Items (maximaler Testscore 28) bei T2 und 30 Items (maximaler Testscore 30) bei T4.

Das Erhebungsinstrument zum Multiplikationsverständnis wurde gemäss Kapitel 4.4 eingesetzt. Es wurden die um zwei Items reduzierte Version des Vortests bei T3 mit insgesamt 24 Items sowie der nicht reduzierte Nachtest bei T4 mit 26 Items verwendet.

Mehrebenen-Modellansatz für diese Arbeit

In den Längsschnittanalysen in dieser Arbeit wird das Mehrebenenmodell gemäss der folgenden Modellgleichung angesetzt:

$$Y = \beta_0 + \beta_1 \cdot X_1 + \beta_2 \cdot X_2 + \varepsilon$$

Dabei steht Y für die abhängige («outcome») Variable der Erhebung nach dem 2. Schuljahr. X_1 ist die unabhängige «Prädiktor»-Variable des Vorwissens, d.h. der Messwert der Erhebung vor dem 2. Schuljahr. Als zweite Prädiktor-Variable auf erster Ebene wird die dichotome Variable X_2 verwendet. Der Wert dieser Variable ist 1, falls die Intervention mit den Fördereinheiten für das Operationsverständnis beim jeweiligen Kind stattgefunden hat. Der Wert von X_2 beträgt 0, falls dies nicht der Fall war, weil das Kind in einer Klasse aus der Kontrollgruppe unterrichtet wurde. Wie üblich bezeichnet β_0 das Intercept und ε den jeweiligen Fehlerterm.

Das verwendete Modell ist das «random intercept»-Modell (Eid et al., 2017, S. 734), bei welchem die Parameter β_1 und β_2 jeweils fixiert sind («fixed effects»). Vergleiche mit den «random slope»-Modellen erfolgen in Kapitel 6.1.3.

Das Vorgehen in allen folgenden Analysen in diesem Kapitel ist es nun, zu bestimmen, ob sich der mit den vorliegenden Daten geschätzte Wert von β_2 im Modellansatz signifikant von 0 unterscheidet. Ist dies der Fall, dann liegt eine signifikante Wirkung der Intervention vor. Das Vorzeichen von β_2 gibt Auskunft über die Richtung dieser Wirkung. Bei positivem Vorzeichen von β_2 war der Lernzuwachs bei Einsatz der Interventionen höher als bei den Kindern der Kontrollgruppe. Bei negativem Vorzeichen von β_2 ist der Lernzuwachs bei den Kindern der Kontrollgruppe grösser als bei den Kindern in den Interventionsklassen.

Die Schulklasse, in welcher die Kinder unterrichtet wurden, wird dabei immer auf Ebene 2 im Mehrebenen-Modell modelliert und auf diese Weise kontrolliert.

Verwendung der Leistung als abhängige Variable anstelle der Leistungszunahme

Gearbeitet wird jeweils ausschliesslich mit Regressionsgleichungs-Modellen. Für die Effektmessung im Längsschnitt kann dabei für die abhängige Variable in äquivalenter Weise der Leistungszuwachs $Y - X$ (bzw. $Y_{post} - Y_{pre}$) als auch einfach die Leistung Y (Y_{post}) zum zweiten Zeitpunkt verwendet werden. Grund dafür ist, dass die Regressionsgleichungen durch ge-

eignete Umformungen ineinander übergehen (Vickers & Altman, 2001). Die entsprechende Umformung lautet

$$Y = \beta_0 + \beta_1 \cdot X + \varepsilon$$

$$Y = \beta_0 + (\beta_1 - 1) \cdot X + X + \varepsilon$$

Dieser Term ist nach Subtraktion von X auf beiden Seiten und nachdem $\beta_2 = \beta_1 - 1$ substituiert wurde äquivalent zur Alternative

$$Y - X = \beta_0 + \beta_2 \cdot X + \varepsilon$$

In dieser Arbeit wird für die Analysen in Kapitel 6 jeweils die Leistung (Y) anstelle des Leistungszuwachses (Y – X) verwendet.

Verfügbare und verwendete Prädiktoren

Als mögliche unabhängige «Prädiktor»-Variablen waren in dieser Arbeit verschiedene Grössen (Ergebnisse aus den Testungen zum Zeitpunkt T1 sowie Strukturdaten) verfügbar. An dieser Stelle wird erläutert, welche Grössen in dieser Arbeit verwendet wurden und weshalb einzelne Grössen auch nicht als Prädiktoren verwendet wurden.

Die Grössen können alle auf Ebene 1 als Prädiktoren im Modell verwendet werden. Einzelne Grössen werden für diesen Zweck vorgängig noch umcodiert. In Kapitel 6.1.3 werden die einzelnen Prädiktoren ins Mehrebenen-Modell eingeführt und auf ihre Nützlichkeit für die Vorhersage des Leistungszuwachses geprüft.

Für die Längsschnittanalyse hauptsächlich verwendete Prädiktoren:

- Vorwissen BASIS-MATH (T2) bzw. Vorwissen Operationsverständnis (T3)
- Durchführung der Interventionen Ja/Nein

Für weitere Analysen zusätzlich verwendete Prädiktoren:

- Arbeitsgedächtnis (T1, siehe Kapitel 4.1.3)
- Kognitive Fähigkeiten (T1, siehe Kapitel 4.1.3)
- Sprachverständnis für die deutsche Sprache (siehe Anhang 8.1.1)
- Ausdrucksfähigkeit für die deutsche Sprache (siehe Anhang 8.1.1)
- Familiensprache (siehe Anhang 8.1.1)
- Geschlecht
- Alter
- Lehrmittel
- Interventionsgruppe (I, II, III oder IV)

Nicht alle verwendeten Prädiktoren erwiesen sich als nützlich für die Varianzaufklärung. Einzelheiten dazu sind in Kapitel 6.1.3 zu finden.

Als Prädiktoren nicht verwendet wurden die Daten aus den verschiedenen MALKA-Tests sowie auch die Daten aus den Befragungen zur sozialen Integration, da diese Daten Bestandteil anderer Untersuchungen bildeten. Aus den vorhandenen Strukturdaten (siehe Anhang 8.1.1) wurde lediglich die Lernziel-Anpassung nicht als Prädiktor verwendet, da Kinder mit angepassten Lernzielen bereits vorgängig aus der Stichprobe für diese Untersuchung ausgeschlossen worden waren (siehe Kapitel 4.2.3). Alle anderen Strukturdaten wurden als Prädiktoren verwendet und auf diese Art auf ihre Nützlichkeit hin geprüft.

Auswahl der Prädiktoren und «lasso»-Regression

Bei der Auswahl und Arbeit mit Prädiktoren sind in der Arbeit mit Mehrebenenmodellen grundsätzlich verschiedene Vorgehensweisen möglich.

Es kann ausgehend vom Nullmodell ein immer komplexeres Modell entworfen werden und die einzelnen Prädiktoren einzeln ins Modell aufgenommen werden. Bei den einzelnen Schritten wird jeweils die Modellverbesserung (oder Modellverschlechterung) durch den neuen Prädiktor mit Hilfe von Kennzahlen wie den AIC und BIC-Werten überprüft und der Prädiktor entweder beibehalten oder verworfen. Bei diesem Vorgehen werden die Modelle mit der Zeit komplexer und umfassen mehr und mehr (hoffentlich nützliche) Prädiktoren. Ein anderes Verfahren ist es, ausgehend von einem Modell mit sehr vielen Prädiktoren einzelne wegzulassen und die Model-Fit-Werte zu prüfen, die sich durch das Weglassen von Prädiktoren ergeben. In beiden Fällen wird häufig theoriegeleitet vorgegangen. Beispielsweise werden Prädiktoren, die aufgrund von theoretischen Erkenntnissen vermutlich hilfreich sein dürften, früh ins Modell aufgenommen.

In der vorliegenden Arbeit werden Intraklassen-Korrelationen («ICC-Werte») mit dem Nullmodell berechnet. Längsschnittanalysen werden mit einem «random intercept»-Modell mit den Prädiktoren «Vorwissen» und «Intervention Ja/Nein» gerechnet. Ausgehend von dem Modell, welches das Vorwissen und die Interventionen als Prädiktoren umfasst, werden zusätzliche Prädiktoren auf ihre Nützlichkeit und Modellverbesserung hin überprüft. Die Ergebnisse hierzu finden sich in Kapitel 6.

Ein rein quantitatives Vorgehen ohne Verwendung von theoretischen Annahmen stellt die «lasso»-Regression («least absolute shrinkage and selection operator»-regression) dar. Diese Art der Regression interpoliert die Datenmenge durch die Inkaufnahme eines leichten Bias und korrigiert den β-Koeffizienten (die «slope») der Regression nach unten. Das Verfahren ist im Gegensatz zur verwandten «ridge»-Regression so ausgestaltet, dass der β-Koeffizient einzelner Prädiktoren auch den Wert 0 annehmen kann. Es wird in der Mehrebenenanalyse gerade deshalb verwendet, um aus einer grossen Menge von denkbaren Prädiktoren diejenigen zu eliminieren, die im Regressions-Modell geringen oder keinen Mehrwert bringen und vom Algorithmus den β-Koeffizienten 0 zugeordnet bekommen.

In der «lasso»-Regression wird die zu minimierende Summe der Residuenquadrate ergänzt durch einen weiteren additiven «Straf»-Term, der proportional zum Betrag $|\beta|$ des Werts von β ist. Damit werden Interpolationen mit tiefem Betrag von β bevorzugt, was auch dazu führen

kann, dass die optimale Wahl für diesen Parameter $\beta = 0$ ist. Die Höhe des Vorfaktors λ im «Straf»-Term $\lambda \cdot |\beta|$ wird zuerst durch Kreuzvalidierungsverfahren anhand des vorliegenden Datensatzes bestimmt und dann in der eigentlichen «lasso»-Regression verwendet. Die Ergebnisse dieses Verfahrens können dabei auch von dem Vorgehen bei der Kreuzvalidierung abhängen. Auch die Ergebnisse dieser Betrachtung finden sich in Kapitel 6.

Zur Bestimmung der ICC-Werte mit dem Nullmodell

In den Kapiteln 6.1, 6.2 und 6.3 werden jeweils auch die ICC-Werte des Nullmodells und somit der Anteil der Varianz zwischen den Klassen berichtet. Dieser Wert dient auch zur Entscheidung, ob der Einsatz von Mehrebenenmodellen sinnvoll bzw. notwendig ist. Die Ansichten darüber, ab welchem ICC-Wert die Mehrebenenanalyse notwendig wird, sind unterschiedlich. Im Umfeld von Kindern in Schulklassen ist zuweilen von einem Wert von 10%, manchmal auch von 5% die Rede, die Angaben schwanken aber und die Formulierungen sind häufig so gewählt, dass ab einem ICC von 10% und höher die Fehler ohne Verwendung des Mehrebenen-Ansatzes beachtlich werden können. In der Literatur werden ICC-Werte von 10% als «tief», solche von 30% als «mittel» und solche von 50% als «hoch» bezeichnet (Hartig & Bechtoldt, 2020, S. 104 ff.). Genannt werden auch Werte von 15% bis 40% (Rödiger, 2017, S. 131; Wieseke et al., 2012) für die Notwendigkeit von Mehrebenenanalysen. Es werden aber auch tiefere Werte als 10% bereits als kritisch angesehen und darauf hingewiesen, dass z.B. Werte über 1% bereits zu Schwierigkeiten führen können, wenn die Gruppierung («Cluster-Struktur») der Daten nicht beachtet wird. Manche Autoren weisen ausdrücklich darauf hin, dass es keine einheitlichen «cut-off-points» gibt bzw. dass die Ansichten zu den ICC-Werten als Entscheidungsmass unterschiedlich sind (Anderson, 2012; Woltman et al., 2012, S. 62). Als alternative Masse werden solche genannt, die zusätzlich die durchschnittliche Gruppengrösse berücksichtigen (Kish, 1965; Schoppek, 2015, S. 201). Angesichts der durchschnittlichen Grösse der Schulklassen von rund 14 bis 17 Kindern (je nach Auswahl) in diesem Projekt würde auch dieses Kriterium in der Form von

$$(\text{durchschnittliche Gruppengrösse} - 1) \cdot \text{ICC} > 1$$

den Einsatz von Mehrebenenmodellen rechtfertigen. In Schoppek (2015) wird weiter dargelegt, dass es Situationen gibt, in denen die Mehrebenenanalyse keinen Mehrwert gegenüber der klassischen Varianzanalyse bringt.

In dieser Arbeit wird der ICC-Wert jeweils berichtet und aufgrund der Grösse dieser Werte (in der Nähe von 10%) jeweils mit Mehrebenenanalysen gearbeitet.

Beurteilung der Modellverbesserungen

Beim Einfügen zusätzlicher Prädiktoren wird mit Modellgütemassen entschieden, ob diese eine zusätzliche Modellverbesserung mit sich bringen. Dafür wurden die Modellgütemasse AIC und BIC (siehe Kapitel 4.5.4) verwendet.

Ein zweites Kriterium für die Nützlichkeit eines neuen Prädiktors ist die Grösse und die statistische Signifikanz für seinen Einfluss. Dieses zweite Kriterium wird in dieser Arbeit ebenfalls betrachtet und verwendet.

Im Übrigen kann auch untersucht werden, ob der Einfluss des Prädiktors bei Anwendung der «lasso»-Regression als vernachlässigbar eingeschätzt (d.h. die «slope» des entsprechenden Prädiktors auf 0 gesetzt) wird. Prädiktoren, bei welchen dies der Fall ist, können aus dem Modell entfernt werden.

Ebenso wird auch der Varianzanteil untersucht, der durch die Ausgestaltung des jeweiligen Modells (inklusive Prädiktorenwahl) aufgeklärt werden kann. Diese Untersuchung erfolgt mit den Mitteln, die in den Kapiteln 4.5.4 und 4.5.5 erläutert wurden.

Software-Einsatz

Die Datensätze wurden von «R» aus Excel eingelesen, nachdem sie mit FileMaker und SPSS erfasst, bearbeitet und aufbereitet worden sind. Alle Auswertungen erfolgten mit R bzw. RStudio. Für die Rechnungen kamen dabei die R-Packages «TAM» (Robitzsch et al., 2020), «lme4» (Bates et al., 2020) und «lmerTest» (Friedman et al., 2020) zum Einsatz. «TAM» wurde v.a. verwendet zur Bestimmung der Personenfähigkeit im Rasch-Modell mit dem WLE-«Weighted Likelihood-Estimation»-Schätzverfahren. Die beiden Packages «lme4» und «lmerTest» und die darin enthaltenen Funktionen (wie z.B. die «lmer»-Funktion) dienten der Arbeit mit den Mehrebenenmodellen. Für die Auswahl der Prädiktoren im Mehrebenenmodell und für die dafür eingesetzte «lasso»-Regression kam das R-Package «glmnet» (Friedman et al., 2020) zum Einsatz. Für die Berechnung der Varianzaufklärung wurde das R-Package «MuMIn» (Barton, 2020) eingesetzt.

Einschränkung der Stichprobe für die Längsschnittanalyse

Die für diese Analyse verwendete Stichprobe besteht – wie in Kapitel 4.2.3 erläutert − aus den Kindern, welche sowohl an beiden BASIS-MATH-Erhebungen (vor und nach dem zweiten Schuljahr) als auch an beiden Erhebungen zum Multiplikationsverständnis mitgemacht haben. Diese Stichprobe besteht aus 680 Kindern. Aus dieser Menge wurden vor Durchführung der Analysen 11 weitere Kinder mit angepassten Lernzielen entfernt, so dass die letztlich verwendete Stichprobe einen Umfang von 669 Kindern hatte. Der Grund für diese Anpassung waren Bedenken, dass die Ergebnisse dieser in den regulären Schulklassen integrierten Kinder Ausreisser darstellen und die Analyse des Lernfortschritts verfälschen könnten.

Die im Bestand verbliebenen 669 Kinder wurden in 46 verschiedenen Klassen unterrichtet, die auf Ebene 2 des Mehrebenen-Modells kontrolliert werden. Um vergleichbare Ergebnisse zu erhalten, wurden insbesondere die Längsschnittanalysen sowohl für den BASIS-MATH als auch für das Operationsverständnis mit genau diesen 669 Kindern durchgeführt, auch wenn für die jeweiligen Instrumente einige Kinder mehr zur Verfügung gestanden wären. Kinder, welche z.B. einmal bei der BASIS-MATH-Erhebung abwesend waren, wurden also auch für

die Analyse des Multiplikationsverständnisses aus dem Bestand entfernt. Grund hierfür war die ausreichend grosse Stichprobe und die Tatsache, dass auf diese Art der Lernfortschritt zwischen den verschiedenen Erhebungsinstrumenten besser verglichen werden konnte.

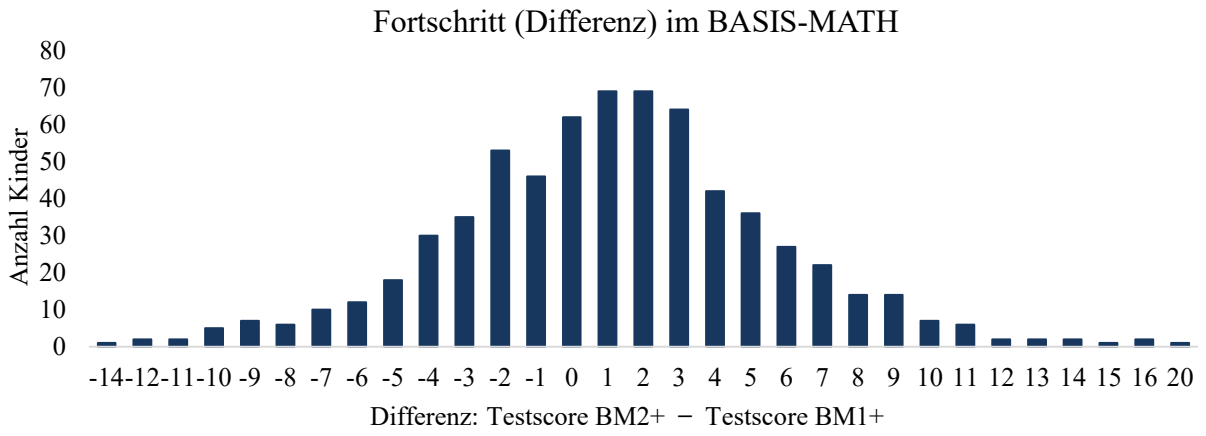


Abbildung 29: Verteilung des Fortschritts der 669 Kinder im zweiten Schuljahr (Differenz der Testscores)

Es wurde überprüft, ob die Stichprobe durch Entfernung von Kindern mit einem besonders hohen Lernfortschritt oder einem besonders hohen Lernrückschritt weiter reduziert werden sollte. Als Mass für den Lernfortschritt wurde die Differenz der Testscores (Punktezahl) in den beiden Erhebungen BASIS-MATH2+ und BASIS-MATH1+ verwendet. Die Bandbreite dieses Werts liegt zwischen -14 («Rückschritt» um 14 Punkte) und +20 («Fortschritt» um 20 Punkte). Die gesamte Verteilung ist in Abbildung 29 dargestellt. Ein sehr hoher Fortschritt könnte auch so gedeutet werden, dass das Kind im Vortest nicht seriös (konzentriert) mitgearbeitet hat, ein hoher Rückschritt könnte analog so gedeutet werden, dass das Kind zum zweiten Testzeitpunkt an der Erhebung nicht konzentriert mitgearbeitet hat.

Die einzelnen Ausreisser in dieser Art der Betrachtung wurden analysiert. Dabei zeigte sich, dass sie sich gleichmässig auf viele verschiedene Klassen und auf alle drei Interventionsgruppen verteilen. Letztlich wurden sie für die weitere Analyse im Bestand beibehalten. Dies aufgrund ihrer gleichmässigen Verteilung und ihrer insgesamt geringen Anzahl mit tiefem Einfluss bei der quantitativen Auswertung. Bedacht wurde auch die Möglichkeit, dass die Fort- und auch Rückschritte bei einzelnen Kindern tatsächlich auch hoch ausfallen können und nicht notwendigerweise damit zusammenhängen, dass die Kinder bei einer der Erhebungen unkonzentriert gewesen wären.

Auswahl der Kinder für die Betrachtung zur unterschiedlichen Rechenleistung

In den Kapiteln 6.1 und 6.2 wird nach der Betrachtung der Gesamtgruppe jeweils geprüft, ob die Interventionen für Kinder mit «schwacher» und mit «starker» Rechenleistung unterschiedlich wirken. Mit dieser Untersuchung wird auch der zweiten Forschungsfrage dieser Arbeit nachgegangen.

Für diese Unterscheidung ist es notwendig, die beiden Gruppen voneinander zu trennen. Die Trennung für den Zweck dieser Arbeit erfolgte auf unterschiedliche Arten.

a1) Bildung des unteren Quartils (Gesamtbestand)

Der Bestand wurde so getrennt, dass die Kinder im untersten Quartil der BASIS-MATH-Erhebung zum Zeitpunkt T2 als «rechenschwach» und die anderen als «rechenstark» betrachtet werden. Es wurde von der Gesamtmenge von 669 Kindern das unterste Quartil gebildet. Damit wurden diejenigen 163 Kinder mit einem Punktwert von 17 oder weniger im BASIS-MATH-1+ der «rechenschwachen» Gruppe zugeordnet. Die Kinder mit einem Punktwert von 18 oder mehr gehörten zur «rechenstarken» Gruppe.

a2) Bildung des unteren Quartils (pro Gruppe)

Die oben beschriebene Aufteilung führte dazu, dass die Kontrollgruppe (Interventionsgruppe III) verhältnismässig klein wurde und lediglich 28 (von 163) Kinder umfasste, was den Nachweis signifikanter Gruppenunterschiede erschwerte. Grund für diese kleine Gruppengrösse war die Tatsache, dass die Kontrollgruppe des 2. Schuljahrs im BASIS-MATH-Test zum Zeitpunkt T2 im Vergleich zu den anderen Gruppen einen relativ hohen Wert aufwies. Als Erklärung hierfür dient die intensive Förderung in dieser Gruppe im Laufe des ersten Schuljahrs. Aus diesem Grund wurde die Einteilung in «rechenstark» und «rechenschwach» noch einmal so vorgenommen, dass in allen vier Interventionsgruppen jeweils einzeln das unterste Quartil gewählt wurde. Damit blieben 42 (von 167) Kindern in der Kontrollgruppe. Bei gleichem Testscore im BASIS-MATH-1+ wurde dabei als Selektionskriterium der Testscore im BASIS-MATH-0+ zu Beginn des ersten Schuljahrs herangezogen und so dafür gesorgt, dass die vier Interventionsgruppen mit 41 (I), 43 (II), 42 (III) bzw. 41 (IV) besetzt wurden. Diese zweite Einteilung bewirkte eine Umteilung von insgesamt 28 Kindern. 12 Kinder wurden neu den «rechenstarken» zugeordnet, 16 Kinder wurden neu den «rechenschwachen» zugeordnet. Für die restlichen 641 Kinder änderte sich die Einteilung nicht.

b) Bildung des unteren 20%-Quantils (anstelle von 25%)

Um zu prüfen, ob die Effekte stabil sind und sich nicht durch eine leichte Verschiebung der verwendeten Quantile verändern, wurden weitere Varianten gerechnet. Eine davon ist die Bildung des unteren 20%-Quantils (anstelle des 25%-Quantils). Auch für diese Variante wurden beide Arten der Aufteilung *b1)* «am Gesamtbestand» oder *b2)* «pro Gruppe» gerechnet.

c) Bildung der unteren «schwächeren» Hälfte der Klasse («50%-Quantil»)

Diese Aufteilung wurde am Gesamtbestand vorgenommen, da sich keine technischen Schwierigkeiten mit der Grösse der Kontrollgruppen ergaben. Es wurde der Testsscore im BASIS-MATH-1+ als Aufteilungskriterium verwendet. Um Zufallseffekte zu vermeiden, wurden

auch hier zwei Varianten mit «cut-off»-Punkt bei Testscore *c1)* 21 («Wahl aller Kindern mit 21 Punkten oder weniger») und *c2)* 22 («Wahl aller Kindern mit 22 Punkten oder weniger»). Die so gebildete «schwächere Hälfte» der Kinder umfasste in den zwei Varianten 317 (47.4%) bzw. 357 (53.4%) der insgesamt 669 Kinder.

Bei allen Varianten a1), a2), b1), b2), c1) und c2) wurden die sinngemäss gleichen Analysen immer auch mit dem komplementären Bestand der «stärkeren» Kinder gerechnet.

5 Empirischer Teil 2: Analyse des Erhebungsinstruments für das Multiplikationsverständnis

In diesem Kapitel wird das Erhebungsinstrument, dessen Entwicklung und Erprobung in Kapitel 4.4 beschrieben wurde, mit Hilfe der im Projekt eingeholten Daten analysiert und besprochen. Analysiert werden sowohl der Vor- als auch der Nachtest dieses Erhebungsinstruments.

Der Vortest wurde im September 2019 mit insgesamt 1'084 Kindern zu Beginn des 2. Schuljahrs in den 68 MALKA-Klassen durchgeführt. Der Nachtest wurde in den Monaten Juni und Juli 2020 mit 738 Kindern am Ende des 2. Schuljahrs in 46 MALKA-Klassen durchgeführt. Die Reduktion der Klassenanzahl hängt – wie in der Einleitung und in Kapitel 4 dargelegt − ausschliesslich mit der Covid-19-Situation und den damit verbundenen Komplikationen zusammen.

Die Analyse der beiden Versionen des Erhebungsinstruments erfolgt jeweils mit denselben Verfahren. Dieses Kapitel ist nach den angewendeten Verfahren gegliedert und die beiden Versionen des Erhebungsinstruments werden jeweils beide im selben Abschnitt besprochen. Das Hauptaugenmerk gilt dabei dem Vortest, welcher unter regulären Umständen (vor Eintreten der Covid-19-Situation) und deshalb auch von mehr Kindern abgelegt wurde als der Nachtest. Trotzdem erfolgt ein grosser Teil der Analysen auch für die Nachtest-Version des Erhebungsinstruments. Dies geschieht insbesondere auch darum, weil mit 738 Kindern aus 46 Klassen ein Grossteil des ursprünglichen Bestands auch den Nachtest abgelegt hat.

5.1 Datenaufbereitung und deskriptive Statistiken (Vor- und Nachtest)

In diesem Abschnitt werden die Codierung des Erhebungsinstruments und die Umwandlung der ursprünglichen Codierung für die weitere Analyse erläutert. Im Anschluss daran werden erste deskriptive Statistiken und Kennzahlen wiedergegeben.

5.1.1 Korrektur und Umwandlung von Sondercodes, Umgang mit fehlenden Daten

Das Vorgehen wird getrennt dargelegt für den Vor- bzw. den Nachtest, auch wenn es in vielen Hinsichten bei den beiden Versionen des Erhebungsinstruments sehr ähnlich war.

Vortest Multiplikationsverständnis

Der in Kapitel 4.4.5 beschriebene Vortest für das Multiplikationsverständnis wurde zu Beginn des Schuljahrs 2019/20 (zum Testzeitpunkt T3) in 68 Schulklassen mit den Kindern der zweiten Klassen durchgeführt. Insgesamt wurden zu diesem Zeitpunkt 1'084 Kinder durch dieses Testinstrument erfasst. Einige Absenzen einzelner Kinder aufgrund von Krankheit zum Testzeitpunkt mussten dabei in Kauf genommen werden. Bei mehr als zwei Absenzen in einer Klasse oder sehr kurzem Anreiseweg für die testleitende Person wurde die Erhebung aber

nachgeholt. Mit der Limitierung der Absenzen (auf maximal eine pro Klasse) wurde sichergestellt, dass keine systematischen Fehler aufgrund dieser Absenzen auftraten. Die Absenzen können als zufällige Ausfälle aus der Stichprobe betrachtet werden.

Die Korrektur erfolgte gemäss der Anleitung, die im Anhang 8.2.2 wiedergegeben ist. Das Ziel der Codierung war es, die Bearbeitung durch die Kinder letztendlich dichotom (jeweils mit dem Code 0 oder mit dem Code 1) zu beurteilen.

Es wurden dabei Sondercodes («4» und «5») verwendet für typische Fehler der Kinder, welche in der ersten Klasse häufig noch «mild» bewertet werden. Der Sondercode «4» wurde verwendet für Bearbeitungen, bei denen die Zehner- und die Einerstelle (z.B. Kinderantwort 12 anstelle der korrekten Zahl 21) verwechselt wurde. Der Sondercode «5» wurde verwendet für Bearbeitungen, bei denen die 6 mit der 9 verwechselt wurde. Diese Codierung kam zur Anwendung, um im Einklang zu bleiben mit dem Instrument BASIS-MATH 1+, welches wenige Monate vorher zum Testzeitpunkt T2 zum Einsatz gekommen war. Insgesamt mussten die Sondercodes «4» und «5» beide sehr selten verwendet werden. Unter 28'184 (1'084 Kinder zu je 26 Items) kam der Sondercode «4» genau 68mal, der Sondercode «5» genau 33mal zum Einsatz. Diese 68 bzw. 33 Verwendungen waren verteilt über verschiedene Items, so dass die quantitativen Auswirkungen der Umcodierung für die beiden Codes gering waren. Diese beiden Codes wurden für die Auswertung genauso wie beim Testinstrument BASIS-MATH 1+ letztendlich zu einer «1» umcodiert, was einer milden Bewertung dieser zwei Fehlerarten entsprach.

Der Sondercode «9» für unleserliche oder unverständliche Bearbeitungen durch die Kinder kam insgesamt 12mal zur Anwendung. Eine Analyse der Bearbeitungen zeigte, dass dieser Code sinnvollerweise in den Code «0» umgewandelt werden kann. Es handelte sich meist um Bearbeitungen, die gänzlich unleserlich waren oder als «Kritzeleien» bezeichnet werden könnten, nicht aber um Bearbeitungen, die bei grosszügiger Betrachtung als «richtig» hätten interpretiert werden können. Auch diese Umcodierung hat keinerlei wesentliche Auswirkung auf die weiteren Analysen, weil die Verwendung des Sondercodes «9» derart selten war.

Die Kinder wurden so instruiert, dass sie Aufgaben auslassen sollen, wenn diese zu schwierig seien. Den Kindern wurde für die Bearbeitung des Testheftes auch ausreichend Zeit gelassen, so dass ausgelassene und leere Bearbeitungen mit 0 codiert wurden.

Diese Art der Umcodierung geschah im Bewusstsein, dass kein Vorgehen bei der Umcodierung «perfekt» oder «optimal» ist. Das Vorgehen wird an dieser Stelle aber der Transparenz willen ausführlich beschrieben (D. H. Rost, 2007, S. 195 ff.). Die Auswirkungen auf die Auswertung sind für alle Sondercodes in dieser Testerhebung vernachlässigbar.

Insgesamt führte dies nach Umcodierung der Sondercodes zu 10'146 Bearbeitungen von Kindern mit der Bewertung «0» und 18'038 Bearbeitungen von Kindern mit der Bewertung «1».

Nachtest Multiplikationsverständnis

Der im Frühjahr 2020 von 738 Kindern abgelegte Nachtest enthielt die beiden im Vortest verwendeten Sondercodes nicht mehr. Bearbeitungen, die im Vortest mit den Sondercodes «4» und «5» codiert wurden, wurden im Nachtest mit 0 codiert, was dem Standard am Ende der zweiten Klasse entspricht.

Im Nachtest wurde aber ein Sondercode «4» für die «trivialen» Bearbeitungen (Multiplikation mit 1 und Addition bzw. Subtraktion der 0) verwendet (siehe 8.2.4). Die Hauptabsicht dabei war es, diese Bearbeitungen, die auch schon im Vortest vorgekommen waren, auf ihre Häufigkeit hin zu untersuchen. Der Sondercode «4» wurde insgesamt lediglich 20mal, verteilt auf verschiedene Items, verwendet. Auch hier ist der quantitative Einfluss der Umcodierung vernachlässigbar. Der Sondercode wurde letztlich in eine «1» umcodiert, womit die Bewertung dieser Antwort mit dem Vortest konsistent gewählt wurde.

Der Sondercode «9» für «unleserlich» wurde nur einmal gewählt bei einer Bearbeitung, in der ein Kind die Multiple-Choice-Items mit Kreuzen statt Kreisen markiert hat. Dieser Code wurde – wiederum ohne Relevanz für die weiteren Rechnungen – in eine 1 umgewandelt.

5.1.2 Testscore und deskriptive Statistiken

Auch in diesem Kapitel erfolgen die Angaben sowohl für den Vor- als auch für den Nachtest.

Testscore (Vortest Multiplikationsverständnis)

Addiert man für jedes Kind die einzelnen Codes der 26 Items, so ergibt sich ein Wert zwischen 0 und 26. In dieser Arbeit wird dieser Wert jeweils als Testscore des jeweiligen Kindes bezeichnet. Die Verteilung der Testscores der 1'084 Kinder im Vortest für das Multiplikationsverständnis kann der Graphik in Abbildung 30 entnommen werden.

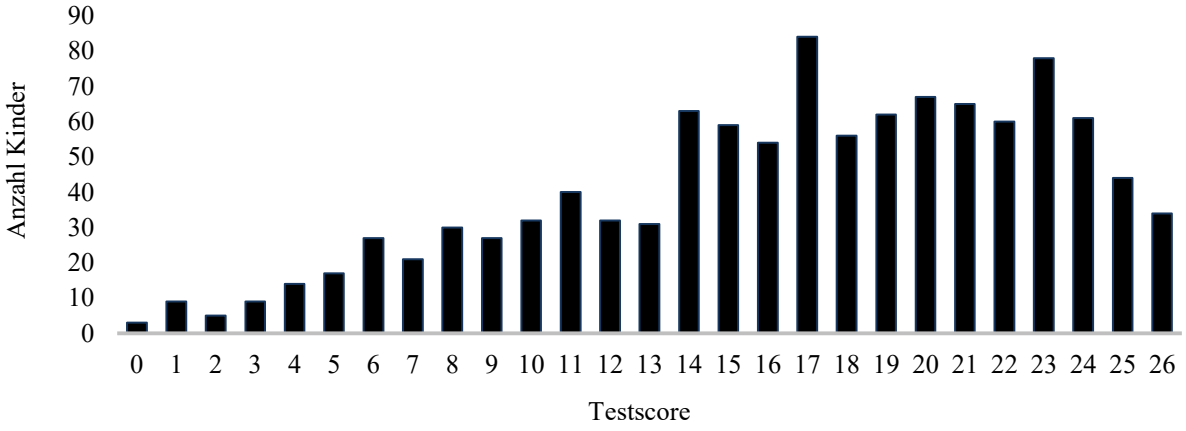


Abbildung 30: Testscore der 1'084 Kinder (Vortest für das Multiplikationsverständnis)

Der Mittelwert der Testscore-Werte lag bei 16.6, die Standardabweichung bei 6.1, der Median bei 17. Der Testscore-Wert von 26 wurde von insgesamt 34 Kindern erreicht, während drei Kinder einen Testscore-Wert von 0 erreicht haben.

Geschlechterunterschiede (Vortest Multiplikationsverständnis)

Der Test wurde von 1'084 Kindern bearbeitet, davon waren 530 Mädchen und 554 Knaben. Die Geschlechter haben den Test unterschiedlich erfolgreich bearbeitet. Die 530 Mädchen haben einen Mittelwert von 15.9 bei einer Standardabweichung von 5.9 erreicht, die 554 Knaben haben einen Mittelwert von 17.3 bei einer Standardabweichung von 6.2 erreicht. Die Unterschiede in den Mittelwerten sind signifikant, was bei der grossen Stichprobe aber auch nicht allzu stark erstaunt. Der t-Test unabhängiger Stichproben verwirft die Nullhypothese gleicher Mittelwerte ($t(1082)=-3.83$, $p<.001$, $d=-0.23$).

Alternative Item-Gestaltung nach Ausgangssituationen (Vortest Multiplikationsverständnis)

Als Alternative zur Addition der 26 Item-Bewertungen wurde die Idee untersucht, anstelle der 26 Items die 19 Ausgangssituationen zu verwenden. Bei dieser Betrachtung würden jeweils diejenigen (insgesamt sieben) Situationen, in denen sowohl nach einer Zahl als auch nach einer Rechnung gefragt wurde, als ein einzelnes «Item» aufgefasst. Eine Ausgangssituation wird dann mit 1 bewertet, wenn beide Fragen – sowohl nach der Zahl als auch nach der Rechnung – mit 1 bewertet werden.

Der Mittelwert der mit 1 bewerteten Bearbeitungen liegt bei dieser Betrachtung bei 11.8, die Standardabweichung bei 4.6, der Median bei 12. Der höchste mögliche Testscore-Wert von 19 wird bei dieser Betrachtung weiterhin von (genau denselben) 34 Kindern erreicht, während sieben Kinder bei dieser Betrachtung einen Testscore-Wert von 0 erreichen. Die Verteilung kann Abbildung 31 entnommen werden.

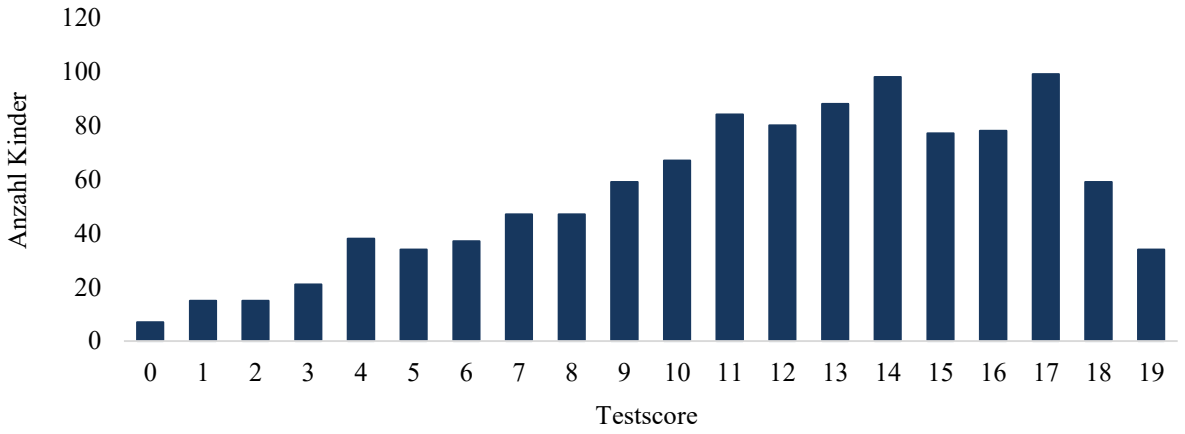


Abbildung 31: Testscore der 1'084 Kinder nach Ausgangssituation (Vortest für das Multiplikationsverständnis)

Bei verschiedenen Auswertungen (z.B. bei der Bestimmung der Reliabilität, der Passung der Items etc.) ergaben sich nur geringfügige Unterschiede zwischen diesen beiden Betrachtungen. Aus diesem Grunde wurde entschieden, alle 26 ursprünglichen Items für die weitere Analyse zu behalten. Auf diese Art geht für weiterführende Untersuchungen am wenigsten Information verloren.

Testscore (Nachtest Multiplikationsverständnis)

Auch im Nachtest lag die Anzahl der verschiedenen Items bei 26. Dies ermöglichte für die einzelnen Kinder als Testscore jeweils einen Wert zwischen 0 und 26. Die Verteilung des Testscores der 738 Kinder, welche den Nachtest abgelegt haben, kann Abbildung 32 entnommen werden.

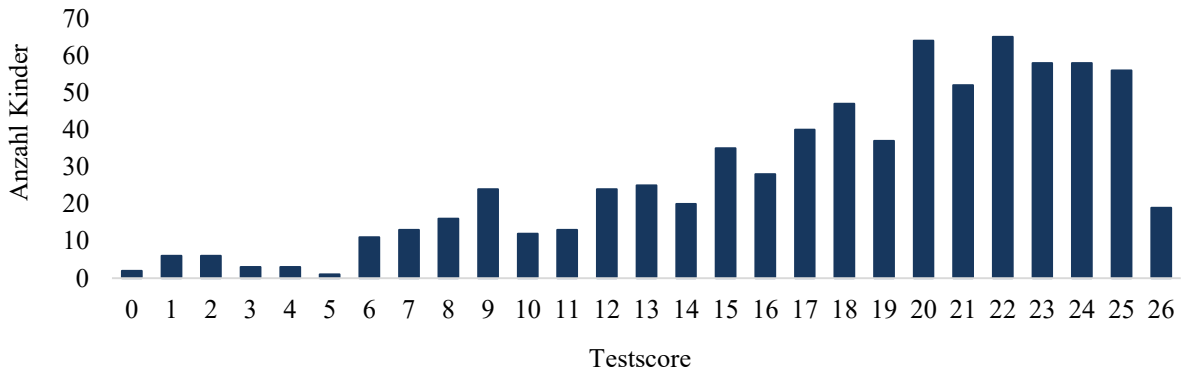


Abbildung 32: Testscore der 738 Kinder (Nachtest für das Multiplikationsverständnis)

Der Nachtest wurde trotz den Schulschliessungen besser gelöst als der Vortest. Die Kinder erreichten nach Umcodierung bei 69.4% der Items eine 1. Dieser Wert lag im Vortest noch bei 64.0% und dürfte bei einem regulären Schuljahr ohne Schulschliessungen noch höher zu liegen kommen.

Der höchste mögliche Testscore-Wert von 26 wurde von 19 der 738 Kinder erreicht, den tiefsten möglichen Wert von 0 erreichten 2 Kinder.

Deckeneffekte bei der ausschliesslichen Betrachtung der Brückenitems

Die 12 verwendeten Brückenitems stellten sich im Nachtest als verhältnismässig leicht heraus. Dies führt zu einem Deckeneffekt, falls man ausschliesslich die 12 Brückenitems betrachtet.

In Abbildung 33 ist die Verteilung des Testscores der Kinder zu sehen, die in Kapitel 6 für die Einschätzung des Längsschnitts herangezogen werden. Die Darstellung zeigt die Werte für

den Testscore, welche die Kinder im Nachtest erreichten, wenn man das Testinstrument auf die 12 Brückenitems reduziert.

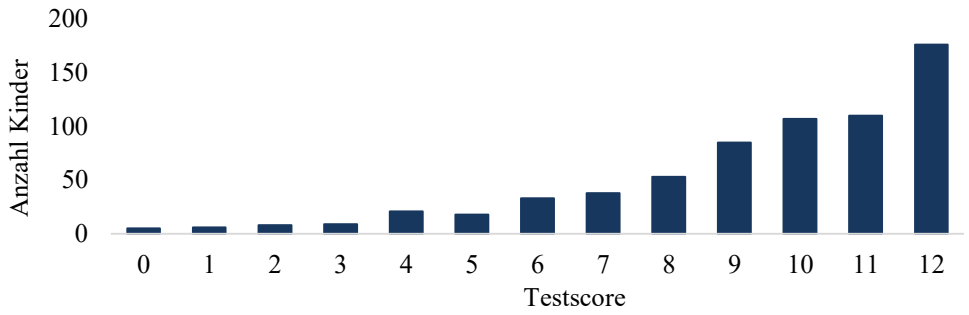


Abbildung 33: Anzahl Kinder pro Testscore im Nachtest für das Multiplikationsverständnis (12 Brückenitems)

Der Item-Schwierigkeitsindex wird im Folgenden insbesondere in Kapitel 5.2.1 vertieft betrachtet und in Tabelle 10 für die Brückenitems wiedergegeben. Die Konsequenzen für die Analyse des Längsschnitts werden in Kapitel 6.3.2 erneut aufgegriffen.

Geschlechterunterschiede (Nachtest Multiplikationsverständnis)

Der Test wurde von 738 Kindern bearbeitet, davon waren 365 Mädchen und 373 Knaben. Die Geschlechter haben den Test unterschiedlich erfolgreich bearbeitet. Die 365 Mädchen haben einen Mittelwert von 17.5 bei einer Standardabweichung von 5.8 erreicht, die 373 Knaben haben einen Mittelwert von 18.6 bei einer Standardabweichung von 5.8 erreicht. Die Unterschiede in den Mittelwerten sind wie schon beim Vortest signifikant. Der t-Test unabhängiger Stichproben verwirft die Nullhypothese gleicher Mittelwerte ($t(736)=-2.44$, $p=.015$, $d=-0.18$).

5.2 Analyse mittels klassischer Testtheorie (Vor- und Nachtest)

In diesem Abschnitt werden sowohl für den Vor- als auch für den Nachtest zum Multiplikationsverständnis die Ergebnisse der Auswertungen mittels klassischer Testtheorie (KTT) präsentiert. Die Analysen umfassen den Item-Schwierigkeitsindex für die jeweiligen 26 Items, die Trennschärfen der Items und die interne Konsistenz des Tests (gemessen mit Cronbachs Alpha). Auch werden die Interitem-Korrelationen präsentiert.

5.2.1 Item-Schwierigkeitsindex

Es folgen die Angaben getrennt für den Vor- bzw. für den Nachtest des Multiplikationsverständnisses.

Vortest Multiplikationsverständnis

Der Item-Schwierigkeitsindex für die 26 Items dieses Testinstruments ergibt sich für jedes einzelne Item aus der relativen Häufigkeit der richtigen Bearbeitung durch die Kinder (Moosbrugger & Kelava, 2012, S. 81 ff.). Die Item-Varianzen ergeben sich für dichotom codierte Items direkt aus den Werten des Schwierigkeitsindexes und liefern keine neue Information, weshalb sie an dieser Stelle nicht angegeben werden.

Tabelle 8: Item-Schwierigkeit im Vortest

Item	Schwierigkeitsindex	Item	Schwierigkeitsindex
1	.91	9a	.82
2	.78	9b	.71
3	.85	10a	.46
4a	.69	10b	.36
4b	.68	11	.72
5a	.78	12	.57
5b	.64	13	.36
6a	.47	14	.74
6b	.44	15	.39
7a	.72	16	.71
7b	.58	17	.74
8a	.54	18	.82
8b	.40	19	.77

Dem Schwierigkeitsindex der einzelnen Items kann entnommen werden, dass viele Items eher «leicht» waren. Werte im Bereich von .50 sind wünschenswert, leichtere Items (mit höheren Werten) eignen sich insbesondere zur Differenzierung zwischen schwächeren Kindern. Der

hohe Wert bei Item 1 ist kein Zufall, da ein besonders leichtes Item zum Einstieg gewählt wurde, damit die Kinder nicht unnötigerweise verunsichert werden.

Nachtest Multiplikationsverständnis

Der Item-Schwierigkeitsindex für die einzelnen Items des Nachtests wurde auf dieselbe Art berechnet wie für den Vortest. Die Ergebnisse lauten:

Tabelle 9: Item-Schwierigkeit im Nachtest

Item	Schwierigkeitsindex	Item	Schwierigkeitsindex
1a	.91	9b	.67
1b	.84	10a	.56
2	.29	10b	.59
3a	.70	11	.92
3b	.70	12	.69
4	.43	13	.68
5a	.70	14	.73
5b	.69	15	.78
6	.74	16	.86
7	.88	17	.71
8a	.76	18	.80
8b	.72	19	.56
9a	.62	20	.53

Wie schon im Vortest wurden viele Items sehr gut gelöst. Wie in Kapitel 5.1.2 beschrieben wurden die Items besser gelöst als beim Vortest. Eine solche Veränderung wurde erwartet, da 12 der 26 Items Brückenitems waren, die unverändert aus dem Vortest übernommen wurden. Die Veränderung bei diesen Brückenitems zeigt sich in der folgenden Übersicht.

Tabelle 10: Item-Schwierigkeit der Brückenitems (Vor- und Nachtest)

Brückenitem	Item-Nr. im Vortest	Item-Nr. im Nachtest	Item-Schwierigkeit im Vortest	Item-Schwierigkeit im Nachtest
Stifte	2	7	.78	.88
Autos (Zahl)	4a	1a	.69	.91
Autos (Rechnung)	4b	1b	.68	.84
Bücher (Zahl)	6a	3a	.47	.70
Bücher (Rechnung)	6b	3b	.44	.70
Stühle (Zahl)	8a	8a	.54	.76
Stühle (Rechnung)	8b	8b	.40	.72
Korb (Zahl)	10a	5a	.46	.70
Korb (Rechnung)	10b	5b	.36	.69
Füsse	13	13	.36	.68
Beeren (3 Teller)	14	11	.74	.92
Beeren (7 Teller)	15	12	.39	.69

Die 14 zusätzlichen Items des Nachtests waren gegenüber den Brückenitems schwieriger zu lösen, was zur Differenzierung zwischen leistungsstärkeren Kindern beitrug.

5.2.2 Trennschärfen und interne Konsistenz

Auch in diesem Kapitel erfolgen die verschiedenen Angaben einzeln für den Vor- bzw. für den Nachtest.

Vortest Multiplikationsverständnis

Für die Erhebung im Vortest ergab sich als Reliabilitätsstatistik ein Wert von $\alpha = .89$ für Cronbachs Alpha. Die Trennschärfen der einzelnen Items können Tabelle 11 entnommen werden. Als «gute» Trennschärfen gelten Werte zwischen .4 und .7 (Moosbrugger & Kelava, 2012, S. 86). Somit weisen bei dieser Erhebung die meisten Items eine «gute» Trennschärfe auf. Ausnahmen bilden das besonders leichte Einstiegsitem «Item 1» mit einem Wert von .26 sowie die beiden Items 5a und 16 mit Werten zwischen .3 und .4.

Nachtest Multiplikationsverständnis

Für die Erhebung im Nachtest ergab sich als Reliabilitätsstatistik ein Wert von $\alpha = .89$ für Cronbachs Alpha. Die Trennschärfen der einzelnen Items können

Tabelle 12 entnommen werden. Von insgesamt 26 Items weisen deren 18 «gute» Trennschärfen im Bereich von .4 bis .7 auf. Alle diese Werte liegen zwischen .4 und .6. Bei den acht anderen Items (Item 1a, 2, 7, 8a, 8b, 11, 18 und 20) wurden Trennschärfen ausserhalb der Bandbreite .4 bis .7 beobachtet. Alle Werte für diese acht Items liegen zwischen .3 und .4.

Tabelle 11: Trennschärfen («Item-Skala-Korrelationen») im Vortest

Item	Korrigierte Item-Skala-Korrelation	Item	Korrigierte Item-Skala-Korrelation
1	.26	9a	.52
2	.45	9b	.54
3	.46	10a	.49
4a	.43	10b	.51
4b	.49	11	.51
5a	.30	12	.51
5b	.46	13	.46
6a	.51	14	.44
6b	.49	15	.42
7a	.43	16	.38
7b	.51	17	.52
8a	.47	18	.48
8b	.47	19	.47

Tabelle 12: Trennschärfen («Item-Skala-Korrelationen») im Nachtest

Item	Korrigierte Item-Skala-Korrelation	Item	Korrigierte Item-Skala-Korrelation
1a	.39	9b	.58
1b	.43	10a	.42
2	.34	10b	.53
3a	.49	11	.35
3b	.48	12	.45
4	.49	13	.54
5a	.53	14	.42
5b	.55	15	.44
6	.59	16	.43
7	.39	17	.44
8a	.33	18	.39
8b	.33	19	.40
9a	.51	20	.34

5.2.3 Inter-Item-Korrelationen

Die Inter-Item-Korrelationsmatrizen des Vortests finden sich der Lesbarkeit halber im Anhang. Die erste Tabelle im Anhang 8.4.1 gibt sämtliche Werte wieder. Es ergeben sich 152 verschiedene Inter-Item-Korrelationswerte (die trivialen Korrelationen von 1.00 eines Items mit sich selbst nicht mitgezählt). Sämtliche Werte sind positiv und alle Werte sind auf einem 1% Niveau signifikant unterschiedlich von 0. Interessant ist es insbesondere, die besonders hohen Werte zu betrachten. In der zweiten Tabelle sind nur die Werte über .30, in der dritten Tabelle sind nur die Werte über .35 dargestellt. Farbig markiert sind die Items mit einem inhaltlichen Zusammenhang bzw. einem sehr ähnlichen Itemformat. An verschiedenen Stellen wird die hohe Korrelation der Kinderantworten bei der Frage nach der Zahl bzw. nach der Rechnung für dieselbe Ausgangssituation ersichtlich. Items 11 bis 15 weisen eine gewisse «Verwandtschaft» auf, da bei diesen fünf Items jeweils nach einer Zahl gefragt wird. Items 16 bis 19 wiederum sind untereinander «verwandt» indem bei ihnen – bei nicht angeleiteter selbstständiger Bearbeitung durch die Kinder – für verschiedene Punktefelder und Punktemuster nach einer Rechnung gefragt wird. Besonders die Einschränkung auf Werte über .35 macht das entstehende Korrelationsmuster besonders gut sichtbar.

Die Inter-Item-Korrelationsmatrizen des Nachtests finden sich ebenfalls im Anhang. Die vierte Tabelle im Anhang 8.4.1 gibt sämtliche Werte wieder. Es ergibt sich dieselbe Anzahl Werte wie beim Vortest, da der Nachtest in der eingesetzten Form dieselbe Anzahl Items enthält wie der Vortest. Wie schon beim Vortest sind sämtliche Werte positiv und auf einem 1%-Niveau signifikant unterschiedlich von 0. Es ergeben sich ähnliche Korrelationsmuster wie im Vortest, insbesondere hohe Zusammenhänge zwischen den Bearbeitungen der Kinder zur selben Ausgangssituation.

Ähnlich wie beim Vortest zeigt sich eine Verwandtschaft zwischen den Items 15 bis 18, in denen mit Punktefeldern gearbeitet wird. Dieser Effekt wird auch bei der folgenden Faktoranalyse in Kapitel 5.3 in Erscheinung treten und dort noch eingehend diskutiert werden.

5.3 Faktoranalysen

Die Faktoranalysen wurden für den Vor- bzw. für den Nachtest getrennt durchgeführt. Die Ergebnisse dieser Analysen werden an dieser Stelle berichtet.

5.3.1 Faktoranalyse für den Vortest

Die Stichprobe des Vortests mit den 1'084 Kinderbearbeitungen und den 26 Items wurde einer explorativen Faktoranalyse unterzogen. Diese Analyse geschah mit den dichotom umcodierten Werten der Kinderbearbeitungen. Das Ziel der Untersuchung war es, Aufschluss darüber zu erhalten, ob und gegebenenfalls wie das mit diesem Instrument erhobene Konstrukt in Faktoren zerfällt. Ein zweites Ziel war es, noch einmal darüber zu befinden, ob einzelne Items für die weiteren Rechnungen aus der Analyse ausgeschlossen werden sollten.

Dabei stellt sich zuerst die Frage, ob diese explorative Faktoranalyse bei dem vorliegenden Datensatz sinnvoll ist, d.h. ob die notwendigen Voraussetzungen für eine explorative Faktoranalyse gegeben sind.

Falls dies der Fall ist, so stellt sich die Folgefrage nach der sinnvollsten Wahl für die Anzahl Faktoren. Für diese Untersuchung werden verschiedene Kriterien herangezogen, welche beim selben Datensatz à priori zu unterschiedlichen Ergebnissen führen können. Die Besprechung dieser verschiedenen Kriterien und das Fazit aus den Ausführungen sind in diesem Abschnitt dargelegt.

In diesem Abschnitt werden also die Voraussetzungen für eine explorative Faktoranalyse geprüft. Sodann wird die Faktoranalyse durchgeführt und die Ergebnisse derselben dargestellt.

Voraussetzungen für eine explorative Faktoranalyse

Als Voraussetzung für die Eignung der Variablen für eine explorative Faktoranalyse mit der Hauptkomponentenanalyse sind drei Kriterien zu prüfen. Einerseits bedarf es einer substantiellen Korrelation zwischen den einzelnen Variablen (Items), im weiteren einen Kaiser Meyer Olkin (KMO)-Kennwert von mindestens .6 sowie einen signifikanten Bartlett-Test auf Sphärizität mit welchem die Nullhypothese der vollkommenen Unkorreliertheit der einzelnen Items verworfen wird (Field, 2013; Schweinberger, 2018, S. 20).

Diese Voraussetzungen wurden geprüft und sind beim vorliegenden Datensatz gegeben. Die Korrelation zwischen den einzelnen Items ist im Anhang 8.4.1 dargelegt. Alle Inter-Items-Korrelationen sind wie in Kapitel 5.2.3 dargelegt positiv und alle Items korrelieren signifikant miteinander. Der KMO-Kennwert für die Stichprobeneignung beträgt .875 und ist somit ausreichend hoch, der Bartlett-Test verwirft (χ^2(df=325) = 8'695.5, p<.001).

In der Literatur wird kontrovers diskutiert, ob die Hauptkomponentenanalyse als Standardverfahren für eine Faktoranalyse bei dichotomen Daten angezeigt ist (stats.stackexchange.com, 2011). Es wird darauf hingewiesen, dass eine solche Faktoranalyse mit dichotomen Daten mit Modellen der IRT in Zusammenhang steht (Takane & De Leeuw, 1987; Kamata & Bauer,

2008). Dieser Zusammenhang kann als Argument für die Verwendung der Hauptkomponentenanalyse gedeutet werden. In dieser Arbeit wird nun die Hauptkomponentenanalyse auf den vorliegenden dichotomen Datensatz angewendet.

Wahl der Anzahl Faktoren

Für die Wahl der Anzahl Faktoren gibt es verschiedene Kriterien, von denen das Kaiser-Kriterium, die Analyse des Screeplots und die Parallelanalyse die zentralen Kriterien darstellen (Hayton et al., 2004; Field, 2009, S. 627 ff.; Moosbrugger & Kelava, 2012, S. 326 ff.; Eid et al., 2017, S. 923 ff.; Schweinberger, 2018).

Der vorliegende Datensatz wurde auf diese Kriterien hin untersucht. Dabei ergaben sich 8 Eigenwerte, die über 1 lagen und damit das Kaiser-Kriterium erfüllen, aber auch ein Screeplot, der so interpretiert werden kann, dass nur ein einzelner Faktor zu wählen ist (siehe Anhang 8.4.3). Die stärkste Beugung («Knick») tritt bei Verbindung der Eigenwerte im Screeplot bereits nach dem ersten Faktor auf.

Tabelle 13: Die grössten zehn Eigenwerte für den Vortest, von denen acht über 1 liegen.

Faktor	Eigenwert	% der Varianz	Kumulierte %
1	7.080	27.231	27.231
2	1.633	6.282	33.512
3	1.387	5.335	38.847
4	1.323	5.087	43.935
5	1.193	4.587	48.522
6	1.136	4.369	52.891
7	1.083	4.164	57.054
8	1.018	3.917	60.971
9	0.919	3.534	64.505
10	0.818	3.147	67.652

In (Field, 2009, S. 662) wird für die Verwendung des Kaisers-Kriteriums als Kriterium an den Bestand eine Grösse von mindestens 250 Probanden und eine durchschnittliche Kommunalität von mindestens .6 genannt. Dieses Kriterium ist bei diesem Bestand mit N=1'084 und einer durchschnittlichen Kommunalität von .61 erfüllt. Die einzelnen Kommunalitäten sind im Anhang 8.4.2 wiedergegeben.

Obwohl mit diesen Angaben und diesem Kriterium auf acht Faktoren geschlossen werden könnte, wurde der Bestand einer zusätzlichen Parallelanalyse unterzogen. Bei dieser wird der Bestand verglichen mit einer «parallelen» Analyse mit Zufallsdaten, um zu eruieren, wie viele

der Faktoren auch durch Zufallseffekte zustande gekommen sein könnten. Auf diese Art lässt sich die Anzahl sinnvoll zu extrahierender Faktoren reduzieren.

Die Parallelanalyse wurde in der Software-Umgebung R mit Hilfe des R-packages «psych» (Revelle, 2020) durchgeführt. Mit dieser Parallelanalyse konnte die Anzahl Faktoren von ursprünglich acht Faktoren nach dem Kaiser-Kriterium auf 5 Faktoren reduziert werden. Die Ergebnisse der Parallelanalyse können als Indiz dafür betrachtet werden, dass bei den zusätzlichen 3 Faktoren auch Zufallseffekte für die Höhe der Eigenwerte als Erklärung herangezogen werden können. Mit der Wahl dieser 5 Faktoren wird ein Varianzanteil von 48.5% erklärt, was auch der Darstellung in Tabelle 13 entnommen werden kann.

Ergebnisse der Faktoranalyse / Ladungen auf den einzelnen Faktoren nach Rotation

Mit den verbleibenden fünf Faktoren wurde sodann die Rotation gerechnet, was zu den Faktorladungen in Tabelle 14 geführt hat.

Für die weitere Besprechung werden die fünf Faktoren so benannt, dass die Interpretation einfacher wird (siehe spätere Ausführungen). Die Namen für die fünf Faktoren lauten «Sachsituationen zur Multiplikation» (Faktor 1), «Operation erster Stufe» (Faktor 2), «Punktebilder und Struktur» (Faktor 3), «Aufmerksamkeit» (Faktor 4), «Abzählen» (Faktor 5).

Auch die Doppelladungen bei den Items 4b, 7b, 10a und 10b können mit dieser Namensgebung und der entsprechenden Interpretation sinnvoll gedeutet werden.

Es fällt auf, dass die Items in Gruppen auf die jeweiligen Faktoren laden. Bemerkenswert ist dabei, dass die ersten vier Items, bei denen noch keine multiplikative Struktur vorliegt und die ausschliesslich die Operationen erster Stufe bedienen, auf den Faktor 2 («Operationen erster Stufe») laden, auf welche die meisten anderen Items nicht laden. Eine Ausnahme bilden die verhältnismässig leichten Items 9a und 9b (Ausgangssituation «Vasen»).

Das Item 5 («Schokolade») lädt als (beinahe) einziges Item auf den Faktor 5 («Abzählen»). Bei diesem Item konnte während der Testdurchführung beobachtet werden, dass es von verschiedenen Kindern zählend durch Abzählen (mit dem Stift oder mit dem Finger) der einzelnen Stücke der Schokolade bearbeitet wurde. Dies stellt einen Vorgang dar, welcher bei vielen anderen Items weniger gut möglich war. Allenfalls noch bei Item 7a (Eier; Frage nach der Anzahl), bei dem ebenfalls eine Ladung auf dem Faktor 5 beobachtet werden kann.

Die Items 6 und 10 laden auf einen weiteren Faktor 4 («Aufmerksamkeit»). Bei diesen beiden Items war jeweils ein hohes Mass an Aufmerksamkeit notwendig, um das Item richtig zu lösen. Bei Item 6 war es notwendig, auch die Frage «Wie viele Bücher haben sie *zusammen*? » bis am Ende aufmerksam mit zu verfolgen. Einige Kinder gaben die Anzahl Bücher von Manuel an und beantworteten Item 6a mit «8» und 6b mit «4+4». Eine vergleichbare Schwierigkeit gab es bei Item 10, bei welchem es notwendig war, auch den Anfang des Satzes zu erfassen. Viele Kinder bearbeiteten dieses Item mit «1+1+1», dies möglicherweise durch stärkeren Fokus auf die optische Präsentation des Items als auf die sprachliche. Durch die Optik konnte bei diesem Item aber nur ein Teil der Aufgabenstellung erfasst werden. Ohne die vier Gänge

in den Keller zu erfassen, konnte dieses Item nicht richtig bearbeitet werden. Auch die Ladung des Items 4 auf diesen Faktor kann durch ähnliche Effekte erklärt werden. Es war bei dem Item mit den Autos notwendig, die Namen der beiden Mädchen genau auseinanderzuhalten.

Auffallend ist ferner die Ladung der Items 8 bis 15 auf denselben Faktor 1 («Sachsituation zur Multiplikation»). Dies ist insofern bemerkenswert als dass bei einigen dieser Items nach einem Term, bei anderen aber lediglich nach einer Zahl gefragt wird.

Die Items 16 bis 19 wiederum laden als einzige auf einen gemeinsamen Faktor 3 («Punktefelder/Struktur»). Diese vier Items bildeten den Abschluss des Tests und wurden wie auch die Items 11 bis 15 von den Kindern selbständig gelöst. Beobachtet wurde dabei eine Häufung von Kinderarbeiten, welche auf die Anzahl Punkte fokussierte und nicht eine passende Rechnung präsentierten, sondern die (häufig korrekte) Anzahl der Punkte (z.B. 18 bei Item 16, 30 bei Item 17 usw.), was aber zur Bewertung 0 führte. Allenfalls kann dies durch Ermüdungseffekte gegen Ende der Testdurchführung erklärt werden. Eine andere Erklärung wäre der Distraktionseffekt, der sich einstellt, nachdem in fünf Items (11 bis 15) nach einer Zahl gefragt wurde, kombiniert mit einem Überfliegen oder einer zumindest nicht sehr aufmerksamen Lektüre der Leitfrage «Finde eine passende Rechnung.» Bemerkenswert auch die Tatsache, dass alle vier Items auf denselben Faktor laden, obwohl bei den Items 16 und 17 die Bewertung «strenger» war und die Struktur des Terms zum Würfelbild passen musste, was bei den letzten beiden Items 18 und 19 nicht mehr der Fall war.

Auch denkbar ist es, dass in dieser Faktoranalyse die verschiedenen theoretisch verankerten Facetten des Konstrukts Operationsverständnis sichtbar werden und die vier Items 16 bis 19 sich deshalb von den anderen unterscheiden, weil nur bei ihnen die Strukturierungsfähigkeiten der Kinder erhoben werden. Die Analyse des Nachtests (siehe Kapitel 5.3.2) wird diese Möglichkeit noch einmal wahrscheinlicher werden lassen, daher die Wahl des Namens «Punktefelder/Struktur» für den Faktor 3.

Die Doppelladung bei Item 4b (Rechnung zur Aufgabe mit den Autos) kann erklärt werden dadurch, dass es sich gleichzeitig um ein Item zu einer Grundoperation erster Stufe (der Subtraktion) handelt sowie auch um ein Item, welches besondere (sprachliche) Aufmerksamkeit verlangt. Die Doppelladung bei Item 7b (Rechnung zur Aufgabe mit den Eiern) kann auf sinngemäss gleiche Art und Weise gedeutet werden. Die Items 10a und 10b (zur Ausgangssituation «Korb») zeigen eine Doppelladung aufgrund der Tatsache, dass es sich um eine Ausgangssituation zur zeitlich-sukzessiven Grundvorstellung handelt. Die Schwierigkeit, die sich den Kindern dabei stellt, ist diejenige, dass sie sich die Geschichte vor Augen führen müssen ohne ein «Endbild» räumlich-simultaner Art vor sich zu sehen. Insofern ist die Ladung auf den Faktor 4 («Aufmerksamkeit») nebst der Ladung auf den Faktor 1 («Sachsituation zur Multiplikation») verständlich.

Tabelle 14: Faktorladungen (>.3) für den Vortest (mit 5 Faktoren; nach Rotation)

Item	Faktoren				
	1 Sachsituationen Multiplikation	2 Operationen erster Stufe	3 Punktebilder & Struktur	4 Aufmerksamkeit	5 Lösen durch Abzählen
1		.520			
2		.668			
3		.741			
4a		.436			
4b		.531		.311	
5a					.834
5b					.744
6a				.797	
6b				.818	
7a					.439
7b		.301		.344	
8a	.506				
8b	.447				
9a	.395	.363			
9b	.362	.450			
10a	.493			.527	
10b	.430			.546	
11	.541				
12	.574				
13	.544				
14	.625				
15	.641				
16			.647		
17			.724		
18			.758		
19			.726		

Anmerkungen. Extraktionsmethode: Hauptkomponentenanalyse. Rotationsmethode: Varimax mit Kaiser-Normalisierung.

5.3.2 Faktoranalyse für den Nachtest

Der Nachtest wurde mithilfe der Bearbeitungen der 738 Kinder im Frühjahr 2020 auf dieselbe Art und Weise wie der Vortest einer exploratorischen Faktoranalyse unterzogen.

Voraussetzungen für eine explorative Faktoranalyse

Alle Items weisen positive Korrelationen zueinander auf (siehe Anhang 8.4.1), auch wenn die Werte dieser Korrelationen zum Teil nicht sehr hoch sind. Viele Inter-Item-Korrelationswerte liegen im Bereich von etwa .2, was leicht unterhalb des wünschenswerten Niveaus von etwa .3 liegt. Einzelne Items korrelieren aber auch mit einem Wert von .3 und höher, Werte unter .15 sind selten. Die Inter-Item-Korrelationstabelle sollte also einer explorativen Faktoranalyse nicht im Wege stehen.

Die weiteren Voraussetzungen für die exploratorische Faktoranalyse in Form eines Kaiser-Meyer-Olkin-Kennwerts über .6 und einem signifikant verwerfenden Bartlett-Test sind gegeben. Das Mass für die Stichprobeneignung nach Kaiser-Meyer-Olkin beträgt .867, der Bartlett-Test auf Sphärizität verwirft (χ^2(df=325) = 6'187.7, p<.001).

Wahl der Anzahl Faktoren

Wie in Kapitel 5.3.1 erfolgt an dieser Stelle die Angabe der Eigenwerte sowohl in tabellarischer Form als auch in Form des Screeplots (siehe Anhang 8.4.3).

Der erste Knick im Screeplot ergibt sich nach dem ersten, ein zweiter leichter Knick zeigt sich nach dem zweiten Eigenwert.

Beim Nachtest beträgt der Mittelwert der Kommunalitäten .58, was knapp unterhalb des gewünschten Werts von .6 liegt. Die Anzahl Kinder, welche den Test abgelegt haben, liegt mit 738 verhältnismässig hoch im Vergleich zur Zielgrösse von 250 (siehe Kapitel 5.3.1).

Wie auch für den Vortest wurde die Anzahl Faktoren reduziert durch einen Vergleich mit einem Paralleltest, der mit der Software R gerechnet wurde. Dabei hat sich gezeigt, dass zwei der sieben Faktoren (Anzahl Eigenwerte grösser als 1) durch Zufallseffekte zustande gekommen sein könnten. Der Vergleich rechtfertigt eine Reduktion der Anzahl zu wählender Faktoren von 7 auf 5. Gemäss Tabelle 15 wird mit dieser Wahl 49.0% der Varianz erklärt.

Ergebnisse der Faktoranalyse / Ladungen auf den einzelnen Faktoren nach Rotation

Aufgrund der obigen Überlegungen wurde die Anzahl von 5 Faktoren gewählt. Das gemessene Konstrukt zerfällt wie schon beim Vortest. Wiederum sind in Tabelle 16 diejenigen Faktorladungen dargestellt, welche den Wert .3 übersteigen.

Tabelle 15: Die grössten zehn Eigenwerte für den Nachtest, von denen sieben über 1 liegen.

Faktor	Eigenwert	% der Varianz	Kumulierte %
1	6.946	26.717	26.717
2	1.752	6.739	33.456
3	1.422	5.470	38.925
4	1.378	5.299	44.225
5	1.248	4.801	49.026
6	1.141	4.388	53.415
7	1.064	4.092	57.507
8	0.951	3.658	61.164
9	0.924	3.555	64.720
10	0.920	3.537	68.257

Es fällt auf wie – in Analogie zum Vortest – die ersten zwei Items 1a und 1b zur Ausgangssituation «Autos» sowie Item 7 zur Ausgangssituation «Stifte» auf einen anderen Faktor laden als sämtliche anderen Items. Diese drei Items sind diejenigen, in denen Sachsituationen zu Grundrechenarten erster Stufe beschrieben werden. Die Items mit multiplikativer Struktur hingegen laden alle hauptsächlich auf die Faktoren 1 bis 4.

Die Brückenitems 3a und 3b («Bücher) laden als einzige auf den Faktor 4.

Es wiederholt sich die Feststellung aus dem Vortest, dass die Items zu den Punktefeldern (Items 15 bis 18) auf einen separaten Faktor laden. Auch hier könnte dies mit dem leicht andersartigen Itemformat (schriftliche ausdrückliche Frage nach einer *Malrechnung*) oder auch damit erklärt werden, dass diese vier Items die Facette der Strukturierungsfähigkeit der Kinder bedient.

Auch im Nachtest zeichnet sich das Item zur Ausgangssituation «Korb und Flaschen» (zur zeitlich-sukzessiven Grundvorstellung der Multiplikation) dadurch aus, dass es sich in den Bearbeitungen der Kinder von der Mehrheit der sprachlich und optisch präsentierten Ausgangssituationen abgrenzt. Interessanterweise lädt es im Nachtest auf denselben Faktor wie die Items zu 11 bis 13 («Beeren» und «Füsse») sowie die beiden Multiple-Choice Items 19 und 20, welche sich trotz andersartigem Itemformat verhältnismässig unauffällig einreihen. Mit den beiden Multiple-Choice-Items scheint es gelungen zu sein, ein für die Kinder etwas anspruchsvolleres Itemformat zum selben Konstrukt zu entwerfen.

In Analogie zum Vortest werden deshalb die fünf Faktoren mit Namen versehen. Für den Nachtest lauten diese fünf Namen «Sachsituationen zur Multiplikation I» (Faktor 1), «Punktebilder und Struktur» (Faktor 2), «Sachsituationen zur Multiplikation II» (Faktor 3), «Bü-

cher» (Faktor 4) und «Operationen erster Stufe» (Faktor 5). Im Gegensatz zum Vortest lassen sich hier die beiden Faktoren 1 und 3 (beide zu Sachsituationen zur Multiplikation» und die Aufteilung der Ladungen auf diese beiden Faktoren nicht gänzlich erklären. Dazu passt auch die Beobachtung, dass drei der vier auftretenden Doppelladungen zwischen diesen beiden Faktoren auftreten. Die vierte Doppelladung ist diejenige bei Item 7 («Stifte»). Dieses Item lädt erwartungsgemäss auf den Faktor für die Operationen erster Stufe. Gleichzeitig und mit einem tieferen Wert lädt es auf den Faktor, der von den «Sachsituationen zur Multiplikation I» gebildet wird. Die Faktorladung von .308 für diesen Effekt liegt sehr nahe am Grenzwert von .3 und wird in dieser Arbeit nicht weiter interpretiert.

Tabelle 16: Faktorladungen (>.3) für den Nachtest (mit 5 Faktoren; nach Rotation)

	Faktoren				
Item	1 Sachsituationen Multiplikation I	2 Punktebilder & Struktur	3 Sachsituationen Multiplikation II	4 Bücher	5 Operationen erster Stufe
1a					.821
1b					.827
2	.329				
3a				.895	
3b				.901	
4	.461				
5a	.420		.426		
5b	.505		.330		
6	.638				
7	.308				.385
8a	.326				
8b	.667				
9a	.644				
9b	.707				
10a	.531				
10b	.661				
11			.499		
12			.509		
13	.470		.311		
14	.432				
15		.642			
16		.757			
17		.718			
18		.771			
19			.671		
20			.683		

Anmerkungen. Extraktionsmethode: Hauptkomponentenanalyse. Rotationsmethode: Varimax mit Kaiser-Normalisierung.

5.4 Analyse mittels IRT (Vortest)

In diesem Abschnitt erfolgt eine Analyse der Testerhebung für den Vortest mittels Item Response Theorie, insbesondere mit dem Rasch-Modell. Wie bereits erwähnt ist das Rasch-Modell bereits bei der Entwicklung dieses Erhebungsinstruments als Auswertungsmodell für die Erhebung festgelegt worden. Items wurden aufgrund ihrer Passung ins Rasch-Modell ausgewählt.

Dargelegt werden Item-Schwierigkeiten und die Personenfähigkeiten, die sich aus der Verwendung des Rasch-Modells ergeben, sowie auch die Reliabilitäten der IRT im Rasch-Modell. Auch werden die Fit-Werte der einzelnen Items berichtet.

Sodann werden die Items im Hinblick auf die Kriterien Geschlecht und Sprachverständnis der Kinder einer «Differential Item Functioning»-Analyse unterzogen. Die Erkenntnisse dieser Analyse können von Interesse sein für eine weitere Verwendung und allfällige Verbesserung des Erhebungsinstruments.

Für die Berechnung der einzelnen Parameter kamen unterschiedliche Software-Packages für die Software R zum Einsatz, insbesondere die beiden packages TAM (Robitzsch et al., 2020) und eRm (Koller et al., 2012; Mair et al., 2020).

5.4.1 Parameter des Rasch-Modells

Die Item-Schwierigkeiten β für die einzelnen Items sowie die Personenfähigkeiten *p* für jedes einzelne Kind wurden mit dem R-package eRm gerechnet.

Item-Schwierigkeiten

Die Item-Schwierigkeiten stehen natürlich in starkem Zusammenhang mit den Item-Schwierigkeits-Werten aus der klassischen Testtheorie. Dies insbesondere darum, weil es sich bei den Item-Schwierigkeiten im Rasch-Modell um erschöpfende (suffiziente) Statistiken handelt (Moosbrugger & Kelava, 2012, S. 243) und es daher nur entscheidend ist, wie viele Kinder − und nicht etwa *welche* Kinder − das jeweilige Item korrekt beantwortet haben. Auch in der klassischen Testtheorie erfolgt die Berechnung der jeweiligen Item-Schwierigkeit mittels Anteil korrekt bearbeiteter (mit 1 codierter) Bearbeitungen. Die Item-Schwierigkeiten für die einzelnen 26 Items finden sich in Tabelle 17. Das Vorzeichen der Item-Schwierigkeiten ist im Rasch-Modell (im Gegensatz zur klassischen Testtheorie) so, dass schwierige Items einen hohen Item-Schwierigkeitswert aufweisen. Das leichte Einstiegsitem 1 macht sich daher in Tabelle 17 mit einem tiefen (negativen) Wert bemerkbar.

Tabelle 17: Item-Schwierigkeiten β im Rasch-Modell (gerechnet mit eRm)

Item	Item-Schwierigkeit β im Rasch-Modell	Item	Item-Schwierigkeit β im Rasch-Modell
1	-2.17	9a	-1.20
2	-0.82	9b	-0.37
3	-1.40	10a	1.12
4a	-0.25	10b	1.67
4b	-0.12	11	-0.38
5a	-0.83	12	0.49
5b	0.07	13	1.69
6a	1.03	14	-0.55
6b	1.20	15	1.51
7a	-0.44	16	-0.35
7b	0.46	17	-0.57
8a	0.66	18	-1.13
8b	1.44	19	-0.74

Personenfähigkeit

Auch die Personenfähigkeit hängt lediglich davon ab, wie viele Items das jeweilige Kind korrekt bearbeitet hat. Es ergeben sich daher bei einem Erhebungsinstrument mit 26 Items und bei einem möglichen Testscore von 0, 1, 2 usw. bis 26 genau 27 Werte für die Personenfähigkeit. Bei Kindern, welche keine oder aber alle Items richtig bearbeitet haben, kann theoretisch keine Personenfähigkeit angegeben werden. Die untenstehenden Zahlen von -4.50 bzw. + 4.52 für diese beiden Personengruppen sind daher als reine Modellwerte zu lesen. Für die übrigen Werte kann, wie in Kapitel 4.5.2 beschrieben, im Zusammenspiel mit der Item-Schwierigkeit die Lösungswahrscheinlichkeit des jeweiligen Kindes für das jeweilige Item eruiert werden. Die Personenfähigkeiten sind in Tabelle 18 wiedergegeben.

Tabelle 18: Personenfähigkeits-Parameter im Rasch-Modell (gerechnet mit eRm)

Anzahl richtig gelöster Items	Personen-fähigkeits-Parameter im Rasch-Modell	Anzahl richtig gelöster Items	Personen-fähigkeits-Parameter im Rasch-Modell	Anzahl richtig gelöster Items	Personen-fähigkeits-Parameter im Rasch-Modell
0	-4.50	9	-0.79	18	0.99
1	-3.66	10	-0.59	19	1.22
2	-2.89	11	-0.39	20	1.46
3	-2.40	12	-0.20	21	1.74
4	-2.03	13	-0.01	22	2.05
5	-1.73	14	0.18	23	2.42
6	-1.46	15	0.37	24	2.90
7	-1.22	16	0.57	25	3.68
8	-1.00	17	0.78	26	4.52

Reliabilität im Rasch-Modell

Im Rasch-Modell können verschiedene Reliabilitätsmasse verwendet werden. Diesen Massen liegen ähnliche Konzepte zugrunde wie Cronbachs α in der klassischen Testtheorie. Ziel ist es, die Varianz der Personenleistung zu beurteilen, welche nicht durch Zufall erklärt werden kann.

Anhand der Bestandsdaten des Vortests wurden die Separation Reliabilität (mit dem package eRm) sowie die EAP-Reliabilität (Expected a Posteriori) und die WLE-Reliabilität (Warm's Weighted Likelihood Estimator) berechnet. Es ergaben sich dabei die Werte, die in Tabelle 19 dargestellt sind.

Tabelle 19: Reliabilitätsmasse im Raschmodell und Werte für den Vortest Multiplikationsverständnis

Reliabilitätsmass	Wert	R-Package	R-Funktion
Person Separation Reliability	.88	eRm	«SepRel()»
EAP Reliability	.88	TAM	«tam.mml()»
WLE Reliability	.85	TAM	«tam.wle()»

Fit-Werte für das Rasch-Modell

Für eine ausreichende Passung ins Rasch-Modell wird wie in Kapitel 4.5.2 beschrieben ein MSW-Infit-Wert für das jeweilige Item in der Nähe von 1 erwartet. Eine Bandbreite zwischen 0.7 und 1.3 ist dabei akzeptabel, eine solche von 0.8 bis 1.2 wünschenswert.

Die Schätzung dieses Werts hängt ab vom jeweils eingesetzten Algorithmus, der sich zwischen den verschiedenen R-Packages unterscheidet. An dieser Stelle werden deshalb die Ergebnisse beider R-Packages TAM und eRm widergegeben:

Package eRm:

Das einzige Item ausserhalb des Korridors von 0.8 bis 1.2 ist das Item 5a mit einem MSQ-Infit-Wert von 1.204. Alle weiteren Items liegen innerhalb dieses Korridors. Insgesamt liegen 20 der 26 Items sogar innerhalb des Korridors von 0.9 bis 1.1, womit der Item-Fit für das Rasch-Modell als sehr gut bezeichnet werden kann.

Package TAM:

Es ergibt sich ein ähnliches Bild. Das «kritische» Item 5a weist einen Wert von 1.191, zwei weitere Items zeigen eine Abweichung von etwas mehr als 0.1 vom gewünschten Wert 1 und 23 Items liegen im (engen) Korridor zwischen 0.9 und 1.1. Die Abweichungen vom Wert 1 für denselben Bestand fallen mit dem Package TAM etwas geringer aus als mit dem Package eRm. Insgesamt kann der ausgezeichnete Model-Fit für das Rasch-Modell hiermit bestätigt werden.

5.4.2 DIF-Analysen für die einzelnen Items

In diesem Kapitel werden DIF-Analysen, insbesondere für die beiden Merkmale «Geschlecht» und «Sprachverständnis», anhand der Vortest-Ergebnisse vorgenommen. Es werden dafür verschiedene Methoden zur Ortung von DIF eingesetzt.

Andersen Likelihood-Ratio-Test

Geschlecht:

Teilt man den Bestand der 1'084 Kinder in die beiden Geschlechter männlich/weiblich, so verwirft der Andersen Likelihood-Ratio-Test (LR-Wert = 64.95, df=25, $p<.01$). Das Antwortverhalten unterscheidet sich also nach Geschlecht.

Sprachverständnis:

Vom Bestand der 1'084 Kinder lag bei 1'073 Kindern eine Einschätzung der Lehrperson betreffend Sprachverständnis der Kinder vor. Die Lehrpersonen wurden aufgefordert, ihre Einschätzung auf einer vierstufigen Skala von 0 bis 3 vorzunehmen. Insgesamt haben die Lehrpersonen 816 Kinder mit dem Maximalwert 3 eingeschätzt. Bei insgesamt 257 Kindern wurde eine 0 (6mal), eine 1 (64mal) oder eine 2 (187mal) als Einschätzung vorgenommen. Bei 11

Kindern fehlte die Einschätzung. Mehrheitlich handelte es sich dabei um Neuzuzüge, bei denen auch auf Nachfrage hin die Einschätzung durch die Lehrperson nicht nachgeliefert worden war. Bei den Analysen im Hinblick auf das Sprachverständnis wurde der Bestand der 1'073 eingeschätzten Kinder verwendet. Das Split-Kriterium für den Bestand war die Einschätzung mit dem Maximalwert (Ja/Nein).

Auch im Hinblick auf das Sprachverständnis verwirft der Andersen Likelihood-Ratio-Test (LR-Wert 112.55, df=25, p<.01). Das Antwortverhalten unterscheidet sich folglich nach Sprachverständnis.

Wald-Test

Mit dem im Folgenden durchgeführten Wald-Test wird nun für beide Split-Kriterien das Verhalten der einzelnen Items genauer untersucht.

Geschlecht:

Der Wald-Test ortet DIF für das Split-Kriterium Geschlecht bei einzelnen Items. Auf dem 5%-Niveau signifikant wird das DIF bei den Items 4b (p=.015), 6a (p<.001), 6b (p=.028), 7a (p=.002) und 13 (p=.002).

Sprachverständnis:

Der Wald-Test ortet DIF auch für das Split-Kriterium Sprachverständnis bei einzelnen Items. Auf dem 5%-Niveau signifikant wird das DIF bei den Items 1 (p=.016), 2 (p=.044), 5a (p=.020), 6a (p=.002), 6b (p=.001), 7a (p=.022), 9a (p<.001), 10a (p<.001) und 10b (p<.001) und 16 (p=.042).

Graphische Überprüfung des Item-Fits

Die graphische Illustration der Items mit DIF findet sich für beide Kriterien (Geschlecht und Sprachverständnis) im Anhang 8.4.4.

Mantel-Haenszel Verfahren

Geschlecht:

Mit dem Mantel-Haenszel-Verfahren wird für 23 der 26 Items kein DIF oder ein solches im Ausmass A (Typ A, «Negligible») ausgewiesen. Die drei Items 6a (p=.001), 7a (p=.005) und 13 (p=.006) zeigen DIF im Ausmass B. Kein Item weist DIF im Ausmass C auf.

Sprachverständnis:

Mit dem Mantel-Haenszel-Verfahren wird für 20 der 26 Items kein DIF oder ein solches im Ausmass A (Typ A, «Negligible») ausgewiesen. Die drei Items 3 (p=.068), 6a (p=.016) und 6b (p=.006) zeigen DIF im Ausmass B. Die drei Items 9a (p<.001), 10a (p<.001) und 10b

(p=.007) zeigen DIF im Ausmass C. Für die Berechnungen wurde die Funktion «difMH» aus dem R-Package difR verwendet (Magis et al., 2018).

Weitere Verfahren

Eine Klassierung nach Ausmass des DIF könnte auch vorgenommen werden, indem ausgezählt wird, bei wie vielen Methoden zur Ortung von DIF die einzelnen Items DIF zeigen. Die Funktion dichoDif aus dem R-Package difR erlaubt eine solche Analyse. Sie wurde mit den vier Verfahren «T.I.D», «MH», «Std.» und «SIBTEST» (Magis et al., 2018, S. 9 ff.) auf die beiden Split-Kriterien Geschlecht und Sprachverständnis angewendet.

Geschlecht:

Bei dieser Analyse werden die Items 6a und 13 auffällig, da bei diesen bei jeweils zwei der vier angewandten Methoden DIF geortet wird.

Sprachverständnis:

Bei dieser Analyse werden die Items 6a, 6b, 8b, 9a und 10a auffällig. Item 9a weist bei drei der vier Methoden DIF auf. Die anderen vier der genannten Items liefern DIF bei zwei der vier Methoden.

Martin Loef-Test

Für den Martin Loef-Test werden die Items anstelle der Personen geteilt. Es wird mit den folgenden beiden Split-Kriterien für die Items gearbeitet:

Präsentationsform:

Die ersten 17 Items wurden den Kindern vorgelesen, die letzten 9 mussten sie selbständig bearbeiten. Während es bei den vorgelesenen Items denkbar wäre, dass die Bearbeitung am Sprachverständnis oder an den auditiven Speicherfähigkeiten der Kinder scheitert, sind für die späteren Items die Lesefähigkeit und die frühe präzise Aufnahme der Instruktionen für die Bearbeitung wesentlich. Der Bestand wird darum für diese Analyse geteilt in die ersten 17 Items und die späteren 9 Items.

Frageform:

Bei einigen Items wird nach einem Ergebnis (einer Zahl), bei anderen nach einem Term gefragt. Verschiedene Ausgangssituationen umfassen beides, nämlich je ein Item mit einer Frage nach einem Ergebnis und ein Item mit einer Frage nach einer Zahl. Denkbar wäre es, dass die jeweiligen Fragen unterschiedlich bearbeitet werden. Dies insbesondere auch darum, weil bei der theoretischen Beschreibung des Konstrukts die Frage nach dem Term zentral ist, das Lösen von Multiplikationsaufgaben und somit also die Frage nach dem Ergebnis der jeweiligen Rechnung aber weitaus geringere Bedeutung hat. Der Bestand wird darum geteilt in die Items mit einer Frage nach einer Zahl (4a, 5a, 6a, 7a, 8a, 9a, 10a, 11, 12, 13, 14, 15) und die anderen Items, bei denen nach einem passenden Term (bzw. in der Sprache der Kinder «nach

einer passenden Rechnung») gefragt wurde. Die Analyse erfolgt mit der Funktion «MLoef» des R-packages eRm. Beide Analysen erfolgten mit dem vollen Bestand der 1'084 Kinder.

Präsentationsform:

Der Martin Loef-Test verwirft die Nullhypothese, dass die beiden Testteile gleich beantwortet werden (LR-Wert 352.2, df=152, p<.001).

Frageform:

Der Martin Loef-Test verwirft die Nullhypothese, dass die beiden Testteile gleich beantwortet werden (LR-Wert 259.8, df=167, p<.001).

5.5 Analyse mittels IRT (Nachtest)

Der Übersichtlichkeit halber werden die Analysen mittels Item Response Theorie für den Nachtest in diesem separaten Kapitel dargestellt. Die Analysemethoden und das Vorgehen sind dabei diejenigen, welche auch im vorangehenden Kapitel für den Vortest gewählt wurden.

5.5.1 Parameter des Rasch-Modells

Wie auch beim Vortest wurde die Item-Schwierigkeiten β für die einzelnen Items sowie die Personenfähigkeiten *p* für jedes einzelne Kind mit dem R-package eRm gerechnet.

Item-Schwierigkeiten

Die Item-Schwierigkeiten für die insgesamt 26 Items des Nachtests sind Tabelle 20 zu entnehmen.

Tabelle 20: Item-Schwierigkeiten β im Rasch-Modell (gerechnet mit TAM)

Item	Item-Schwierigkeit β im Rasch-Modell	Item	Item-Schwierigkeit β im Rasch-Modell
1a	-2.96	9b	-0.97
1b	-2.18	10a	-0.30
2	1.24	10b	-0.49
3a	-1.12	11	-3.15
3b	-1.12	12	-1.03
4	0.38	13	-1.02
5a	-1.14	14	-1.31
5b	-1.08	15	-1.67
6	-1.36	16	-2.37
7	-2.61	17	-1.21
8a	-1.55	18	-1.86
8b	-1.22	19	-0.29
9a	-0.64	20	-0.13

Personenfähigkeit

Die Personenfähigkeitsparameter können der Tabelle 21 entnommen werden.

Tabelle 21: Personenfähigkeits-Parameter im Rasch-Modell (gerechnet mit eRm)

Anzahl richtig gelöster Items	Personen-fähigkeits-Parameter im Rasch-Modell	Anzahl richtig gelöster Items	Personen-fähigkeits-Parameter im Rasch-Modell	Anzahl richtig gelöster Items	Personen-fähigkeits-Parameter im Rasch-Modell
0	-4.48	9	-0.75	18	0.96
1	-3.63	10	-0.56	19	1.18
2	-2.85	11	-0.37	20	1.41
3	-2.37	12	-0.18	21	1.68
4	-1.99	13	-0.00	22	1.99
5	-1.69	14	0.18	23	2.36
6	-1.42	15	0.37	24	2.86
7	-1.18	16	0.56	25	3.65
8	-0.96	17	0.75	26	4.51

Reliabilität im Rasch-Modell

Für den Nachtest können die Reliabilitäten im Rasch-Modell der Tabelle 22 entnommen werden.

Tabelle 22: Reliabilitätsmasse im Raschmodell und Werte für den Nachtest Multiplikationsverständnis

Reliabilitätsmass	Wert	R-Package	R-Funktion
Person Separation Reliability	.85	eRm	«SepRel()»
EAP Reliability	.86	TAM	«tam.mml()»
WLE Reliability	.82	TAM	«tam.wle()»

Fit-Werte für das Rasch-Modell

Package eRm:

Alle Items liegen innerhalb dieses Korridors von 0.8 bis 1.2. Insgesamt liegen 19 der 26 Items sogar innerhalb des Korridors von 0.9 bis 1.1. Für vier Items berechnet der Algorithmus Werte zwischen 0.8 und 0.9. Für drei Items liefert er Werte zwischen 1.1 und 1.2.

Package TAM:

Alle Items liegen innerhalb dieses Korridors von 0.8 bis 1.2. Insgesamt liegen 22 der 26 Items sogar innerhalb des Korridors von 0.9 bis 1.1. Die Item-Fit-Werte für das Rasch-Modell können als sehr gut bezeichnet werden.

5.5.2 DIF-Analysen für die einzelnen Items

Das Geschlecht war für alle 738 Kinder vorhanden. Die Strukturvariable des Sprachverständnisses war bei 11 der 738 Kinder fehlend, so dass diese Analyse lediglich mit den 727 Kindern durchgeführt werden konnte, für welche die Variable vorlag. Alle Rechnungen unten wurden mit 738 Kindern (für das Geschlecht) bzw. 727 Kindern (für das Sprachverständnis) durchgeführt. Die Verwendung der Strukturvariable Sprachverständnis erfolgte wie beim Vortest (siehe Kapitel 5.4.2).

Andersen Likelihood-Ratio-Test

Geschlecht:

Der Andersen Likelihood-Ratio-Test verwirft die Nullhypothese, dass die Bearbeitung durch die Kinder nicht vom Geschlecht abhängt (LR-Wert = 82.65, df=25, p<.001).

Sprachverständnis:

Der Andersen Likelihood-Ratio-Test verwirft die Nullhypothese, dass die Bearbeitung durch die Kinder nicht vom Sprachverständnis abhängt (LR-Wert = 6.11, df=25, p<.001).

Wald-Test

Geschlecht:

Der Wald-Test ortet DIF für das Split-Kriterium Geschlecht bei verschiedenen Items. Auf dem 5%-Niveau signifikant wird das DIF bei den Items 3a (p=.043), 3b (p=.043), 6 (p=.021), 9a (p=.047), 10a (p=.043), 13 (p=.033) sowie 15 (p=.028), 16 (p<.001), 17 (p=.016) und 18 (p=.001).

Sprachverständnis:

Der Wald-Test ortet auch DIF für das Split-Kriterium Sprachverständnis bei verschiedenen Items. Auf dem 5%-Niveau signifikant wird das DIF bei den Items 5a (p=.012), 5b (p=.002), 6 (p=.014), 10a (p=.002), 15 (p=.042) und 20 (p=.035).

Graphische Überprüfung des Item-Fits

Die graphische Illustration der Items mit DIF findet sich für beide Kriterien (Geschlecht und Sprachverständnis) im Anhang 8.4.4.

In den Graphen im Anhang sind für den Nachtest die Items mit markantem DIF (gerechnet mit dem LR-Test) dargestellt. Für das Geschlecht weisen die Items 16, 17 und 18 (Punktefelder) deutlich sichtbares DIF auf, beim Sprachverständnis sticht insbesondere Item 10a (Frage nach der Anzahl Scheiben beim Item «Ananas auf den Pizzas») heraus.

Mantel-Haenszel Verfahren

Geschlecht:

Mit dem Mantel-Haenszel-Verfahren wird für 21 der 26 Items kein DIF oder ein solches im Ausmass A (Typ A, «Negligible») ausgewiesen. Die drei Items 6 (p=.038), 13 (p=.012) und 17 (p=.021) zeigen DIF im Ausmass B. Die zwei Items 16 (p<.001) und 18 (p=.003) weisen DIF im Ausmass C auf.

Sprachverständnis:

Mit dem Mantel-Haenszel-Verfahren wird für 20 der 26 Items kein DIF oder ein solches im Ausmass A (Typ A, «Negligible») ausgewiesen. Sechs Items zeigen DIF im Ausmass B, bei den folgenden vier dieser sechs Items sind die Werte signifikant: Item 5a (p=.041), 5b (p=.009), 9a (p=.018) und 10a (p=.009).

Für die Berechnungen verwendet wurde wie schon beim Vortest die Funktion «difMH» aus dem R-Package difR.

Weitere Verfahren

Wie bereits beim Vortest wird die Funktion dichoDif aus dem R-Package difR verwendet, um verschiedene Methoden zur Ortung von DIF für die beiden Split-Kriterien Geschlecht und Sprachverständnis einzusetzen. Die Analyse geschieht wiederum mit den vier Verfahren «T. I. D», «MH», «Std.» und «SIBTEST».

Geschlecht:

Alle vier Verfahren orten DIF für Item 16, drei der vier Verfahren orten DIF für Item 18 und zwei der vier Verfahren orten DIF für die Item 10a, 13 und 17. Bei den weiteren Items wird nur mit einem oder häufig auch mit keinem Verfahren DIF nachgewiesen.

Sprachverständnis:

Bei dieser Analyse werden die Items 5b, 9a und 10a auffällig. Für Item 5b und 10a weisen drei der vier Verfahren DIF nach. Für Item 9a tun dies zwei der vier Verfahren. Bei allen anderen Items sind es entweder ein oder kein Verfahren.

Martin Loef-Test

Die mit dem Martin Loef-Test durchgeführten Analysen betreffend Präsentation- und Frageform wurden mit dem Nachtest nicht wiederholt. Grund dafür ist die Erkenntnis, dass auch ein Verwerfen des Tests für die Weiterentwicklung des Erhebungsinstruments nicht verwertbar ist, da die Gründe für das Verwerfen unklar bleiben.

5.6 Zusammenhänge mit anderen an der Stichprobe erfassten Grössen (Vortest)

Anhand des Vortests wurden zusätzliche Zusammenhangsmasse mit anderen Grössen gerechnet, die im Rahmen des Projekts MALKA erhoben wurden. In diesem Kapitel erfolgt eine Übersicht über die Ergebnisse dieser Rechnungen. Die Rechnungen wurden lediglich für den Vortest vorgenommen, weil im Vortest die grössere Stichprobe zur Verfügung stand und dieser Bestand zum Testzeitpunkt T3 noch nicht durch die Covid-19-Situation im Frühjahr 2020 beeinflusst war. Bei diesem Testzeitpunkt ist es also wahrscheinlicher, dass die Ergebnisse unter regulären Umständen repliziert werden können. Der Testzeitpunkt T4, bei welchem der Nachtest durchgeführt wurde, lag so kurz nach Wiedereröffnung der Schulen, dass die Erhebungen (möglicherweise) auch andere Effekte widerspiegeln. Die Erfassung solcher «Covid-19-Effekte» war aber nicht das Ziel dieser Arbeit.

5.6.1 Korrelationen

Für viele der 1'084 Kinder, welche zu Beginn ihres zweiten Schuljahrs den Test für das Multiplikationsverständnis abgelegt haben, liegen aus dem ersten Schuljahr weitere Testergebnisse vor. In diesem Kapitel werden die Ergebnisse dieser vorgängigen Tests mit den Ergebnissen des Vortests für das Operationsverständnis in Zusammenhang gebracht. Mit solchen Zusammenhangsanalysen kann gleichzeitig der Test für das Operationsverständnis validiert werden, indem geprüft wird, ob die Ergebnisse mit anderen Ergebnissen korrelieren (wenn dies zu erwarten wäre, Stichwort: konvergente Validität) oder nicht bzw. nur schwach korrelieren (wenn dies zu erwarten wäre, Stichwort: diskriminante Validität). Verwendet werden dafür die Tests, die in Kapitel 4.1.3 beschrieben wurden. Als Mass für den Zusammenhang werden hauptsächlich Korrelationen eingesetzt. Für diesen Zweck werden die Kinder mit ihren jeweiligen Ergebnissen aus den einzelnen Erhebungsinstrumenten ohne Rücksicht auf ihre Klassenzugehörigkeit untersucht. Die Berücksichtigung der Klassenzugehörigkeit erfolgt dann mit Hilfe von Mehrebenenanalysen. Angesichts der eher tiefen Intraklassen-Korrelation (ICC), die dort berechnet und angegeben wird, erscheint dieses Vorgehen an dieser Stelle berechtigt.

Tabelle 23: Korrelationen des Vortests (OVpre) mit anderen Erhebungsinstrumenten

	Alter OVpre	BM 0+	BM 1+	CFT	AG	AE T1	KR T1	AE T2	KR T2	KR nicht-zählend	KR zählend	SFON
OVpre	-.02	.63**	.70**	.52**	.48**	.36**	.57**	.35**	.58**	.53**	-.09**	.18**
N	1'042	1'009	1'026	1'011	1'008	1'007	1'007	1'003	1'003	1'003	1'003	1'007

Anmerkungen. OV = Operationsverständnis. BM = BASIS-MATH CFT = Culture Fair Intelligence Test. AG = Arbeitsgedächtnis. AE = Anzahlerfassung. KR = Kopfrechnen. SFON = Spontaneous Focusing on Numerosity. *p < .05. **p < .01.

Im Folgenden werden verschiedene Werte aus Tabelle 23 noch einzeln kommentiert und in Zusammenhang mit anderen Erkenntnissen aus dieser Arbeit gestellt.

Alter zum Zeitpunkt des Vortests (Testzeitpunkt T3)

Von den 1'084 Kindern lagen bei 1'042 das genaue Geburtsdatum und damit das genaue Alter zum Testzeitpunkt T3 vor. Bei den anderen Kindern handelte es sich um Neuzugänge, bei denen die entsprechenden Daten nicht erfasst worden waren. Alter und Testergebnis im Test für das Operationsverständnis wiesen für die 1'042 Kinder eine äusserst geringe Korrelation von -.023 (p=.467) auf. Dieses Ergebnis ist insofern interessant, als dass der Lernstand der Kinder im Vortest nicht mit dem Alter in Zusammenhang stand, das Alter der Kinder sich dann aber für den *Lernzuwachs* im zweiten Jahr als ein hilfreicher Prädiktor entpuppte (siehe Kapitel 6.1.3).

Gesamtwert BASIS-MATH 0+ sowie BASIS-MATH 1+ (Testzeitpunkte T1, T2)

Von den 1'084 Kindern lag bei 1'009 der Gesamtwert des Erhebungsinstruments BASIS-MATH 0+ zum Testzeitpunkt T1 (rund ein Jahr vor T3) vor. Die beiden Instrumente weisen eine signifikante (p<.001) Korrelation von .627 auf. Bei 1'026 Kindern lag der Gesamtwert des Erhebungsinstruments BASIS-MATH 1+ zum Testzeitpunkt T2 (rund vier Monate vor T3) vor. Die beiden Instrumente weisen eine signifikante (p<.001) Korrelation von .698 auf. Bemerkenswert ist die Beobachtung, dass der unterschiedliche Testzeitpunkt und die unterschiedliche zeitliche Nähe zum Testzeitpunkt T3 sich nicht stärker bemerkbar machten.

Die weiteren Werte aus Tabelle 23 erscheinen plausibel und stützen die Überlegungen zur Validität des Erhebungsinstruments. Insbesondere kann es als Aspekt der *divergenten* Validität betrachtet werden, dass die zählend erbrachten Leistungen («KR zählend») der Kinder im MALKA-Test nur in geringem Masse (und mit negativem Vorzeichen) mit deren Leistung im Erhebungsinstrument für das Operationsverständnis korrelieren. Alle anderen Korrelationen können im Lichte der *konvergenten* Validität betrachtet werden und entsprechen sowohl hinsichtlich Vorzeichen als auch hinsichtlich Grössenordnung den Erwartungen für ein solches Erhebungsinstrument.

5.6.2 Einfluss der Einteilung in die Schulklassen

Anhand des vorliegenden Datensatzes kann untersucht werden, welcher Anteil der Gesamtvarianz einer abhängigen Variable durch die Zugehörigkeit zu einer Schulklasse bzw. im Gegensatz dazu durch die individuellen Unterschiedlichkeiten («Schwankungen») zwischen einzelnen Kindern erklärt werden kann.

Für diese Betrachtung eignet sich die Kennzahl der Intraklassen-Korrelation (ICC: «intraclass-correlation») wie sie in der Mehrebenenanalyse (siehe Kapitel 4.5.5 und Kapitel 6) verwendet wird (Eid et al., 2017, S. 730 ff.). Diese Kennzahl wird in einem Mehrebenen-

Nullmodell für unterschiedliche abhängigen Variablen bestimmt. Die Werte dieser Berechnung erfolgen an dieser Stelle – der Vergleichbarkeit halber − für den Bestand der Kinder, die alle fünf unten aufgeführten Erhebungen abgelegt haben. Der Kinderbestand umfasst 1'003 Kinder aus 68 Klassen. Die Werte sind von der Covid-19-Situation im Frühjahr 2020 nicht tangiert, da die drei Erhebungen T1, T2 und T3 alle früher stattgefunden haben.

Die ICC-Werte in diesem Kapitel brauchen nicht notwendigerweise übereinzustimmen mit den Werten, die für diese Kennzahl an anderer Stelle (insbesondere bei den Längsschnittanalysen) berechnet wurden. Grund dafür sind die unterschiedlichen Kinderbestände, die für die verschiedenen Betrachtungen herangezogen werden.

Tabelle 24: ICC-Werte für verschiedene Erhebungen

Erhebung	Testzeitpunkt	Intraklassen-Korrelation (ICC)
Arbeitsgedächtnis	T1	10.3%
Kognitive Fähigkeiten (CFT)	T1	7.5%
BASIS-MATH 0+	T1	20.9%
BASIS-MATH 1+	T2	12.5%
Vortest Multiplikationsverständnis	T3	11.1%

5.6.3 Einfluss des Lehrmittels

Die Schulklassen im Projekt haben im ersten Schuljahr alle entweder mit dem Schulbuch «Primarstufe Mathematik 1» oder mit dem «Schweizer Zahlenbuch 1» gearbeitet. Die Aufteilung der Klassen und Kinder auf die beiden Lehrmittel ist in Kapitel 4.1.2 beschrieben.

Anhand der Stichprobe wurde untersucht, ob sich die Klassen bei den Ergebnissen für den Test für das Multiplikationsverständnis im Hinblick auf das verwendete Schulbuch unterscheiden. Die entsprechende Hypothese kann verneint werden. Die Mittelwerte bei den Testscores (Punktezahlen) liegen bei 16.74 bei den Kindern, die mit dem Lehrmittel «Primarstufe Mathematik 1» unterrichtet wurden sowie bei 16.77 bei den Kindern, die mit dem «Schweizer Zahlenbuch 1» unterrichtet wurden. Von Varianzhomogenität kann nach Durchführung eines Levene-Tests ausgegangen werden ($F(1,1022)=0.23$, $p=.64$), der folgende t-Test ist nicht signifikant ($t(1022)=-0.065$, $p=.95$).

5.7 Itemauswahl, Konstruktion einer Testskala

In diesem Kapitel werden die gesammelten Beobachtungen aus den vorgängigen Kapiteln verwendet, um basierend auf den 26 Items der beiden Erhebungsinstrumente eine Item-Auswahl zu treffen. Verwendet werden dabei die Ergebnisse aus den Betrachtungen mit Hilfe der klassischen Testtheorie als auch die Ergebnisse, die mit Hilfe der Item Response Theorie erreicht wurden. Wie in den vorangehenden Kapiteln dargelegt wurde, greifen die Ergebnisse aus der KTT und aus der IRT natürlicherweise ineinander und ergänzen sich. Als Beispiel hierfür können Items betrachtet werden, die aufgrund von besonderen Herausforderungen bei der Aufmerksamkeit oder bei der Aufnahmefähigkeit von sprachlichen Informationen in einer Faktoranalyse auf einen separaten Faktor laden. Bei solchen Items überrascht es nicht, wenn sie innerhalb der IRT bei der Analyse von DIF mit Split-Kriterium «Sprachverständnis» ebenfalls auffällig werden. In diesem Kapitel werden nun sämtliche Ergebnisse aus den Kapiteln 5.1 bis 5.6 verwendet, um über die Konstruktion der Testskala zu entscheiden. Es geht dabei insbesondere darum, zu entscheiden, ob und wenn ja welche Items für die weiteren Auswertungen aus dem Erhebungsinstrument ausgeschlossen werden sollen.

5.7.1 Konstruktion einer Testskala (Vortest)

Aus theoretischen Überlegungen (Lösen der Items durch sichtbares «Abzählen») sowie auch als Folge der Faktoranalyse empfiehlt es sich, auf die beiden Items 5a und 5b («Schokolade») zu verzichten. Für die weiteren Analysen werden diese beiden Items aus dem Erhebungsinstrument entfernt.

Die Items 16 bis 19 des Vortests wurden anders bearbeitet als die Items vorher. Unklar bleibt, ob es sich dabei um einen Ermüdungseffekt der Kinder (gegen Ende der Erhebung hin), um eine Unklarheit bei der Anweisung oder allenfalls um ein andersartiges Lösungsverhalten aufgrund anderer Fähigkeiten (die vier Items bedienen stärker als die anderen die «Strukturierungsfähigkeiten» der Kinder) handelt. Auch die Tatsache, dass ähnliche Effekte im Nachtest bei den vier Items mit den Punktefeldern wieder auftreten, trägt nicht dazu bei, diese Frage zu beantworten. Für die weiteren Analysen werden diese vier Items beibehalten.

Ebenfalls beibehalten werden die Items 4, 6 und 10, auch wenn möglicherweise Sprachverständnis bzw. Aspekte der Aufmerksamkeit bei der Bearbeitung dieser Items eine Rolle spielen. Die DIF-Analysen weisen vor allem bei Item 10 darauf hin, dass dies der Fall war. Beibehalten wird Item 10 vor allem auch deshalb, weil es aus theoretischen Gründen bedeutsam erscheint, zumindest ein Item beizubehalten, welches die zeitlich-sukzessive Grundvorstellung der Multiplikation bedient.

Letztlich werden also mit Ausnahme des Items 5 sämtliche Items beibehalten. Item 5 wird gänzlich (sowohl 5a als auch 5b) entfernt. Die Entscheidung, die Mehrheit der Items beizubehalten, geschieht im Bewusstsein, dass Operationsverständnis ein vielschichtiges Konstrukt mit mehreren Facetten darstellt. Die Items können verschiedene Distraktoren enthalten, welche kaum gänzlich aus einem solchen Erhebungsinstrument entfernt werden können. Bei den Erhebungen wurde soweit wie möglich versucht, Distraktoren auszuschalten. Es wurde auf

starke optische Unterstützung, keine zu häufigen Wechsel des Frageformates und auf mehrmaliges langsames Vorlesen der Items geachtet. Optimierungspotential besteht möglicherweise noch bei der Testdurchführung. Man könnte den Test kürzen und/oder die Kinder bis am Ende durch den Test durchführen. Allerdings erscheint es wahrscheinlich, dass auch dann einzelne Items DIF aufweisen und das Konstrukt bei der Faktoranalyse zerfällt, weil das Konstrukt Operationsverständnis unterschiedliche Facetten umfasst. Aus diesem Grunde wird für diese Arbeit ein «grosszügiger» Umgang mit verschiedenen kritischen Items gewählt.

5.7.2 Konstruktion einer Testskala (Nachtest)

Beim Nachtest fiel keines der Items bei der Testdurchführung derart negativ auf wie das «Schokoladen»-Item 5 im Vortest. Auch die Faktoranalyse und auch DIF-Analysen zeigen keine Items auf, welche aus dem Test entfernt werden müssten.

Diskutiert werden könnten am ehesten die Items 15 bis 18 («Punktefelder») bei denen DIF nach Geschlecht auftritt, was schwierig zu erklären ist. Aus theoretischen Gründen wurde entschieden, diese Items, die mit den Strukturierungsfähigkeiten einen wesentlichen Teil des Konstrukts Operationsverständnis bedienen, beizubehalten.

Der Nachtest wird somit mit allen 26 Items für die weiteren Analysen, insbesondere die Analysen des Längsschnitts in Kapitel 6, beibehalten.

Erwähnenswert ist, dass bei der Korrektur des Items 4 einige unerwartete Bearbeitungen durch die Kinder beobachtet worden sind. Die Codieranleitung wurde deshalb aber nicht angepasst und die Korrektur wurde gemäss der ursprünglichen Codieranleitung durchgeführt. Trotzdem ist das Item bei der quantitativen Analyse in keiner Art auffällig geworden und wird für die Auswertungen in dieser Arbeit beibehalten. Für etwaige zukünftige Erhebungen mit diesem Instrument empfiehlt es sich aber, auch dieses Item mit einer Frage nach der Gesamtanzahl Flaschen zu versehen. Bearbeitungen wie « 2 · 6 2 · 7 » deuten darauf hin, dass es für die Kinder nicht selbstverständlich war, dass dies der Fall war. Ebenfalls unerwartet und in der Codieranleitung nicht abgebildet (wenn auch selten) waren bei diesem Item Bearbeitungen der Art «7+7+6+6». Auch die Bearbeitung mit «13+13» sollte in einer Neuauflage der Codieranleitung erwähnt werden.

Einige Kinder haben die Items unerwarteterweise auch mit Hilfe von Divisionen bearbeitet. Dies kam selten und nur in einzelnen Klassen vor, die mit dem Stoff bereits so weit fortgeschritten waren, dass diese Operation im Unterricht bereits eingeführt worden ist. Entsprechende Bearbeitungen wurden (obwohl in der Codieranleitung nicht erwähnt) mit 1 codiert, wenn sie die Umkehrung einer Multiplikation beschrieben, die mit 1 codiert worden wäre (z.B. bei «28:7» anstelle von «4·7»).

6 Empirischer Teil 3: Ergebnisse der Längsschnittanalysen

In dieser Arbeit wurde der Frage nachgegangen, ob die Interventionen in den Unterricht in Form von Fördereinheiten eine Veränderung des Lernzuwachses bei den Kindern bewirken. In diesem Kapitel werden die Ergebnisse der diesbezüglichen Analyse wiedergegeben und damit die Forschungsfragen dieser Arbeit beantwortet.

Die Forschungsfragen zielen einerseits auf den Zuwachs der Rechenleistungen, welche mit dem Instrument BASIS-MATH zu den Testzeitpunkten T2 und T4 erfasst wurden. Die Ergebnisse dieser Untersuchung werden in Kapitel 6.1 berichtet.

Zudem zielen die Forschungsfragen auf den Lernzuwachs der Kinder, der mit dem Erhebungsinstrument für das Operationsverständnis zu den Testzeitpunkten T3 und T4 gemessen werden konnte. Von den Ergebnissen dieser Betrachtung wird in den Kapiteln 6.2 und 6.3 berichtet.

Die Analyse erfolgte mit den Auswertungsmethoden, die in Kapitel 4.5 dargelegt wurden.

6.1 Längsschnittanalyse mit dem Instrument BASIS-MATH (BM)

In diesem Kapitel wird der Lernfortschritt bei den Rechenleistungen untersucht. Die entsprechenden Fähigkeiten der Kinder wurden mit dem Instrument BASIS-MATH erfasst. In Kapitel 6.1.1 erfolgt zuerst eine Untersuchung mit der gesamten Stichprobe (zur Forschungsfrage FF1), während in Kapitel 6.1.2 dann Teilbestände von Kindern mit schwacher bzw. starker Rechenleistung (zur Forschungsfrage FF2) einzeln untersucht werden. In Kapitel 6.1.3 werden ergänzend hierzu die Ergebnisse von verschiedenen Modell-Variationen berichtet.

6.1.1 Lernfortschritt aller Kinder (BM)

Der Lernfortschritt der Kinder wird im Folgenden einzeln für die vier Interventionsgruppen I, II, III und IV des ersten Schuljahrs tabellarisch dargestellt. Von Interesse ist hauptsächlich die Frage, ob die Kinder der Interventionsklassen des zweiten Schuljahrs (I, II und IV) einen anderen Lernfortschritt zeigen als diejenigen der Kontrollgruppe (III). Tabellarisch dargestellt wird neben der Anzahl Kinder in den einzelnen Gruppen jeweils auch das Vorwissen in Form des mittleren Testscores im BASIS-MATH zum Testzeitpunkt T2. Der Lernfortschritt wird dargestellt als Zuwachs des Testscores (d.h. als Testscore BASIS-MATH 2+ abzüglich des Vorwissens). Diese Differenz ist leicht zu interpretieren, wurde aber für die Berechnung des Fortschritts ersetzt durch die Veränderung der Personenfähigkeit (mit der WLE-Schätzung) für die einzelnen Kinder. Der Durchschnitt dieses Werts für die jeweiligen Gruppen ist den Tabellen ebenfalls zu entnehmen.

In Tabelle 25 sind diese Werte nun für den Gesamtbestand der 669 Kinder dargestellt. Zu beachten ist, dass die Werte für den «Zuwachs» auch negativ werden können und dies keinen «Rückschritt» der Kinder darstellt. Bei der Differenz der Testscores werden die Testscores aus zwei unterschiedlichen Instrumenten (BASIS-MATH 2+ und BASIS-MATH 1+) gebildet.

Die Personenfähigkeit wiederum ist ein Parameter, der im Mittel über die Gesamtheit aller Kinder bei beiden Testzeitpunkten 0 beträgt. Negative Werte beim Zuwachs zwischen T2 und T4 bedeuten hier lediglich eine unterdurchschnittliche Zunahme der Fähigkeiten (gemessen am Gesamtbestand).

Tabelle 25: Übersicht Lern-Fortschritt im BASIS-MATH (alle 669 Kinder)

Gruppe	Einsatz FE 2. SJ	Anzahl Kinder	Mittlerer Testscore BM T2	Mittlerer Zuwachs BM	Mittlerer Zuwachs PF
I	Ja	165	19.9	1.0	-0.02
II	Ja	173	20.8	1.2	-0.01
III	Nein	167	22.2	1.1	-0.01
IV	Ja	164	20.4	1.2	0.03
alle		669	20.8	1.1	0.00

Anmerkungen. FE = Fördereinheiten, BM = BASIS-MATH, PF = Personenfähigkeit im Rasch-Modell («WLE»-Wert)

Für die Signifikanzmessung wird der in Kapitel 4.5 erläuterte Mehrebenen-Modellansatz mit dem Vorwissen und der Anwendung der Interventionen als Prädiktoren gewählt.

Das Vorwissen in Form der Personenfähigkeit zum Testzeitpunkt T2 stellt sich als höchst signifikanter Prädiktor heraus (β_1=0.588, p<.001). Die Anwendung der Interventionen weist eher auf einen geringeren Lernzuwachs bei Einsatz der Interventionen hin (negativer Wert von β_2), der Effekt ist aber nicht signifikant (β_2=-0.151, p=.289).

Tabelle 26: Ergebnisse der Mehrebenenanalyse (BASIS-MATH, Gesamtbestand)

Fixe Effekte	Schätzung	Std.fehler	df	t-Wert	p	β standardisiert
Vorwissen	0.588	0.028	665.7	21.278	<.001	.635
Intervention	-0.151	0.141	44.5	-1.072	.289	-.112

Anmerkungen. Anzahl Kinder: 669. Anzahl Klassen: 46.

Die fixen Effekte (Vorwissen und Intervention) korrelieren mit einem Wert von .069.

Die Ergebnisse zeigen, dass für den Gesamtbestand kaum Unterschiede zwischen IG und KG bestehen beim Lernfortschritt der Kinder. Das Vorzeichen des (nicht signifikanten und geringen) Effekts ist so, dass die Kinder in den Interventionsklassen geringeren Fortschritt gemacht haben als jene der Kontrollgruppe.

Der ICC-Wert im Nullmodell für diese Auswertung beträgt 9.6%. Es werden also 90.4% der Varianz in den Kinderleistungen durch individuelle Effekte und lediglich 9.6% der Varianz durch die Zugehörigkeit zur jeweiligen Klasse erklärt.

Durch die Modellausgestaltung inklusive Wahl der beiden Prädiktoren Vorwissen und Intervention werden R^2_m=41.1% bzw. R^2_c=46.5% der Gesamtvarianz erklärt.

Ohne diese beiden Prädiktoren (im Nullmodell) wird durch das Modell mit R^2_m= 0% kein Varianzanteil bzw. ein solcher von R^2_c=9.6% (wenn man den «random»-Effekt der Schwankung zwischen den Klassen mitberücksichtigt) erklärt. Der Wert von R^2_c entspricht in diesem Modell dem oben angegebenen ICC.

6.1.2 Lernfortschritt von Kindern mit unterschiedlicher Rechenleistung (BM)

Die Einteilung der Kinder in die jeweiligen Quartile erfolgte wie in Kapitel 4.5.6 beschrieben. Die Ergebnisse in diesem Kapitel beziehen sich auf die entsprechenden Teilbestände. Die Anzahl Kinder ist jeweils den Tabellen zu entnehmen. Ebenso ist die Anzahl Klassen angegeben, die sich ändern kann, da es möglich ist, dass einzelne Klassen keine («schwachen» oder «starken») Kinder zur jeweiligen Auswertung beitragen. Bei grösseren (oder stark negativen) Werten für β kann die Unterschiedlichkeit zwischen der Interventions- und der Kontrollgruppe jeweils auch der Zuwachstabelle entnommen werden.

Unteres Quartil (Gesamtbestand)

Tabelle 27: Übersicht Lern-Fortschritt im BASIS-MATH (unteres Quartil des Gesamtbestands)

Gruppe	Einsatz FE 2. SJ	Anzahl Kinder	Mittlerer Testscore BM T2	Mittlerer Zuwachs BM	Mittlerer Zuwachs PF
I	Ja	53	12.2	3.1	0.57
II	Ja	43	12.7	3.6	0.64
III	Nein	28	14.4	1.0	0.08
IV	Ja	39	12.5	2.9	0.59
alle		163	12.8	2.8	0.51

Das Vorwissen in Form der Personenfähigkeit zum Testzeitpunkt T2 zeigt sich als signifikanter Prädiktor (β_1=0.656, p<.001). Die Anwendung der Interventionen zeigt keinen signifikanten Effekt (β_2=0.370, p=.137).

Tabelle 28: Ergebnisse der Mehrebenenanalyse (BASIS-MATH, unteres Quartil des Gesamtbestands)

Fixe Effekte	Schätzung	Std.fehler	df	t-Wert	p	β standardisiert
Vorwissen	0.656	0.097	159.4	6.744	<.001	.482
Intervention	0.370	0.245	55.7	1.509	.137	.318

Anmerkungen. Anzahl Kinder: 163. Anzahl Klassen: 43.

Die fixen Effekte (Vorwissen und Intervention) korrelieren mit einem Wert von .133.

Die Ergebnisse zeigen, dass sich die Interventionen in diesem Teilbestand positiv auf den Lernfortschritt der Kinder ausgewirkt haben. Das Ergebnis ist allerdings statistisch nicht signifikant, d.h. es könnte auch durch Zufall entstanden sein. Der Nachweis der statistischen Signifikanz wird dadurch erschwert, dass die Kontrollgruppe in diesem Teilbestand mit lediglich 28 Kindern nicht mehr sehr umfangreich ist. Die andersartige Auswahl der «schwächeren» 25% der Kinder aus dem Gesamtbestand, die im Folgenden untersucht wurde, geschah auch um zu prüfen, ob eine andere Zuteilung den positiven Effekt der Interventionen verstärkt sichtbar oder sogar signifikant werden lässt. Insofern ging es auch darum, zufällige Effekte bei der Auswahl der Kinder zu erkennen und wenn möglich zu eliminieren.

Durch die Modellausgestaltung inklusive Wahl der beiden Prädiktoren Vorwissen und Intervention werden R^2_m=22.4% bzw. R^2_c=31.3% der Gesamtvarianz in diesem Modell erklärt. Zu beachten ist, dass in den Teilbeständen jeweils auch die Gesamtvarianz des Modells eine andere ist und sich die Werte daher zwischen Beständen nicht vergleichen lassen.

Unteres Quartil (pro Gruppe)

Tabelle 29: Übersicht Lern-Fortschritt im BASIS-MATH (unteres Quartil jeder Interventionsgruppe)

Gruppe	Einsatz FE 2. SJ	Anzahl Kinder	Mittlerer Testscore BM T2	Mittlerer Zuwachs BM	Mittlerer Zuwachs PF
I	Ja	41	10.9	3.1	0.74
II	Ja	43	12.7	3.6	0.64
III	Nein	42	15.7	2.1	0.06
IV	Ja	41	12.8	3.0	0.56
alle		167	13.0	3.0	0.50

Das Vorwissen in Form der Personenfähigkeit zum Testzeitpunkt T2 zeigt sich als signifikanter Prädiktor (β_1=0.711, p<.001). Die Anwendung der Interventionen zeigt keinen signifikanten Effekt (β_2=0.095, p=.692).

Tabelle 30: Ergebnisse der Mehrebenenanalyse (BASIS-MATH, unteres Quartil jeder Interventionsgruppe)

Fixe Effekte	Schätzung	Std.fehler	df	t-Wert	p	β standardisiert
Vorwissen	0.711	0.101	164.0	7.034	<.001	.507
Intervention	0.095	0.239	49.0	0.399	.692	.077

Anmerkungen. Anzahl Kinder: 167. Anzahl Klassen: 44.

Die fixen Effekte (Vorwissen und Intervention) korrelieren mit einem Wert von .254.

Die Ergebnisse zeigen, dass die Interventionen in diesem Teilbestand einen leichten positiven Einfluss auf den Lernfortschritt der Kinder hatten. Das Ergebnis ist allerdings statistisch eben-

falls nicht signifikant und könnte auch durch Zufall entstanden sein. Die andersartige Aufteilung des Bestands ändert also nichts an der früheren Erkenntnis.

Durch die Modellausgestaltung inklusive Wahl der beiden Prädiktoren Vorwissen und Intervention werden R^2_m=24.4% bzw. R^2_c=35.2% der Gesamtvarianz in diesem Modell erklärt.

Die «rechenstarken» Kinder «oberhalb» des unteren Quartils (Gesamtbestand)

Tabelle 31: Übersicht Lern-Fortschritt im BASIS-MATH (stärkere 75% des Gesamtbestands)

Gruppe	Einsatz FE 2. SJ	Anzahl Kinder	Mittlerer Testscore BM T2	Mittlerer Zuwachs BM	Mittlerer Zuwachs PF
I	Ja	112	23.5	0.0	-0.31
II	Ja	130	23.5	0.4	-0.23
III	Nein	139	23.8	1.2	-0.03
IV	Ja	125	22.8	0.6	-0.12
alle		506	23.4	0.6	-0.17

Das Vorwissen in Form der Personenfähigkeit zum Testzeitpunkt T2 zeigt sich als signifikanter Prädiktor (β_1=0.423, p<.001). Die Anwendung der Interventionen zeigt keinen signifikanten Effekt (β_2=-0.268, p=.084).

Tabelle 32: Ergebnisse der Mehrebenenanalyse (BASIS-MATH, stärkere 75% des Gesamtbestands)

Fixe Effekte	Schätzung	Std.fehler	df	t-Wert	p	β standardisiert
Vorwissen	0.424	0.040	492.0	10.537	<.001	.412
Intervention	-0.268	0.151	40.4	-1.772	.084	-.238

Anmerkungen Anzahl Kinder: 506. Anzahl Klassen: 46.

Die fixen Effekte (Vorwissen und Intervention) korrelieren mit einem Wert von .023.

Die Rechnungen bestärken die Vermutung, die sich aus Tabelle 31 ergibt, nämlich dass sich die Interventionen in diesem Teilbestand negativ auf den Lernfortschritt der Kinder ausgewirkt haben. Beobachtet wurde, dass Kinder in der Kontrollgruppe bei dieser Messung den höheren Lernzuwachs hatten als Kinder in den Interventionsgruppen des 2. Schuljahrs. Das Ergebnis ist aber statistisch nicht signifikant, d.h. es kann nicht ausgeschlossen werden, dass es durch Zufall entstanden ist. Die Wahrscheinlichkeit für eine derart ausgeprägte (oder noch ausgeprägtere) Beobachtung, wie sie hier gemacht wurde, liegt bei 8.4%, was bei der Konvention eines 5%-Testniveaus nicht ausreicht, um einen Zufallseffekt auszuschliessen. Das Vorzeichen soll aber vorerst Beachtung finden, ebenso die Tatsache, dass der p-Wert nicht weit von der 5%-Schwelle entfernt liegt. Wie sich noch zeigen wird, wird das 5%-Niveau bei anderen Bestands-Aufteilungen unterschritten werden.

Durch die Modellausgestaltung inklusive Wahl der beiden Prädiktoren Vorwissen und Intervention werden R^2_m=18.7% bzw. R^2_c=27.2% der Gesamtvarianz in diesem Modell erklärt.

Die «rechenstarken» Kinder «oberhalb» des unteren Quartils (pro Gruppe)

Tabelle 33: Übersicht Lern-Fortschritt im BASIS-MATH (stärkere 75% jeder Interventionsgruppe)

Gruppe	Einsatz FE 2. SJ	Anzahl Kinder	Mittlerer Testscore BM T2	Mittlerer Zuwachs BM	Mittlerer Zuwachs PF
I	Ja	124	22.8	0.3	-0.28
II	Ja	130	23.5	0.4	-0.23
III	Nein	125	24.4	0.8	-0.03
IV	Ja	123	22.9	0.6	-0.12
alle		502	23.4	0.5	-0.17

Das Vorwissen in Form der Personenfähigkeit zum Testzeitpunkt T2 zeigt sich als signifikanter Prädiktor (β_1=0.431, p<.001). Die Anwendung der Interventionen zeigt keinen signifikanten Effekt (β_2=-0.278, p=.075).

Tabelle 34: Ergebnisse der Mehrebenenanalyse (BASIS-MATH, stärkere 75% jeder Interventionsgruppe)

Fixe Effekte	Schätzung	Std.fehler	df	t-Wert	p	β standardisiert
Vorwissen	0.431	0.040	483.6	10.737	<.001	.419
Intervention	-0.278	0.153	43.6	-1.822	.075	-.245

Anmerkungen. Anzahl Kinder: 502. Anzahl Klassen: 46.

Die fixen Effekte (Vorwissen und Intervention) korrelieren mit einem Wert von .083.

Die Ergebnisse der Rechnung in Tabelle 34 bestätigen die Vermutung, die sich auch aus Tabelle 33 ergibt, nämlich dass in den Interventionsklassen in diesem Teilbestand der Lernfortschritt der Kinder geringer war als in den Klassen der Kontrollgruppe. Das Ergebnis ist aber wiederum statistisch nicht signifikant. Es wurde wiederum beobachtet, dass Kinder in der Kontrollgruppe den höheren Lernzuwachs hatten als Kinder in den Interventionsgruppen des 2. Schuljahrs, die Unterschiedlichkeit könnte aber – gemäss der 5%-Konvention − wiederum durch Zufall erklärt werden.

Durch die Modellausgestaltung inklusive Wahl der beiden Prädiktoren Vorwissen und Intervention werden R^2_m=20.1% bzw. R^2_c=28.1% der Gesamtvarianz in diesem Modell erklärt.

Fazit

Die Interventionen wirkten in den Stärkeklassen unterschiedlich. Die «schwächeren» Kinder konnten von den Interventionen «profitieren». Die «schwachen» Kinder in den Interventionsklassen wiesen einen höheren Lernfortschritt auf als die Kinder in der Kontrollgruppe. Bei den «starken» Kindern war es umgekehrt. Diese wiesen in der Kontrollgruppe einen höheren Lernzuwachs auf als in den drei Interventionsgruppen. Bei allen Aufteilungen in «schwache» und «starke» Kinder sind die gemessenen Effekte auf dem 5%-Niveau nicht signifikant. Es konnte also beobachtet werden, dass für die jeweiligen Stärkegruppe in der Interventionsgruppe mehr (bzw. weniger) und in der anderen Kontrollgruppe weniger (bzw. mehr) Lernzuwachs stattfand. Es konnte aber rechnerisch nicht nachgewiesen werden, dass diese Unterschiedlichkeit beim Lernzuwachs nicht auch durch Zufall entstanden sein könnte.

Unterteilung bei «20%/80%»

Die Effekte wurden alle auch bei Wahl der «schwächsten» 20% anstelle der «schwächsten» 25% (sowie auch für die «stärkeren» 80% anstelle der «stärkeren» 75%) gerechnet und verändern sich dabei nicht wesentlich. Die Effekte haben dasselbe Vorzeichen, eine ähnliche Grössenordnung und sind ebenfalls alle nicht signifikant.

Unterteilung bei «50%/50%»

Ebenfalls wurde der Bestand noch einmal zerlegt in zwei (in etwa) gleich grosse Bestände der «schwächeren» und der «stärkeren» Kinder des Gesamtbestands. Diese Zweiteilung geschah wie bereits erläutert auf mehrere Arten. Die Ergebnisse der sich daraus ergebenden vier Varianten lauten:

Die «schwächere» Hälfte der Kinder (Gesamtbestand), Split bei 21 (Testscore ≤ 21)

Tabelle 35: Übersicht Lern-Fortschritt im BASIS-MATH (schwächere Hälfte, Split bei 21)

Gruppe	Einsatz FE 2. SJ	Anzahl Kinder	Mittlere Punktezahl BM T2	Mittlerer Zuwachs BM	Mittlerer Zuwachs PF
I	Ja	85	15.1	2.5	0.47
II	Ja	82	16.0	2.6	0.46
III	Nein	62	17.2	2.0	0.35
IV	Ja	88	16.5	2.7	0.54
alle		317	16.1	2.5	0.46

Das Vorwissen in Form der Personenfähigkeit zum Testzeitpunkt T2 zeigt sich als signifikanter Prädiktor (β_1=0.797, p<.001). Die Anwendung der Interventionen zeigt keinen signifikanten Effekt (β_2=0.121, p=.514).

Tabelle 36: Ergebnisse der Mehrebenenanalyse (BASIS-MATH, schwächere Hälfte, Split bei 21)

Fixe Effekte	Schätzung	Std.fehler	df	t-Wert	p	β standardisiert
Vorwissen	0.797	0.066	310.9	12.153	<.001	.583
Intervention	0.121	0.184	51.1	0.657	.514	.099

Anmerkungen. Anzahl Kinder: 317. Anzahl Klassen: 46.

Die Korrelation der fixen Effekte (Vorwissen und Intervention) beträgt .006.

Die Ergebnisse deuten darauf hin, dass die Interventionen in diesem Teilbestand einen (leicht) positiven Einfluss auf den Lernfortschritt der Kinder hatten. Das Ergebnis ist statistisch nicht signifikant und kann somit auch durch Zufall entstanden sein.

Durch die Modellausgestaltung inklusive Wahl der beiden Prädiktoren Vorwissen und Intervention werden R^2_m=32.6% bzw. R^2_c=40.6% der Gesamtvarianz in diesem Modell erklärt.

Die «stärkere» Hälfte der Kinder (Gesamtbestand), (Testscore >21)

Tabelle 37: Übersicht Lern-Fortschritt im BASIS-MATH (stärkere Hälfte, Split bei 21)

Gruppe	Einsatz FE 2. SJ	Anzahl Kinder	Mittlere Punktezahl BM T2	Mittlerer Zuwachs BM	Mittlerer Zuwachs PF
I	Ja	80	25.0	-0.6	-0.54
II	Ja	91	25.2	0.0	-0.44
III	Nein	105	25.2	0.5	-0.23
IV	Ja	76	24.8	-0.6	-0.52
alle		352	25.1	-0.1	-0.42

Das Vorwissen in Form der Personenfähigkeit zum Testzeitpunkt T2 zeigt sich als signifikanter Prädiktor (β_1=0.399, p<.001). Die Anwendung der Interventionen wirkt in die negative Richtung und ist auf dem 5%-Niveau signifikant (β_2=-0.372, p=.034).

Tabelle 38: Ergebnisse der Mehrebenenanalyse (BASIS-MATH, stärkere Hälfte, Split bei 21)

Fixe Effekte	Schätzung	Std.fehler	df	t-Wert	p	β standardisiert
Vorwissen	0.399	0.056	341.9	7.097	<.001	.345
Intervention	-0.372	0.169	38.2	-2.205	.034	-.336

Anmerkungen. Anzahl Kinder: 352. Anzahl Klassen: 46.

Die fixen Effekte (Vorwissen und Intervention) korrelieren mit einem Wert von -.008.

Die Ergebnisse zeigen, dass die Interventionen in diesem Teilbestand einen negativen Einfluss auf den Lernfortschritt der Kinder hatten bzw. dass Kinder in der Kontrollgruppe den höheren Lernzuwachs hatten als Kinder in den Interventionsgruppen des 2. Schuljahrs. Das Ergebnis ist mit $\beta = -0.372$ verhältnismässig ausgeprägt und auf dem 5%-Niveau auch statistisch signifikant. Vergleicht man die beiden standardisierten β-Werte, so unterstreicht dies die negative Wirkung der Intervention im Vergleich zur positiven Wirkung des Vorwissens.

Durch die Modellausgestaltung inklusive Wahl der beiden Prädiktoren Vorwissen und Intervention werden R^2_m=14.4% bzw. R^2_c=23.9% der Gesamtvarianz in diesem Modell erklärt.

Die «schwächere» Hälfte der Kinder (Gesamtbestand), Split bei 22 (Testscore ≤ 22)

Tabelle 39: Übersicht Lern-Fortschritt im BASIS-MATH (schwächere Hälfte, Split bei 22)

Gruppe	Einsatz FE 2. SJ	Anzahl Kinder	Mittlere Punktezahl BM T2	Mittlerer Zuwachs BM	Mittlerer Zuwachs PF
I	Ja	95	15.8	2.0	0.36
II	Ja	91	16.5	2.5	0.43
III	Nein	74	17.9	1.5	0.26
IV	Ja	97	17.1	2.6	0.53
alle		357	16.8	2.2	0.41

Das Vorwissen in Form der Personenfähigkeit zum Testzeitpunkt T2 zeigt sich als signifikanter Prädiktor (β_1=0.760, p<.001). Die Anwendung der Interventionen zeigt keinen signifikanten Effekt (β_2=0.141, p=.417).

Tabelle 40: Ergebnisse der Mehrebenenanalyse (BASIS-MATH, schwächere Hälfte, Split bei 22)

Fixe Effekte	Schätzung	Std.fehler	df	t-Wert	p	β standardisiert
Vorwissen	0.760	0.059	352.4	12.825	<.001	.578
Intervention	0.141	0.172	50.0	0.818	.417	.117

Anmerkungen. Anzahl Kinder: 357. Anzahl Klassen: 46.

Die fixen Effekte (Vorwissen und Intervention) korrelieren mit einem Wert von .079.

Die Ergebnisse zeigen, dass die Interventionen in diesem Teilbestand einen leicht positiven Einfluss auf den Lernfortschritt der Kinder hatten. Das Ergebnis ist statistisch nicht signifikant.

Durch die Modellausgestaltung inklusive Wahl der beiden Prädiktoren Vorwissen und Intervention werden R^2_m=32.1% bzw. R^2_c=39.8% der Gesamtvarianz in diesem Modell erklärt.

Die «stärkere» Hälfte der Kinder (Gesamtbestand), (Testscore>22)

Tabelle 41: Übersicht Lern-Fortschritt im BASIS-MATH (stärkere Hälfte, Split bei 22)

Gruppe	Einsatz FE 2. SJ	Anzahl Kinder	Mittlere Punktezahl BM T2	Mittlerer Zuwachs BM	Mittlerer Zuwachs PF
I	Ja	70	25.4	-0.4	-0.55
II	Ja	82	25.6	-0.2	-0.51
III	Nein	93	25.6	0.7	-0.23
IV	Ja	67	25.2	-1.0	-0.64
alle		312	25.5	-0.1	-0.46

Das Vorwissen in Form der Personenfähigkeit zum Testzeitpunkt T2 zeigt sich als signifikanter Prädiktor (β_1=0.331, p<.001). Die Anwendung der Interventionen wirkt in die negative Richtung und ist auf dem 5%-Niveau signifikant (β_2=-0.503, p=.015*).

Tabelle 42: Ergebnisse der Mehrebenenanalyse (BASIS-MATH, stärkere Hälfte, Split bei 22)

Fixe Effekte	Schätzung	Std.fehler	df	t-Wert	p	β standardisiert
Vorwissen	0.331	0.062	299.9	5.355	<.001	.278
Intervention	-0.503	0.197	35.5	-2.548	.015	-.458

Anmerkungen. Anzahl Kinder: 312. Anzahl Klassen: 46.

Die fixen Effekte (Vorwissen und Intervention) korrelieren mit einem Wert von -.011.

Die Ergebnisse zeigen, dass die Interventionen in diesem Teilbestand einen negativen Einfluss auf den Lernfortschritt der Kinder hatten bzw. dass Kinder in der Kontrollgruppe den höheren Lernzuwachs hatten als Kinder in den Interventionsgruppen des 2. Schuljahrs. Das Ergebnis ist mit β = -0.503 ausgeprägt und auf dem 5%-Niveau auch statistisch signifikant.

Durch die Modellausgestaltung inklusive Wahl der beiden Prädiktoren Vorwissen und Intervention werden R^2_m=12.2% bzw. R^2_c=27.1% der Gesamtvarianz in diesem Modell erklärt.

Zusätzliche gerechnete Variante

Die beobachteten Effekte wurden zur Vermeidung von Zufallseffekten auch für eine Aufteilung im höheren Bereich («cut-off»-Punkt beim Testscore 23 im BASIS-MATH-1+) gerechnet und ändern sich weder beim Vorzeichen, bei der Grössenordnung noch bei der Signifikanz.

6.1.3 Variationen des gewählten Mehrebenen-Modells

An dieser Stelle werden Variationen und Alternativen zum gewählten Mehrebenen-Modellansatz besprochen und die Änderung der sich dabei ergebenden Ergebnisse jeweils dargelegt. Es sind zum Teil erst Ergebnisse der vorgenommenen Berechnungen, welche die jeweilige Modellwahl rechtfertigen. Beispielsweise wäre ein «random slope»-Modell aus rein theoretischen Gesichtspunkten genauso vertretbar wie ein «random intercept»-Modell. Erst die Anwendung am vorliegenden Datensatz zeigt auf, dass der zusätzliche Freiheitsgrad bei der Parameterwahl im «random slope»-Modell in diesem konkreten Fall keinen Mehrwert liefert. Aus diesem Grund erfolgt dieser Teil der Modellwahl-Besprechung ergänzend zu den Erläuterungen in Kapitel 4.5.6 an dieser Stelle.

Da sich der BASIS-MATH im Vergleich zum Erhebungsinstrument für das Multiplikationsverständnis als das sensiblere Erhebungsinstrument herausgestellt hat, erfolgt diese Besprechung ausschliesslich in Kapitel 6.1 und somit ausschliesslich unter Verwendung des Instruments BASIS-MATH.

Differenz anstelle der Nachtest-Leistung als Zielvariable

Das Modell wurde auch gerechnet mit dem Zuwachs in der Personenfähigkeit anstelle der Personenfähigkeit im Nachtest als Zielvariable. Wie in Kapitel 4.5.5 erläutert ergibt sich in den Werten für den Parameter β_2 sowie auch im p-Wert dadurch keine Veränderung gegenüber dem gewählten Ansatz.

Tabelle 43: Änderung einzelner Parameter-Werte bei Änderung der abhängigen Variablen

Abhängige Variable	Leistung im Nachtest (Personenfähigkeit)	Leistungszuwachs (Zuwachs bei der Personenfähigkeit)
Korrelation X_1 und X_2	.069	.069
Schätzung für β_1 (X_1: Vorwissen)	0.588	-0.412
p-Wert für β_1	<.001	<.001
Schätzung für β_2 (X_2: Interventionen)	-0.151	-0.151
p-Wert für β_2	.289	.289
ICC-Wert	9.6%	8.8%
Erklärte Varianz R^2_m	41.1%	24.8%
Erklärte Varianz R^2_c	46.5%	31.6%

Anmerkungen. Anzahl Kinder: 669. Anzahl Klassen 46.

Der ICC-Wert des Nullmodells bei Verwendung der Personenfähigkeit im Nachtest liegt bei 9.6%, der ICC-Wert des Nullmodells bei Verwendung des Zuwachses in der Personenfähigkeit liegt bei 8.8% (siehe Tabelle 43).

Auch wenn sich der mit dem ICC-Wert gemessene Anteil der Varianz, welcher durch die Klassenzuteilung erklärt werden kann, zwischen den Modellen unterscheidet, so ergibt sich für die Frage der Wirksamkeit der Interventionen keine Unterschiedlichkeit zwischen den beiden Modellen. In beiden Fällen kann für den Parameter β_2 der identische Wert für β_2=-0.151 (mit p=.289 nicht signifikant) errechnet werden. Zu erkennen ist auch, wie der Parameter β_1 sich (wie in Kapitel 4.5.6 erläutert) um einen additiven Summanden von 1 unterscheidet. Bei der Verwendung des Zuwachses anstelle der Personenfähigkeit reduziert sich β_1 genau um den Wert 1.

Was sich bei der Verwendung der Differenzen von Personenfähigkeiten anstelle der Personenfähigkeit im Nachtest aber ändert, sind die durch die Modelle erklärten Varianzanteile.

Verwendung des «random slope»-Modells anstelle des «random intercept»-Modells

Mit der Verwendung des «random slope»-Modells anstelle des «random intercept»-Modells werden zusätzliche Freiheitsgrade im Modell aufgenommen. Die «slopes» können sich im «random slope»-Modell zwischen den einzelnen Klassen unterscheiden.

Die Verwendung des «random slope»-Modells lässt die Werte AIC und BIC für die Modellgüte jedoch schlechter werden. Illustriert wird dies in Tabelle 44 anhand des gesamten Datensatzes für den Lernfortschritt gemessen mit dem BASIS-MATH. Die Werte für die Kennzahlen AIC und BIC steigen. Dies kann so interpretiert werden, dass die höhere Komplexität des «random slope»-Modells stärker negativ ins Gewicht fällt als der Gewinn an Genauigkeit, welcher durch die zusätzlichen Freiheitsgrade im «random slope»-Modell erreicht wird.

Tabelle 44: Modellgüte-Masse „random intercept“-Modell und „random slope“-Modell

Modell	df	AIC	BIC	logLik	Deviance	Chisq	Chi df	p(>Chisq)
random intercept	5	1'938.3	1'960.8	-964.1	1'928.3			
random slope	10	1'945.8	1'990.9	-962.9	1'925.8	2.45	5	.785

Die durch das Modell erklärten Varianzanteile steigen bei Verwendung des «random slope»-Modells, allerdings in geringem Masse:

Tabelle 45: erklärte Varianzanteile („random intercept“ und „random slope“-Modelle)

Modell	R^2_m	R^2_c
random intercept	41.1%	46.5%
random slope	41.9%	48.5%

In dieser Arbeit wird als Folge dieser Beobachtung für sämtliche Analysen mit Teilbeständen und auch für die Erfassung des Lernfortschritts mit dem Erhebungsinstrument für das Multiplikationsverständnis das «random intercept»-Modell verwendet.

Verwendung der Interventionen als Prädiktor

Als signifikanter Prädiktor für die Leistungen im Nachtest wurde die Leistung im Vortest geortet. Die Interventionen als zusätzlicher Prädiktor − so wie im Modellansatz verwendet – liefern für den Gesamtbestand keine Verbesserung des Modells (siehe Tabelle 46).

Tabelle 46: Modellgüte-Masse ohne bzw. mit Intervention auf Ebene 1

Modell	df	AIC	BIC	logLik	Deviance	Chisq	Chi df	P(>Chisq)
nur Vorwissen	4	1'937.5	1'955.5	-964.7	1'929.5			
Vorwissen und Interventionen	5	1'938.3	1'960.8	-964.1	1'928.3	1.193	1	.275

Auch der erklärte Anteil an der Gesamtvarianz sinkt nur unwesentlich, wenn man die Interventionen als Prädiktor entfernt (siehe Tabelle 47). Der Unterschied wird etwas grösser, wenn man einen Teilbestand betrachtet, in dem die Interventionen eine signifikante Wirkung gezeigt haben (siehe Tabelle 48).

Tabelle 47: Veränderung der erklärten Varianzanteile (Gesamtbestand der 669 Kinder)

Modell	R^2_m	R^2_c
Vorwissen und Interventionen	41.1%	46.5%
nur Vorwissen (ohne Interventionen)	40.6%	46.1

Tabelle 48: Veränderung der erklärten Varianzanteile (stärkere Hälfte, Split bei 21)

Modell	R^2_m	R^2_c
Vorwissen und Interventionen	14.4%	23.9%
nur Vorwissen (ohne Interventionen)	11.8%	22.5%

Verwendung weiterer Prädiktoren im Mehrebenenmodell

Als weitere Prädiktoren wurden die acht unten aufgeführten zusätzlich verfügbaren Variablen auf ihre Nützlichkeit hin durch Einführen ins Mehrebenen-Modell geprüft. Das Vorwissen und die Information, ob die Interventionen durchgeführt wurden, sind dabei als Prädiktoren im Modell beibehalten worden.

Zusätzliche Variablen als Kandidaten für Prädiktoren im Mehrebenenmodell sind das Arbeitsgedächtnis, die kognitiven Fähigkeiten, das Sprachverständnis für die deutsche Sprache, die Ausdrucksfähigkeit in der deutschen Sprache, die Familiensprache des Kindes, das Geschlecht, das Alter des Kindes sowie das in der Klasse eingesetzte Lehrmittel.

Für das Arbeitsgedächtnis und die kognitiven Fähigkeiten wurde der Testscore bei der Erfassung zum Zeitpunkt T1 verwendet. Die Variablen Sprachverständnis und Ausdrucksfähigkeit wurden in dichotome Variablen umcodiert. Dies geschah so, dass beim Sprachverständnis und bei der Ausdrucksfähigkeit nur die höchste Stufe der Lehrpersonen-Einschätzung in eine 1 umgewandelt wurde. Ebenso wurde die Familiensprache in eine dichotome Variable umcodiert (0: «deutsch/schweizerdeutsch ist (auch) Familiensprache mit Einschätzung «1» oder «2» durch die LP, 1: «deutsch als Zweitsprache», Einschätzung «3», siehe Anhang 8.1.1). Geschlecht wurde als dichotome Variable (0: «weiblich», 1: «männlich») verwendet. Für das Alter wurde das Alter der Kinder beim Ablegen des Tests für das Vorwissen (also bei der BASIS-MATH-Erhebung zum Zeitpunkt T2) gewählt. Lehrmittel ist eine dichotome Variable (0: «Primarstufe Mathematik», 1: «Schweizer Zahlenbuch»).

Verzichtet wurde auf die Interventionsgruppe (I, II, III, IV) als Prädiktor-Variable, weil der Einfluss des ersten Schuljahrs auf den Verlauf des zweiten als weniger zentral als andere Prädiktoren (z.B. die Durchführung der Interventionen im zweiten Schuljahr) eingeschätzt wurde und weil die Wirkung der Interventionen des ersten Schuljahrs sich bereits im Vorwissen zeigt. Die Interventionen des zweiten Schuljahrs (und damit der Vergleich zwischen Gruppe III und den anderen drei Gruppen) sind im Modell durchgehend abgebildet.

Die Ergebnisse des einzelnen Hinzunehmens dieser Prädiktoren sind die folgenden:

Arbeitsgedächtnis

Tabelle 49: Parameterschätzungen der β-Werte bei Hinzunahme des Arbeitsgedächtnisses als Prädiktor

Fixe Effekte	Schätzung	Std.fehler	df	t-Wert	p	β standardisiert
Vorwissen	0.531	0.030	652.2	17.835	<.001	.575
Intervention	-0.102	0.139	44.8	-0.732	.468	-.075
Arbeitsgedächtnis	0.084	0.019	650.9	4.474	<.001	.145

Kognitive Fähigkeiten

Tabelle 50: Parameterschätzungen der β-Werte bei Hinzunahme der kognitiven Fähigkeiten als Prädiktor

Fixe Effekte	Schätzung	Std.fehler	df	t-Wert	p	β standardisiert
Vorwissen	0.520	0.030	653.7	17.583	<.001	.563
Intervention	-0.142	0.136	44.7	-1.045	.302	-.105
CFT (kogn. Fähigk.)	0.049	0.009	654.9	5.498	<.001	.174

Sprachverständnis für die deutsche Sprache

Tabelle 51: Parameterschätzungen der β-Werte bei Hinzunahme des Sprachverständnisses als Prädiktor

Fixe Effekte	Schätzung	Std.fehler	df	t-Wert	p	β standardisiert
Vorwissen	0.567	0.028	663.8	20.279	<.001	.613
Intervention	-0.104	0.145	45.3	-0.718	.476	-.077
Sprachverständnis	0.389	0.108	650.4	3.602	<.001	.113

Ausdrucksfähigkeit für die deutsche Sprache

Tabelle 52: Parameterschätzungen der β-Werte bei Hinzunahme der Ausdrucksfähigkeit als Prädiktor

Fixe Effekte	Schätzung	Std.fehler	df	t-Wert	p	β standardisiert
Vorwissen	0.566	0.028	664.1	20.184	<.001	.612
Intervention	-0.127	0.142	45.2	-0.897	.375	-.094
Ausdrucksfähigkeit	0.337	0.096	647.1	3.501	<.001	.110

Familiensprache (bzw. Deutsch als Zweitsprache «DaZ»)

Tabelle 53: Parameterschätzungen der β-Werte bei Hinzunahme der Familiensprache als Prädiktor

Fixe Effekte	Schätzung	Std.fehler	df	t-Wert	p	β standardisiert
Vorwissen	0.587	0.028	664.1	21.144	<.001	.635
Intervention	-0.153	0.142	44.7	-1.077	.287	-.113
Familiensprache	-0.022	0.110	613.4	-0.204	.838	-.006

Geschlecht

Tabelle 54: Parameterschätzungen der β-Werte bei Hinzunahme des Geschlechts als Prädiktor

Fixe Effekte	Schätzung	Std.fehler	df	t-Wert	p	β standardisiert
Vorwissen	0.580	0.028	665.0	20.596	<.001	.627
Intervention	-0.153	0.141	44.5	-1.084	.284	-.113
Geschlecht	0.111	0.080	650.8	1.393	.164	.041

Alter

Tabelle 55: Parameterschätzungen der β-Werte bei Hinzunahme des Alters als Prädiktor

Fixe Effekte	Schätzung	Std.fehler	df	t-Wert	p	β standardisiert
Vorwissen	0.592	0.027	664.8	21.568	<.001	.640
Intervention	-0.140	0.140	44.7	-1.005	.320	-.103
Alter bei T2	-0.347	0.105	663.7	-3.298	<.001	-.097

Lehrmittel

Tabelle 56: Parameterschätzungen der β-Werte bei Hinzunahme des Lehrmittels als Prädiktor

Fixe Effekte	Schätzung	Std.fehler	df	t-Wert	p	β standardisiert
Vorwissen	0.588	0.028	664.9	21.272	<.001	.635
Intervention	-0.174	0.144	43.1	-1.203	.235	-.128
Lehrmittel	0.100	0.125	41.6	0.797	.430	.037

Fazit

Es zeigt sich, dass das Arbeitsgedächtnis einen signifikanten und positiven (je höher das Arbeitsgedächtnis desto höher der Lernzuwachs) Einfluss auf den Lernzuwachs hat. Dieselbe Aussage kann für die kognitiven Fähigkeiten gemacht werden. Der Effekt ist signifikant und

positiv (je höher die bei T1 gemessenen kognitiven Fähigkeiten desto höher der Lernzuwachs). Sprachverständnis und Ausdrucksfähigkeit in der deutschen Sprache haben beide ebenfalls einen positiven, signifikanten Einfluss auf den Lernzuwachs. Der Einfluss der Familiensprache ist gering und nicht signifikant. Der Einfluss des Geschlechts ist positiv (d.h. die mit 1 codierten Knaben lernen mehr dazu als die Mädchen), der Effekt ist aber nicht signifikant. Das Alter zum Zeitpunkt T2 hat einen negativen (jüngere Kinder lernen mehr dazu als ältere) und ebenfalls signifikanten Einfluss auf den Lernzuwachs. Der Einfluss des verwendeten Lehrmittels ist nicht signifikant.

Arbeitsgedächtnis, kognitive Fähigkeiten, Sprachverständnis, Ausdrucksfähigkeit in der deutschen Sprache und auch Geschlecht korrelieren mit dem Vorwissen der Kinder zum Messzeitpunkt T2. Beim Alter ist diese Korrelation gering, beim Lehrmittel ist sie nicht vorhanden.

Tabelle 57: Veränderung der Modellgüte-Masse AIC und BIC durch Einfügen von Prädiktoren

		alle 669 Kinder		657 Kinder	
Prädiktoren	df	AIC	BIC	AIC	BIC
Nullmodell (ohne Prädiktoren)	3	2'285.2	2'298.7	2'239.2	2'252.7
nur Interventionen	4	2'283.1	2'301.1	2'237.4	2'255.4
nur Vorwissen	4	1'937.5	1'955.5	1'898.4	1'916.3
Vorwissen und Interventionen	5	1'938.3	1'960.8	1'899.1	1'921.5
Vorwissen, Interventionen und Arbeitsgedächtnis	6	---[a]	---[a]	1'881.2	1'908.1
Vorwissen, Interventionen und kognitive Fähigkeiten	6	---[a]	---[a]	1'871.6	1'898.5
Vorwissen, Interventionen und Sprachverständnis	6	1'927.5	1'954.6	1'890.4	1'917.3
Vorwissen, Interventionen und Ausdrucksfähigkeit	6	1'928.1	1'955.2	1'889.2	1'916.1
Vorwissen, Interventionen und Familiensprache	6	1'940.2	1'967.3	1'900.9	1'927.9
Vorwissen, Interventionen und Geschlecht	6	1'938.3	1'965.3	1'899.2	1'926.2
Vorwissen, Interventionen und Alter	6	1'929.4	1'956.4	1'888.9	1'915.8
Vorwissen, Interventionen und Lehrmittel	6	1'939.6	1'966.6	1'900.5	1'927.4

Anmerkungen. [a] Die Werte können aufgrund von fehlenden Werten nicht bestimmt werden.

Die Modellgüte-Masse AIC und BIC zeigen auf, dass durch die Hinzunahme weiterer Prädiktoren zum Teil Verbesserungen des Modells erreicht werden können. Sie zeigen aber auch

auf, dass vor allem die Verwendung des Vorwissens im Modell die Güte des Modells bestimmt.

Da bei insgesamt 12 der 669 Kinder aufgrund von Absenzen zum Testzeitpunkt T1 die Werte für das Arbeitsgedächtnis und zum Teil auch für die kognitiven Fähigkeiten nicht vorhanden waren, wurden die Modellvergleiche auch mit den 657 Kindern mit vollem Datensatz nochmals gerechnet. Die Ergebnisse beider Rechnungen können der Tabelle 57 entnommen werden.

Varianzaufklärung eines Modells mit allen 10 Prädiktoren

Verwendet man alle zehn Prädiktoren gleichzeitig, so kann mit dem Modell ein Anteil von R^2_m=46.3% bzw. R^2_c=51.2% der Gesamtvarianz aufgeklärt werden. Es zeigt sich also auch hier, dass der Anteil an der erklärten Varianz, den das Vorwissen alleine bereits liefert, auch mit neun weiteren Prädiktoren nicht mehr sehr stark gesteigert werden kann. Wählt man unter den verschiedenen Prädiktoren diejenigen aus, die sich bei einzelner Hinzunahme besonders bewährt haben (CFT, AG, Sprachverständnis, Ausdrucksfähigkeit und Alter) so ergibt sich (inkl. Vorwissen, aber ohne die Interventionen) ein Modell mit insgesamt sechs Prädiktoren und einer Varianzaufklärung von R^2_m=45.8% bzw. R^2_c=50.5%.

Insgesamt kann gefolgert werden, dass das Vorwissen einen sehr grossen Teil der Varianz beim Testzeitpunkt T4 aufklärt. Mit zusätzlichen Prädiktoren kann wenig gewonnen werden. Diejenigen Prädiktoren, die vielversprechend aussehen und einen signifikant von 0 unterschiedlichen β-Wert aufweisen, tun dies u.a. auch deshalb, weil sie mit dem Vorwissen korrelieren. Dies ist insbesondere bei den beiden Formen der Sprachfähigkeiten, beim Arbeitsgedächtnis und bei den kognitiven Fähigkeiten der Fall.

Die etwas überraschende Erkenntnis ist, dass mit dem Alter etwas zusätzliche Varianz erklärt werden kann. Der Lernzuwachs der jüngeren Kinder war im Laufe des 2. Schuljahrs etwas höher als derjenige der älteren Kinder. Nimmt man nur das Alter zusätzlich zum Vorwissen ins Modell auf, so ergibt sich ein erklärter Varianzanteil von immerhin R^2_m=42.1% bzw. R^2_c=47.2%.

Reduktion der Anzahl Prädiktoren mit Hilfe der «lasso»-Regression

Für die «lasso»-Regression wurde das Modell von Beginn weg mit allen zehn Prädiktoren (siehe oben) ausgestattet. Weil bei 12 Kindern die Ergebnisse des Arbeitsgedächtnisses und bei einigen von diesen 12 auch die Ergebnisse des CFT-Tests nicht vorlagen, wurde der Datensatz um diese 12 Kinder auf die verbleibenden 657 Kinder reduziert.

Die «lasso»-Regression geht von allen Prädiktoren aus und reduziert die β-Werte für die Prädiktoren so weit, wie es möglich ist, ohne den Bias allzu stark zu erhöhen. Dies kann so weit gehen, dass für die β-Werte die bestmögliche Schätzung 0 beträgt. Prädiktoren, bei welchen dies eintritt, können so als erklärende Variablen aus dem Modell eliminiert werden. Der Algo-

rithmus kennt keinerlei Theorie und Vorlieben für einzelne Prädiktoren und beschäftigt sich «blind» mit dem vorliegenden Datensatz.

Der Algorithmus ortet vier (der zehn) Prädiktoren, welche aus dem Modell entfernt werden können. Es sind dies die Prädiktoren Interventionen Ja/Nein, Schulbuch, Sprachverständnis und Familiensprache. Die Ergebnisse sind allerdings nicht stabil und hängen auch von der Partitionierung des Kinderbestands (Anteil der Kinder im jeweiligen Train- und im Test-Set und genaue Zufalls-Zuteilung zu den beiden Gruppen) für die Kreuzvalidierung ab. Es werden aber immer wieder die oben genannten vier Prädiktoren als redundant geortet, bei verschiedenen Durchläufen aber auch nur eine Teilmenge davon. Vier typische Ergebnisse der «lasso»-Regression sind in Tabelle 58 aufgeführt. Bei Durchlauf 4 werden alle Prädiktoren (z.T. mit sehr tiefen β-Werten) beibehalten.

Wichtigster Prädiktor bleibt auch hier das Vorwissen (Prädiktor 10) mit β-Werten im Bereich von 0.45. Das Alter zeigt sich auch hier als überraschend nützlicher Prädiktor mit negativem Vorzeichen, die übrigen vier Prädiktoren (Geschlecht, Ausdrucksfähigkeit, Arbeitsgedächtnis und kognitive Fähigkeiten) zeichnen sich durch tiefe Werte von β aus. Die tiefen β-Werte der meisten Prädiktoren deuten auch darauf hin, dass die Varianzaufklärung des Modells mit diesen Prädiktoren nicht wesentlich verbessert werden kann.

Tabelle 58: Schätzung für β mit «lasso»-Regression (λ gemäss Kreuzvalidierung)

Prädiktor	Bedeutung des Prädiktors	β-Werte nach «lasso»-Regression Durchlauf 1	β-Werte nach «lasso»-Regression Durchlauf 2	β-Werte nach «lasso»-Regression Durchlauf 3	β-Werte nach «lasso»-Regression Durchlauf 4
1	Interventionen Ja/Nein	---	---	-0.059	0.029
2	Geschlecht	0.224	0.206	0.100	0.233
3	Schulbuch	0.085	0.063	---	0.006
4	Alter bei T2	-0.354	-0.330	-0.348	-0.311
5	Sprachverständnis (dichotom)	---	---	---	0.051
6	Ausdrucksfähigkeit (dichotom)	0.192	0.195	0.203	0.303
7	Familiensprache (dichotom)	0.109	0.021	---	0.193
8	Kognitive Fähigkeiten	0.040	0.043	0.043	0.040
9	Arbeitsgedächtnis	0.086	0.078	0.060	0.078
10	Vorwissen (PF im BM bei T2)	0.446	0.449	0.461	0.455

«lasso»-Regression mit strengerer Selektion der Prädiktoren

Bei der «lasso»-Regression wird ein Parameter λ verwendet, welcher als Vorfaktor hohe Werte von β gegenüber der Quadratsumme der Residuen negativ gewichtet. Die optimale Wahl von λ wird durch ein Kreuzvalidierungsverfahren am vorhandenen Datensatz festgelegt. Die obigen vier Durchläufe wurden auf diese Art gerechnet.

Eine Möglichkeit, den Algorithmus etwas «strenger» bzw. «resoluter» werden zu lassen gegenüber Prädiktoren, die wenig beitragen, ist es, den vorgeschlagenen Wert des Parameters λ künstlich zu erhöhen. An dieser Stelle werden weitere vier Durchläufe der «lasso»-Regression angegeben. Bei diesen neuen Durchläufen ist der «optimale» Parameter λ um den multiplikativen Faktor 10 erhöht worden. Die Auswahl aus den obigen 10 Prädiktoren durch den Algorithmus geschieht so deutlich strenger.

Tabelle 59 verdeutlicht die Erkenntnis, dass auf die Interventionen, das Schulbuch, das Sprachverständnis und die Familiensprache als Prädiktoren verzichtet werden kann. Das Vorwissen zeigt sich auch bei dieser Betrachtung als der wichtigste Prädiktor. Beibehalten werden aber immer auch die kognitiven Fähigkeiten und das Arbeitsgedächtnis.

Tabelle 59: Schätzung für β mit «lasso»-Regression (λ mit 10 multipliziert)

Prädiktor	Bedeutung des Prädiktors	β-Werte nach «lasso»-Regression Durchlauf 5	β-Werte nach «lasso»-Regression Durchlauf 6	β-Werte nach «lasso»-Regression Durchlauf 7	β-Werte nach «lasso»-Regression Durchlauf 8
1	Interventionen Ja/Nein	---	---	---	---
2	Geschlecht	0.062	---	---	---
3	Schulbuch	---	---	---	---
4	Alter bei T2	-0.064	-0.045	-0.139	---
5	Sprachverständnis (dichotom)	---	---	---	---
6	Ausdrucksfähigkeit (dichotom)	0.107	0.013	0.052	---
7	Familiensprache (dichotom)	---	---	---	---
8	Kognitive Fähigkeiten	0.030	0.019	0.030	0.014
9	Arbeitsgedächtnis	0.055	0.046	0.036	0.006
10	Vorwissen (PF im BM bei T2)	0.437	0.419	0.447	0.397

Diese Sicht auf die verschiedenen Prädiktoren wird bestätigt, wenn man die standardisierten β-Werte betrachtet. Die Interpretation der β-Werte ist bei den standardisierten Werten schwieriger, dafür sind die Werte besser untereinander vergleichbar. In Tabelle 60 sind sowohl die standardisierten als auch die nicht-standardisierten β-Werte dargestellt. Erkennbar wird, dass nebst dem Vorwissen die kognitiven Fähigkeiten, das Arbeitsgedächtnis und an vierter Stelle das Alter (mit negativem Vorzeichen) den Lernzuwachs im zweiten Schuljahr erklären. Die anderen Prädiktoren tragen wenig zur Erklärung des Lernzuwachses bei.

Tabelle 60: Nicht-standardisierte und standardisierte β-Werte für das Modell mit allen 10 Prädiktoren

Prädiktor	Bedeutung des Prädiktors	Schätzung β-Wert	Std.fehler	p	β-Wert standardisiert
1	Interventionen Ja/Nein	-0.063	0.139	.654	-.020
2	Geschlecht	0.177	0.077	.022	.066
3	Schulbuch	0.058	0.120	.632	.021
4	Alter bei T2	-0.373	0.104	<.001	-.102
5	Sprachverständnis (dichotom)	0.142	0.159	.372	.041
6	Ausdrucksfähigkeit (dichotom)	0.234	0.138	.090	.076
7	Familiensprache (dichotom)	0.183	0.127	.151	.052
8	Kognitive Fähigkeiten	0.041	0.009	<.001	.143
9	Arbeitsgedächtnis	0.065	0.019	<.001	.113
10	Vorwissen (PF im BM bei T2)	0.464	0.031	<.001	.503

Mehrebenen-Modell mit 3 Ebenen anstelle von 2 Ebenen

In Kapitel 4.5.5 wurde erläutert, weshalb ein Modellansatz mit zwei Ebenen gewählt wurde. Würde man hingegen 3 Ebenen im Modell zulassen, so böten sich die Interventionsgruppen I, II, III, IV als «oberste» dritte Ebene an. In jeder dieser vier Gruppen befinden sich jeweils 11 oder 12 Klassen. Die Gruppen unterscheiden sich durch die Förderungen, welche die Kinder im ersten Schuljahr erhalten haben. Dass dieser Ansatz keine Verbesserung des Modells darstellt, kann auch den Werten in Tabelle 61 entnommen werden.

Tabelle 61: Modellgüte-Masse bei Hinzunahme einer dritten Ebene

Modell	df	AIC	BIC	logLik	Deviance	Chisq	Chi df	p(>Chisq)
2 Ebenen	5	1'938.3	1'960.8	-964.1	1'928.3			
3 Ebenen	6	1'940.3	1'967.3	-964.1	1'928.3	0	1	1

Die erklärten Varianzanteile bleiben mit R^2_m=41.1% und R^2_c=46.5% bei dem Modell mit drei Ebenen gegenüber demjenigen mit zwei Ebenen unverändert.

6.2 Längsschnittanalyse für das Operationsverständnis (gesamte Skala)

Ergänzend zur Hauptanalyse mit dem BASIS-MATH-Erhebungsinstrument wird in diesem Kapitel der Lernzuwachs beschrieben, der mit dem neuen Erhebungsinstrument für das Operationsverständnis erfasst werden konnte. Bei der Wirkung der Interventionen ist anzumerken, dass in den Fördereinheiten (Interventionen) intensiv mit denselben Fragen an die Kinder («Finde eine passende Rechnung. Welche Rechnung passt zum Bild? » usw.) gearbeitet wurde, die auch in den Erhebungsinstrumenten verwendet wurden. Ein Teil des Lernfortschritts der Kinder in den Interventionsklassen kann also sicherlich auch durch Gewöhnung der Kinder ans Itemformat und die im Erhebungsinstrument verwendete Art der Frage erklärt werden (siehe Diskussion in Kapitel 7.1). Auch in diesem Kapitel erfolgt eine Trennung zwischen Betrachtungen des gesamten Bestands (für die Betrachtungen zur Forschungsfrage FF1) und Teilbeständen (für die Betrachtungen zur Forschungsfrage FF2).

6.2.1 Lernfortschritt aller Kinder (OV)

Es wird mit demselben Mehrebenen-Modellansatz wie im vorangehenden Kapitel gearbeitet. Ausgewertet wird an dieser Stelle der Lernfortschritt der Kinder, gemessen an den Personenfähigkeiten, die sich aus den Erhebungsinstrumenten für das Multiplikationsverständnis (MV) zu den Zeitpunkten T3 und T4 ergeben.

Tabelle 62: Übersicht Lern-Fortschritt im Multiplikationsverständnis (alle 669 Kinder)

Gruppe	Einsatz FE 2. SJ	Anzahl Kinder	Mittlerer Testscore MV T2	Mittlerer Zuwachs MV	Mittlerer Zuwachs PF
I	Ja	165	15.8	1.8	-0.06
II	Ja	173	16.3	2.1	-0.04
III	Nein	167	17.4	2.0	-0.09
IV	Ja	164	15.5	2.6	0.19
alle		669	16.2	2.1	0.00

Das Vorwissen in Form der Personenfähigkeit zum Testzeitpunkt T3 zeigt sich auch hier als signifikanter Prädiktor (β_1=0.627, p<.001). Die Anwendung der Interventionen zeigt keinen signifikanten Effekt (β_2=-0.024, p=.876).

Tabelle 63: Ergebnisse der Mehrebenenanalyse (Multiplikationsverständnis, Gesamtbestand)

Fixe Effekte	Schätzung	Std.fehler	df	t-Wert	p	β standardisiert
Vorwissen	0.627	0.027	662.5	23.25	<.001	.666
Intervention	-0.024	0.155	42.7	-0.158	.876	-.017

Anmerkungen. Anzahl Kinder: 669. Anzahl Klassen: 46.

Die fixen Effekte (Vorwissen und Intervention) korrelieren mit einem Wert von .073.

Die Ergebnisse zeigen, dass für den Gesamtbestand kaum Unterschiede beim Lernfortschritt zwischen den Kindern der verschiedenen Interventionsgruppen bestehen.

Der ICC-Wert im Nullmodell für diese Auswertung beträgt 10.2%. Es werden also 89.8% der Varianz in den Kinderleistungen durch individuelle Effekte und lediglich 10.2% der Varianz durch die Zugehörigkeit zur jeweiligen Klasse erklärt.

Durch die Modellausgestaltung inklusive Wahl der beiden Prädiktoren Vorwissen und Intervention werden R^2_m=44.2% bzw. R^2_c=50.8% der Gesamtvarianz erklärt.

6.2.2 Lernfortschritt von Kindern mit unterschiedlicher Rechenleistung (OV)

Bei gleicher Einteilung der Kinder in die jeweiligen Quartile wie auch für die Analysen des Lern-Fortschritts beim BASIS-MATH wurde folgendes beobachtet:

Unteres Quartil (Gesamtbestand)

Tabelle 64: Übersicht Lern-Fortschritt im Multiplikationsverständnis (unteres Quartil des Gesamtbestands)

Gruppe	Einsatz FE 2. SJ	Anzahl Kinder	Mittlerer Testscore OV T3	Mittlerer Zuwachs OV	Mittlerer Zuwachs PF
I	Ja	53	10.9	1.5	-0.01
II	Ja	43	12.2	1.0	-0.16
III	Nein	28	11.8	1.3	-0.05
IV	Ja	39	10.7	2.3	0.21
alle		163	11.4	1.5	-0.01

Das Vorwissen in Form der Personenfähigkeit zum Testzeitpunkt T3 zeigt sich als signifikanter Prädiktor (β_1=0.658, p<.001). Die Anwendung der Interventionen zeigt keinen signifikanten Effekt (β_2=0.090, p=.723).

Tabelle 65: Ergebnisse (Multiplikationsverständnis, unteres Quartil des Gesamtbestands)

Fixe Effekte	Schätzung	Std.fehler	df	t-Wert	p	β standardisiert
Vorwissen	0.658	0.069	159.8	9.493	<.001	.611
Intervention	0.090	0.253	53.1	0.356	.723	.069

Anmerkungen. Anzahl Kinder: 163. Anzahl Klassen: 43.

Die fixen Effekte (Vorwissen und Intervention) korrelieren mit einem Wert von .017.

Der beobachtete Lernzuwachs der Kinder zeigt sich unter Anwendung der Interventionen in diesem Teilbestand als (sehr leicht) höher als ohne die Anwendung der Interventionen. Das Ergebnis ist statistisch aber nicht signifikant und kann durchaus zufällig entstanden sein. Der Nachweis der statistischen Signifikanz wird − wie bereits in Kapitel 6.1.2 erläutert – durch die kleine Kontrollgruppe erschwert. Es erfolgen darum dieselben weiteren Rechnungen analog zu Kapitel 6.1.2.

Durch die Modellausgestaltung inklusive Wahl der beiden Prädiktoren Vorwissen und Intervention werden R^2_m=35.8% bzw. R^2_c=44.4% der Gesamtvarianz in diesem Modell erklärt.

Unteres Quartil (pro Gruppe)

Tabelle 66: Übersicht Lern-Fortschritt Multiplikationsverständnis (unteres Quartil jeder Int.gruppe)

Gruppe	Einsatz FE 2. SJ	Anzahl Kinder	Mittlerer Testscore OV T3	Mittlerer Zuwachs OV	Mittlerer Zuwachs PF
I	Ja	41	10.0	1.7	0.04
II	Ja	43	12.2	1.0	-0.16
III	Nein	42	12.6	1.8	0.01
IV	Ja	41	11.0	2.2	0.20
alle		167	11.5	1.7	0.02

Das Vorwissen in Form der Personenfähigkeit zum Testzeitpunkt T3 zeigt sich als signifikanter Prädiktor (β_1=0.636, p<.001). Die Anwendung der Interventionen zeigt keinen signifikanten Effekt (β_2=-0.066, p=.782).

Tabelle 67: Ergebnisse (Multiplikationsverständnis, unteres Quartil jeder Interventionsgruppe)

Fixe Effekte	Schätzung	Std.fehler	df	t-Wert	p	β standardisiert
Vorwissen	0.636	0.070	163.9	9.051	<.001	.585
Intervention	-0.066	0.238	39.8	-0.279	.782	-.049

Anmerkungen. Anzahl Kinder: 167. Anzahl Klassen: 44.

Die fixen Effekte (Vorwissen und Intervention) korrelieren mit einem Wert von .090.

Der Lernzuwachs scheint in der Kontrollgruppe leicht höher gewesen zu sein als in der Interventionsgruppe. Das Ergebnis ist allerdings statistisch ebenfalls nicht signifikant.

Durch die Modellausgestaltung inklusive Wahl der beiden Prädiktoren Vorwissen und Intervention werden R^2_m=33.6% bzw. R^2_c=42.3% der Gesamtvarianz in diesem Modell erklärt.

Die «rechenstarken» Kinder «oberhalb» des unteren Quartils (Gesamtbestand)

Tabelle 68: Übersicht Lern-Fortschritt im Multiplikationsverständnis (stärkere 75% des Gesamtbestands)

Gruppe	Einsatz FE 2. SJ	Anzahl Kinder	Mittlerer Testscore OV T3	Mittlerer Zuwachs OV	Mittlerer Zuwachs PF
I	Ja	112	18.1	2.0	-0.08
II	Ja	130	17.6	2.5	0.01
III	Nein	139	18.5	2.1	-0.10
IV	Ja	125	16.9	2.7	0.18
alle		506	17.8	2.3	0.00

Das Vorwissen in Form der Personenfähigkeit zum Testzeitpunkt T3 zeigt sich als signifikanter Prädiktor (β_1=0.483, p<.001). Die Anwendung der Interventionen zeigt keinen signifikanten Effekt (β_2=-0.043, p=.768).

Tabelle 69: Ergebnisse (Multiplikationsverständnis, stärkere 75% des Gesamtbestands)

Fixe Effekte	Schätzung	Std.fehler	df	t-Wert	p	β standardisiert
Vorwissen	0.483	0.032	501.5	14.861	<.001	.551
Intervention	-0.043	0.145	44.3	-0.296	.7685	-.036

Anmerkungen. Anzahl Kinder: 506. Anzahl Klassen: 46.

Die fixen Effekte (Vorwissen und Intervention) korrelieren mit einem Wert von .074.

Die Ergebnisse deuten darauf hin, dass sich die Interventionen in diesem Teilbestand und bei der Verwendung dieses Erhebungsinstruments kaum auf den Lernfortschritt der Kinder ausgewirkt hatten. Das Ergebnis ist statistisch nicht signifikant.

Durch die Modellausgestaltung inklusive Wahl der beiden Prädiktoren Vorwissen und Intervention werden R^2_m=30.3% bzw. R^2_c=37.0% der Gesamtvarianz in diesem Modell erklärt.

Die «rechenstarken» Kinder «oberhalb» des unteren Quartils (pro Gruppe)

Tabelle 70: Übersicht Lern-Fortschritt Multiplikationsverständnis (stärkere 75% jeder Int.gruppe)

Gruppe	Einsatz FE 2. SJ	Anzahl Kinder	Mittlerer Testscore OV T3	Mittlerer Zuwachs OV	Mittlerer Zuwachs PF
I	Ja	124	17.7	1.9	-0.09
II	Ja	130	17.6	2.5	0.01
III	Nein	125	19.0	2.1	-0.13
IV	Ja	123	17.0	2.7	0.19
alle		502	17.8	2.3	-0.01

Das Vorwissen in Form der Personenfähigkeit zum Testzeitpunkt T3 zeigt sich als signifikanter Prädiktor (β_1=0.492, p<.001). Die Anwendung der Interventionen zeigt keinen signifikanten Effekt (β_2=-0.111, p=.477).

Tabelle 71: Ergebnisse (Multiplikationsverständnis, stärkere 75% jeder Interventionsgruppe)

Fixe Effekte	Schätzung	Std.fehler	df	t-Wert	p	β standardisiert
Vorwissen	0.492	0.032	494.8	15.207	<.001	.561
Intervention	-0.110	0.154	45.8	-0.718	.477	-.094

Anmerkungen. Anzahl Kinder: 502. Anzahl Klassen: 46.

Die fixen Effekte (Vorwissen und Intervention) korrelieren mit einem Wert von .105.

Auch hier sind die Schätzwerte, welche den Einfluss der Interventionen messen, negativ und so nahe bei 0, dass Zufallseffekte nicht ausgeschlossen werden können. Das Ergebnis ist statistisch nicht signifikant.

Durch die Modellausgestaltung inklusive Wahl der beiden Prädiktoren Vorwissen und Intervention werden R^2_m=31.9% bzw. R^2_c=39.8% der Gesamtvarianz in diesem Modell erklärt.

Fazit

Eine Unterschiedlichkeit zwischen der Interventions- und der Kontrollgruppe konnte bei den Erhebungen für das Operationsverständnis nicht beobachtet werden. Dies überrascht insofern, als sich die Förderung des Operationsverständnisses eigentlich eher bei der Erhebung des Operationsverständnisses als bei denjenigen mit dem BASIS-MATH-Instrument zeigen sollte.

Auch der beim BASIS-MATH beobachtete Unterschied zwischen den Stärkeklassen konnte mit diesem Instrument nicht nachgewiesen werden. Die p-Werte sämtlicher Aufteilungen in Stärkeklassen sind hoch und lassen die gemachten Beobachtungen unter der Nullhypothese «es gibt keinen Unterschied zwischen der Kontroll- und den Interventionsgruppen» als plausibel erscheinen.

Auch eine Trennung der «schwächeren» 20% anstelle von 25% ändert nichts an dieser Situation. Ebenso lassen die verschiedenen Aufteilungen in eine «schwächere» und eine «stärkere» Hälfte der 669 Kinder keine signifikante Wirkung der Interventionen (gemessen mit diesem Erhebungsinstrument) nachweisen. Wie auch bei den oben aufgeführten Ergebnissen bleiben die β_2-Werte bei den zwei Aufteilungen in eine «schwächere» und eine «stärkere» Hälfte sehr tief und die entsprechenden p-Werte hoch. Auf die Angabe der gesamten Tabellen wird an dieser Stelle der Übersicht halber verzichtet. Eine Kurzübersicht kann Tabelle 72 entnommen werden.

Tabelle 72: Ergebnisübersicht Multiplikationsverständnis bei Einteilung in zwei „Hälften“

Aufteilung in zwei Hälften (Split bei Testscore)	Stärkegruppe	Anzahl Kinder	β_2-Wert (Einfluss der Intervention)	p-Wert	β_2-Wert standardisiert
≤ 21	«Schwächere Hälfte»	317	-0.011	.950	-.009
> 21	«Stärkere Hälfte»	352	-0.035	.824	-.031
≤ 22	«Schwächere Hälfte»	357	-0.008	.966	-.006
> 22	«Stärkere Hälfte»	312	-0.107	.496	-.097

6.3 Längsschnittanalyse für das Operationsverständnis (Brückenitems)

Unter den Items des Vor- und des Nachtests gab es 12 Items, welche unverändert übernommen wurden. Diese 12 Brückenitems erlauben eine weitere Längsschnittmessung mit einer auf diese Items reduzierten Skala. Auch ermöglichen es die Brückenitems, DIF über die Zeit zu orten. Diese DIF-Analyse erfolgt in Kapitel 6.3.1, worauf die Ergebnisse der Längsschnittanalyse anhand aller Kinder im folgenden Kapitel 6.3.2 wiedergegeben werden.

Natürlicherweise sind die Ergebnisse mit denjenigen aus Kapitel 6.2 verwandt, da es sich bei den 12 Brückenitems um eine Auswahl aus den Items des dort verwendeten Erhebungsinstruments handelt. Die Brückenitems bilden rund die Hälfte des gesamten Erhebungsinstruments (12 von 24 verwendeten Item im Vortest, 12 von 26 verwendeten Items im Nachtest). Deren Bearbeitung korreliert in sehr hohem Masse mit der Bearbeitung des gesamten Tests.

Die Personenfähigkeiten im Rasch-Modell korrelieren mit .95 für den Vortest und .87 für den Nachtest zwischen dem gesamten Erhebungsinstrument und der auf die Brückenitems reduzierten Version des Erhebungsinstruments.

Auf die Betrachtung von Teilbeständen für die Brückenitems wird verzichtet. Die Begründung dafür erfolgt in Kapitel 6.3.2.

6.3.1 DIF-Analyse der 12 Brückenitems im Längsschnitt

Die DIF-Analyse im Längsschnitt erlaubt es, zu untersuchen, ob eines oder mehrere der 12 Brückenitems eine andere zeitliche Entwicklung aufweist als die anderen. Wie schon bei DIF nach anderen Kriterien (z.B. Geschlecht oder Sprachverständnis) ist DIF alleine kein Ausschlusskriterium für eines der Items aus der Skala. DIF kann aber als Anregung dazu dienen, zu überlegen, weshalb ein Item in der zeitlichen Entwicklung ein anderes Verhalten zeigt als die anderen.

Die Ortung von DIF über die Zeit erfolgt mit denselben Verfahren, die in Kapitel 4.5.3 erläutert und in den Kapiteln 5.4.2 und 5.5.2 auf die Bestände des Vor- und Nachtests angewandt wurden.

Der Bestand, der an dieser Stelle verwendet wird, ist derjenige der 680 Kinder aus 46 Klassen, welcher auch für alle anderen Längsschnittanalysen herangezogen wird. Die Ergebnisse der DIF-Analysen sind wie folgt:

Andersen Likelihood-Ratio-Test

Der Andersen Likelihood-Ratio-Test verwirft die Nullhypothese, dass die Gesamtheit der 12 Items sich über die Zeit gleich entwickelt (LR-Wert = 43.25, df=21, $p<.001$).

Wald-Test

Der Wald-Test ortet DIF über die Zeit bei einigen Items. Auf dem 5%-Niveau signifikant wird das DIF bei den Items Stifte» ($p<.001$), «Auto a: Frage nach der Zahl» ($p=.019$), «Bücher a: Frage nach der Zahl» ($p=.020$) und «Tische b: Frage nach der Rechnung» ($p=.007$).

Graphische Überprüfung des Item-Fits

Dieselben vier Items wie beim Wald-Test werden auch bei einer graphischen Überprüfung des DIF auffällig.

Mantel-Haenszel Verfahren

Mit diesem Verfahren wird DIF geortet bei den Items «Stifte», «Bücher a», «Korb a», «Auto a» und «Tische b». Die ersten drei erhalten für den Effekt die Bewertung B, die anderen beiden die Bewertung A.

Weitere Verfahren

Vergleich der Verfahren «T.I.D», «MH», «Std.» und «SIBTEST» mit dem R-Package difR: Die Items mit DIF bei zwei Verfahren sind «Stifte», «Bücher a», «Korb a» und «Tische b».

Kurze Diskussion

Die DIF-Analyse sollte nicht zu weiterem Ausschluss von Items verwendet werden. Es erscheint vielmehr plausibel, dass die Items «Stifte» und «Auto» als die einzigen beiden Items, welche die Subtraktion bedienen, beim Fortschritt auffällig werden. Diese beiden Items passen zu Grundoperationen, welche bereits erlernt und im 2. Schuljahr nur noch gefestigt werden, während alle anderen Items zur neu erlernten Grundoperation der Multiplikation gehören und deshalb im 2. Schuljahr einen anderen (vermutlich höheren) Lernfortschritt mit sich bringen.

Bei Item «Tische» kann das DIF durch die strenge Codierung für dieses Item erklärt werden. Im Gegensatz zu anderen Items muss dieses Item mit einer Antwort bearbeitet werden, die zur Struktur des Bilds passt. Es ist anzunehmen, dass die Kinder in dieser Hinsicht im 2. Schuljahr dazu gelernt haben und dieses Item deshalb besser bearbeitet wird. «Bücher» hingegen ist auch am Ende des 2. Schuljahrs ein kniffliges Item, das hohe Aufmerksamkeit verlangt und häufig (falsch) mit $8 = 4 + 4$ bearbeitet wurde. An diesem Stolperstein hat sich auch im Laufe des 2. Schuljahrs wenig geändert.

Die 12 Items werden für die Erfassung des Lernfortschritts für den Gesamtbestand der Kinder im folgenden Kapitel beibehalten.

6.3.2 Lernfortschritt aller Kinder (Brückenitems)

Auf die Unterscheidung von Kindern in unterschiedlichen Stärkeklassen wird an dieser Stelle verzichtet. Grund dafür ist die Tatsache, dass die Brückenitems sich im Nachtest – wie in Kapitel 5.1.2 bereits berichtet − als verhältnismässig leicht herausstellten.

Sehr viele Kinder erreichten unter den 12 Brückenitems im Nachtest bei T4 hohe Testscores. Insgesamt 176 von 669 Kinder erreichten im Nachtest den höchsten möglichen Testscore von 12, weitere 110 Kinder einen solchen von 11 und 107 Kinder einen solchen von 10 (siehe Abbildung 33). Die Brückenitems waren vom Schwierigkeitsgrad her im Vortest angemessen, erwiesen sich dann im Nachtest aber als zu leicht. Vermutlich kamen bei diesen Items zusätzlich auch Erinnerungseffekte an den Vortest, der 9 Monate vorher abgelegt wurde, zum Tragen.

Diese Situation akzentuiert sich zusätzlich, wenn man nur die «stärkeren» Kinder betrachtet. Aus diesem Grund wird an dieser Stelle auf weitere Analysen mit Teilbeständen der 669 Kinder verzichtet. Die Brückenitems sind als Teil des gesamten Erhebungsinstruments in die Ergebnisse in Kapitel 6.2 eingeflossen. Eine detailliertere Analyse mit ausschliesslich Brückenitems erscheint angesichts dieses Deckeneffekts aber für Teilbestände nicht als sinnvoll, da dieser Teil des Testinstruments nicht optimal differenziert. Der Vollständigkeit halber werden die Ergebnisse der Analysen für den Gesamtbestand der 669 Kinder im Folgenden aber dargestellt. Auch diese Ergebnisse sind aufgrund des Deckeneffekts mit etwas Vorsicht zu betrachten.

Lernfortschritt für den Gesamtbestand

Es wird derselbe Mehrebenen-Modellansatz wie in den vorherigen Kapiteln gewählt.

Tabelle 73: Übersicht Lern-Fortschritt bei den Brückenitems im Operationsverständnis (Gesamtbestand)

Gruppe	Einsatz FE 2. SJ	Anzahl Kinder	Mittlerer Testscore OV BI T3	Mittlerer Zuwachs OV BI	Mittlerer Zuwachs PF
I	Ja	165	6.5	2.4	-0.11
II	Ja	173	6.8	2.6	-0.03
III	Nein	167	7.6	2.5	-0.04
IV	Ja	164	6.3	2.9	0.19
alle		669	6.8	2.6	0.00

Das Vorwissen in Form der Personenfähigkeit zum Testzeitpunkt T3 zeigt sich als signifikanter Prädiktor (β_1=0.553, p<.001). Die Anwendung der Interventionen zeigt keinen signifikanten Effekt (β_2=-0.189, p=.175).

Tabelle 74: Ergebnisse (Multiplikationsverständnis, nur Brückenitems, Gesamtbestand)

Fixe Effekte	Schätzung	Std.fehler	df	t-Wert	p	β standardisiert
Vorwissen	0.553	0.028	666.0	19.770	<.001	.608
Intervention	-0.189	0.137	39.4	-1.382	.175	-.132

Anmerkungen. Anzahl Kinder: 669. Anzahl Klassen: 46.

Die fixen Effekte (Vorwissen und Intervention) korrelieren mit einem Wert von .108.

Die Ergebnisse zeigen, dass für den Gesamtbestand kaum Unterschiede bestehen beim Lernfortschritt der Kinder. Das schwache und nicht signifikante Ergebnis deutet darauf hin, dass die Kinder in der Kontrollgruppe mehr dazu gelernt haben als jene in den Interventionsgruppen, ohne dass jedoch der Zufall als erklärender Effekt ausgeschlossen werden kann.

Durch die Modellausgestaltung inklusive Wahl der beiden Prädiktoren Vorwissen und Intervention werden R^2_m=38.3% bzw. R^2_c=41.8% der Gesamtvarianz erklärt.

Der ICC-Wert für das Nullmodell gemessen mit den 12 Brückenitems beträgt 8.7%.

7 Diskussion und Ausblick

Kapitel 7.1 beinhaltet die Diskussion der Ergebnisse, welche in den Kapiteln 5 und 6 berichtet wurden. In Kapitel 7.2 folgt ein Ausblick auf allfällige weitere Forschungsarbeiten.

7.1 Diskussion

Diskutiert werden in Kapitel 7.1.1 zuerst die Ergebnisse der Analyse des neuen Erhebungsinstruments, sodann in Kapitel 7.1.2 die Resultate der Längsschnittanalyse. Das abschliessende Kapitel 7.1.3 umfasst einige weitere Beobachtungen, die im Laufe dieser Arbeit gemacht wurden und erwähnenswert erscheinen.

7.1.1 Erhebungsinstrument für das Multiplikationsverständnis

Im Rahmen dieser Arbeit wurde für den Einsatz im Forschungsprojekt ein neues Erhebungsinstrument für das Multiplikationsverständnis entwickelt. Die Entwicklungsarbeiten und die Diskussion der Testgütekriterien für dieses Instrument sind in Kapitel 4.4 dieser Arbeit erfolgt. Dasselbe Kapitel umfasst auch Bemerkungen zu den Einschränkungen, die hinsichtlich des Geltungsbereichs des neuen Erhebungsinstruments gemacht werden müssen (siehe Kapitel 4.4.4).

Zusammenfassend kann gesagt werden, dass mit dem neuen Erhebungsinstrument ein valides Testinstrument vorliegt, mit dem das Multiplikationsverständnis auch für eine grosse Anzahl von Kindern in zufriedenstellender Art und Weise erfasst werden kann. Es kann somit als Vorlage für die Weiter- oder Neuentwicklung ähnlicher Instrumente herangezogen werden. Die Erfahrungen, die bei der Entwicklung des Instruments gemacht wurden, sind in Kapitel 4.4 ausführlich beschrieben. Auch auf einige denkbare Stolpersteine wird dort ausdrücklich hingewiesen.

Zufriedenstellend sind insbesondere der Item-Schwierigkeitsgrad im Vortest sowie (mit Ausnahme der Brückenitems) auch derjenige im Nachtest. Ebenfalls zufriedenstellend sind die weiteren Kennzahlen wie z.B. die verschiedenen Reliabilitätsmasse, die Item-Trennschärfen oder auch die Passung der einzelnen beibehaltenen Items ins Rasch-Modell. Das Ausmass an DIF der einzelnen Items war derart, dass die meisten Items für die Auswertungen beibehalten werden konnten.

Das nun vorliegende Erhebungsinstrument war allerdings in mancherlei Hinsicht sehr spezifisch auf die Anforderungen im beschriebenen Forschungsprojekt zugeschnitten. Die Erhebung des Multiplikationsverständnisses vor Einführung der Malrechnung in den getesteten Klassen war beispielsweise eine der grösseren Herausforderungen bei der Entwicklung dieses Instruments. Dessen Wiederverwendung unter anderen Bedingungen wäre deshalb mit Schwierigkeiten bzw. Gefahren verbunden und sollte vorsichtig und gegebenenfalls unter Anpassung einzelner Items erfolgen.

Bedauerlicherweise stellte sich das Instrument in der Längsschnittanalyse als nicht sehr sensitiv bezüglich Unterschiedlichkeit der verschiedenen Interventionsgruppen heraus. Diese Schwäche des Instruments wäre bei einer allfälligen Replikation ebenfalls im Auge zu behalten.

Eine interessante Erkenntnis aus den Analysen mit dem neuen Erhebungsinstrument ist aber die Beobachtung, dass sich verschiedene theoretische Aspekte zum Operationsverständnis auch rein quantitativ zu zeigen scheinen. Zum einen scheinen die Ergebnisse der Erhebungen bei einer Faktoranalyse nach der Grundrechenart in einzelne Faktoren zu zerfallen. Zum anderen kann man die Ergebnisse der Faktoranalyse auch so interpretieren, dass sie die unterschiedliche Bearbeitung der Kinder bei verschiedenen inhaltlichen Grundvorstellungen belegen. Auch laden die Items, bei denen Punktefelder mit Malrechnungen in Zusammenhang gebracht werden mussten, auf einen anderen Faktor als die übrigen. Dies könnte als leichte quantitative Evidenz für das theoretische Konstrukt der Strukturierungsfähigkeit und dessen Rolle bei der Beschreibung von Operationsverständnis betrachtet werden. Um diese Vermutungen mit Sicherheit zu belegen, wäre aber eine weitere Arbeit mit ähnlichen Instrumenten notwendig. Eine Querschnittanalyse würde hierfür genügen; mehrere Messzeitpunkte wären dazu nicht erforderlich. Notwendig wären aber weitere Items zu unterschiedlichen inhaltlichen Grundvorstellungen bzw. weitere Items mit Punktefeldern und dem Ziel, die Strukturierungsfähigkeiten der Kinder zu erfassen. Möchte man belegen, dass die beobachteten Effekte tatsächlich mit dem theoretischen Konzept der Grundvorstellungen bzw. dem Konstrukt Strukturierungsfähigkeit zusammenhängen, so wäre eine breiter abgestützte Beobachtung erforderlich. Die beiden Items zur zeitlich-sukzessiven Grundvorstellung (zur Ausgangssituation «Korb») müssten durch weitere Items zur selben Grundvorstellung ergänzt werden. Ähnliches kann über die Items zu den Punktefeldern gesagt werden. Es sollten weitere solche Items ins Erhebungsinstrument aufgenommen werden. Auch sollte deren Itemformat demjenigen der anderen Items entsprechen, um auszuschliessen, dass das Itemformat für die anders gelagerten Faktorladungen ausschlaggebend ist.

7.1.2 Längsschnittanalysen

In dieser Arbeit wurde der Lernzuwachs durch den Einsatz von Fördereinheiten für das Operationsverständnis untersucht. Die Ergebnisse dieser Analyse wurden in Kapitel 6 präsentiert. Die Längsschnittanalyse wurde hauptsächlich mit dem standardisierten Testinstrument BASIS-MATH vorgenommen, da vor allem der Zuwachs bei den allgemeinen Rechenleistungen von Interesse war. Ergänzend wurde der Lernzuwachs auch mit dem neuen Erhebungsinstrument für das Multiplikationsverständnis erhoben.

Effekte bei den allgemeinen Rechenleistungen

Es zeigten sich bei den allgemeinen Rechenleistungen im Durchschnitt über die ganze Klasse keine Effekte bzw. keine Unterschiedlichkeit zwischen den verschiedenen Interventionsgruppen einerseits und der Kontrollgruppe andererseits. Hingegen gab es Unterschiede bei der

Leistungsentwicklung bei Trennung der Kinder in eine «schwache» bzw. eine «starke» Gruppe. Der Einsatz der Fördereinheiten hat sich für die rechenschwachen Kinder vorteilhaft ausgewirkt. Die rechenstarken Kinder haben jedoch vom «normalen» Unterricht in der Kontrollgruppe mehr profitiert als vom Unterricht mit den Fördereinheiten in den Interventionsgruppen. Die meisten dieser Effekte sind nicht sehr stark ausgeprägt. In dieser Arbeit ist es nur an einzelnen Stellen gelungen, die Signifikanz dieser Effekte auf dem 5%-Niveau nachzuweisen. Statistisch signifikant belegt wurde lediglich die Beobachtung, dass die stärkere Hälfte der Kinder in der Kontrollgruppe den höheren Lernzuwachs hatte als in den Interventionsgruppen. Aussagen zu den stärkeren Teilen der Klasse unterliegen aber immer der Randbemerkung, dass das eingesetzte Instrument BASIS-MATH vor allem auf die schwächeren Kinder ausgerichtet ist und daher besonders bei diesen Kindern optimal differenziert.

Covid-19-Situation

Zudem müssen noch einmal in der Diskussion der Ergebnisse an dieser Stelle die sehr speziellen Umstände des Schuljahrs 2019/20 aufgrund der Schulschliessungen zwischen März und Mai 2020 erwähnt werden. Dieses Schuljahr war das Beobachtungsjahr, welches den meisten Analysen dieser Arbeit zugrunde lag. Es ist offensichtlich, dass in diesem Zeitraum durch die Covid-19-Situation verschiedene Effekte auf den Schulbetrieb und auch auf den Lernzuwachs der Kinder eingewirkt haben. Einige dieser Effekte haben möglicherweise sogar eine stärkere Auswirkung als die eingesetzten 16 Fördereinheiten. Erwähnt seien an dieser Stelle noch einmal die in der Einleitung angeschnittene unterschiedliche Qualität des Unterrichts zu Hause, die unterschiedliche Ausgestaltung des Fernunterrichts und die kantonalen Unterschiede nach Wiederöffnung der Schulen am 11. Mai 2020. Im Weiteren haben die Schulschliessungen dazu geführt, dass die Fördereinheiten nicht in allen Klassen vollständig bearbeitet werden konnten. Auch der Lernfortschritt der rechenschwächeren Kinder, der in dieser Arbeit besondere Aufmerksamkeit erhalten hat, ist von den Schulschliessungen betroffen. Es wird davon ausgegangen, dass sich die zwei Monate ohne Unterricht für die schwächeren Kinder in besonders starkem Ausmass negativ ausgewirkt haben. Auch dieser Einfluss schlägt sich in den Messungen der vorliegenden Arbeit nieder. Der in dieser Arbeit gemessene Lernzuwachs unterliegt daher immer dem Vorbehalt, dass das Schuljahr 2019/20 ein sehr aussergewöhnliches war. Allerdings muss auch angemerkt werden, dass die verschiedenen Klassen alle gleichermassen von der Schulschliessung betroffen waren und dass die Schulschliessungen in allen Kantonen der Schweiz gleich lange gedauert haben (nämlich acht Wochen abzüglich zwei Wochen Frühlingsferien).

Trotz der Covid-19-Situation ist es im Frühjahr 2020 gelungen, eine ausreichende Anzahl Klassen für die Schlusserhebungen zu gewinnen und einige Effekte zu orten, auch wenn diese mehrheitlich nicht auf einem signifikanten Niveau nachgewiesen werden konnten. In welchem Ausmass diese Effekte durch die Covid-19-Situation beeinflusst wurden, lässt sich an dieser Stelle nicht beantworten.

Weitere Einflussfaktoren

Nebst der Covid-19-Situation gibt es weitere Einflussfaktoren, die sich in den Werten dieser Arbeit niederschlagen. Die Lehrpersonen waren eine sicherlich sehr motivierte, belastbare und an fachdidaktischen Fragen interessierte Gruppe. Diese Tatsache lässt − zusammen mit dem Einblick in die einzelnen Protokolle − vermuten, dass die Umsetzungstreue der Fördereinheiten sehr hoch war. Vermutlich tragen auch die Auswahl der Lehrpersonen und die Qualität ihres Unterrichts positiv zum Lerneffekt der Kinder bei. Viele Lehrpersonen haben angegeben, sie seien kooperative Arbeitsformen gewohnt. Es ist anzunehmen, dass auch verschiedene Klassen der Kontrollgruppe im Schuljahr 2019/20 in kooperativen Settings gearbeitet haben, was die beobachteten Unterschiede zwischen Interventions- und Kontrollgruppen vermindert haben könnte.

Möglich ist es aber auch, dass die Einführungsveranstaltungen zu den Fördereinheiten dazu geführt haben, dass die Lehrpersonen der Interventionsgruppe sich intensiver mit dem Themenkreis Operationsverständnis befasst haben als jene der Kontrollgruppe, welche nicht auf dieselbe Art und Weise mit der Thematik konfrontiert wurden. Somit kann nicht ausgeschlossen werden, dass Fachwissen und fachdidaktisches Wissen der Lehrpersonen «auf Umwegen» in die Studienergebnisse mit eingeflossen sind und einen Teil der Effekte in den Interventionsgruppen erklären.

Effekte beim Multiplikationsverständnis

Im Gegensatz zur Messung mit dem standardisierten Messinstrument BASIS-MATH für die allgemeinen Rechenleistungen war es nicht möglich, mit dem Erhebungsinstrument für das Multiplikationsverständnis signifikante Unterschiede zwischen der Interventions- und der Kontrollgruppe zu orten. Beobachtet wurden durchgehend sehr tiefe Werte für β und hohe für p, was die Nullhypothese «keine Wirkung der Interventionen» (bei Messung mit diesem Instrument) plausibel erscheinen lässt.

Diese Beobachtungen sind insofern erstaunlich, als dass das Erhebungsinstrument (bis auf die Unterschiedlichkeit bei den Übersetzungsrichtungen) genau diejenigen Fähigkeiten erfasst, die im Laufe des Schuljahrs in den Interventionsklassen eingeübt wurden.

Das Messinstrument für das Multiplikationsverständnis scheint für diese Art der Gruppenvergleiche zu wenig sensitiv zu sein. Diese Schwäche des Erhebungsinstruments müsste bei einer allfälligen weiteren Verwendung oder Weiterentwicklung beachtet werden.

Vergleich und Besprechung der Effekte bei BASIS-MATH bzw. Multiplikationsverständnis

Eine vertiefte Analyse des Lernfortschritts bei BASIS-MATH zeigt auf, dass dieses Instrument die Differenzen zwischen den Interventionsklassen bei verschiedenen Themen der Grundschulmathematik – und insbesondere nicht nur beim Operationsverständnis und den Grundvorstellungen – ortet. BASIS-MATH diagnostiziert unterschiedliche Lernfortschritte auch bei den Themen «Dezimales Stellenwertsystem», «Zahl zerlegen» (d.h. bei Rechenauf-

gaben des Typs «70 = 15 + ?»), «Zählen in Schritten» und «Sachrechnen». Nebst der zu geringen Sensitivität des Erhebungsinstruments für das Multiplikationsverständnis kann also davon ausgegangen werden, dass die Förderung auch bei Themenbereichen gewirkt hat, welche nicht explizit gefördert worden waren. Mit dem Instrument BASIS-MATH werden solche Themen erfasst, mit dem Instrument für das Multiplikationsverständnis hingegen nicht.

Das Erhebungsinstrument für das Multiplikationsverständnis war aus den dargelegten Gründen auf eine Übersetzungsrichtung («zum Term hin») beim Darstellungswechsel eingeschränkt. Möglicherweise erklärt auch diese Einschränkung die geringe Sensitivität des Erhebungsinstruments. Denkbar wäre, dass vor allem diejenige Übersetzungsrichtung, welche nur in den Fördereinheiten, nicht aber im Erhebungsinstrument für das Multiplikationsverständnis vorkam («vom Term weg», mit Aufforderungen an die Kinder in der Art von «Erzähle eine Geschichte, die zur Rechnung $4 \cdot 5$ passt.»), den Kindern hilfreiche denkerische Anstösse gegeben hat. Diese könnten zum Verständnis der Kinder und letztlich auch zu ihren allgemeinen Rechenfähigkeiten (nachgewiesen mit BASIS-MATH) beigetragen haben, ohne dass sie vom Erhebungsinstrument für das Multiplikationsverständnis erfasst werden.

Eine weitere Möglichkeit, die geringen Effekte der Interventionen zu erklären, wäre die Berücksichtigung der Fähigkeiten der Lehrpersonen. Es ist davon auszugehen, dass mehrheitlich überdurchschnittlich motivierte, belastbare und an fachdidaktischen Fragen interessierte Lehrpersonen am Projekt teilgenommen haben. Man kann annehmen, dass diese den Kindern bereits im Normalunterricht guten Unterricht bieten, der sich durch Interventionen nicht leicht verbessern lässt. Es kann also sein, dass die Kontrollgruppe, in welcher die Lehrpersonen im zweiten Schuljahr so unterrichten konnten, wie sie es gewohnt sind, eine besonders starke Vergleichsgruppe darstellte. In den Interventionsgruppen mussten die Lehrpersonen bei der Verwendung der Fördereinheiten einem Skript folgen, das sie zuerst ausführlich studieren mussten und welches auch eine ungewohnte Umstellung ihres Unterrichts zur Folge hatte. Die Rückmeldungen zu den Fördereinheiten waren zwar durchwegs positiv und die Umsetzung scheint den Lehrpersonen gelungen zu sein. Trotzdem kann es sein, dass der «normale» Unterricht in den Klassen der Kontrollgruppen effizienter und vor allem für die stärkeren Kinder optimaler ablief als derjenige mit den Fördereinheiten. Einzelne Rückmeldungen von Lehrpersonen zu den Fördereinheiten gingen auch in die Richtung, dass sie den Schwierigkeitsgrad des Materials für starke Kinder als zu tief einschätzten.

Denkbar wäre auch, dass sich vor allem das kooperative Setting positiv auf den Lernfortschritt der schwachen Kinder auswirkt. In der vorliegenden Studie wurde unter Einsatz dieses Unterrichts-Settings mit den Kindern verhältnismässig lange und behutsam an einem zentralen Thema der Grundschulmathematik gearbeitet. Schwache Kinder konnten von diesem Setting profitieren und ihr Fortschritt zeigt sich nun in verschiedenen Themenfeldern der Mathematik. Gefördert wurde also ein Grundverständnis, das sich positiv auswirkt, auch wenn es sich im Erhebungsinstrument für das Multiplikationsverständnis nicht direkt nachweisen lässt. Das Setting war also für die schwachen Kinder hilfreich, für die stärkeren (denen in diesem Setting eine Art Tutorenrolle zukam) hingegen weniger.

Die für diese Arbeit formulierten Forschungsfragen aus Kapitel 4.2.1 können also beide mit «Ja» beantwortet werden. Durch gezielte Interventionen zur Förderung des Operationsver-

ständnisses kann der Lernzuwachs bezüglich der Rechenleistung in Mathematik beeinflusst werden. Im Hinblick auf die Wirkung der Interventionen gibt es tatsächlich Unterschiede zwischen den Kindern mit schwacher und jenen mit starker Rechenleistung. Der Unterricht mit den vorliegenden Fördereinheiten hat den Lernzuwachs der schwachen Kinder begünstigt. Die starken Kinder haben jedoch beim «gewöhnlichen» Unterricht in den Klassen der Kontrollgruppe grössere Fortschritte gemacht als beim Unterricht mit den Fördereinheiten.

Bemerkenswert ist auch, dass diese Unterschiede messbar wurden, obwohl das Schuljahr 2019/20 aufgrund der Covid-19-Situation ein sehr aussergewöhnliches darstellte.

Vorschlag zur Optimierung des Forschungsdesigns

Für die vorliegende Arbeit etwas bedauerlich ist die Tatsache, dass im zweiten Beobachtungsjahr des Forschungsprojekts bei insgesamt drei Interventionsgruppen lediglich eine Kontrollgruppe zur Verfügung stand (siehe Abbildung 14). Ein Design mit je zwei Interventions- und auch zwei Kontrollgruppen wäre für diese Arbeit vorteilhaft gewesen. Das vorliegende Design stellte eine Konzession an das Gesamtprojekt dar. Es musste beachtet werden, dass alle Lehrpersonen mindestens in einem der beiden Schuljahre eine mathematische Intervention durchführen konnten. Gleichzeitig durfte die Arbeitslast für die einzelnen Lehrpersonen nicht zu hoch werden. Ebenfalls beachtet werden mussten die ursprünglichen Ankündigungen und Versprechen an die Lehrpersonen durch das Projektteam, welche strikt eingehalten wurden. Mit diesen Massnahmen wurde im Projektteam versucht, dafür zu sorgen, dass die Lehrpersonen zufrieden bleiben und nicht aus dem Projekt aussteigen. Für die vorliegende Arbeit ergaben sich durch dieses Design einige Nachteile. Einerseits wurde durch die geringe Grösse der Kontrollgruppe (die sich durch die Covid-19-Situation noch weiter akzentuierte) das Erreichen signifikanter Messergebnisse erschwert. Zudem stand als Kontrollgruppe lediglich diejenige Gruppe zur Verfügung, die im ersten Schuljahr die bestmögliche Förderung erhalten hatte. Dies war etwas bedauerlich, auch wenn die gewählten Auswertungsmethoden diesen Effekt angemessen berücksichtigen. Zusätzlich war der Aufwand für die Betreuung von drei Interventionsgruppen grösser als derjenige für zwei Interventionsgruppen. Für ein zuklünftiges Forschungsprojekt könnte an dieser Stelle also noch eine Optimierung erreicht werden.

7.1.3 Weitere Beobachtungen

Im Laufe der Auswertungen wurden zusätzliche Beobachtungen gemacht, die nicht direkt mit den Forschungsfragen dieser Arbeit in Zusammenhang standen. Über manche wurde in den Kapiteln 5 und 6 berichtet. Einige davon sollen hier in aller Kürze noch einmal erwähnt werden.

Lehrmittel

Ein viel diskutiertes Thema in der Schweiz ist das Lehrmittel, welches im Unterricht eingesetzt wird. Die Beobachtungen in dieser Arbeit deuten darauf hin, dass es weder beim Wissensstand (siehe Kapitel 5.6.3) noch beim Lernfortschritt (siehe Kapitel 6.1.3) zwischen den Klassen Unterschiede gibt, die durch das Lehrmittel erklärt werden könnten.

Varianzanteile der Klassenzuteilung

In Kapitel 5.6.2 wurden die ICC-Werte des Nullmodells bei verschiedenen abhängigen Variablen berichtet. Etwas überraschend ist die Beobachtung, dass die durch die Klassenzuteilung erklärten Varianzanteile bei den allgemeinen Rechenleistungen zwischen T1 und T2 gesunken sind. Zu erwarten wäre, dass der Effekt der gemeinsamen Klassenzugehörigkeit über die Zeit zunimmt, zumal der Testzeitpunkt T1 nur kurz nach Schulbeginn (im 1. Schuljahr) angesetzt war. Die hohen ICC-Werte bei T1 sind also kaum durch die bisherige Beschulung, sondern durch andere Effekte (Einfluss des Kindergartens oder Zusammensetzung der Klassen hinsichtlich städtischem/ländlichem Gebiet bzw. Elternhaus) zu erklären. Über das erste Schuljahr hinweg vermindert sich der durch die Klassenzuteilung erklärte Varianzanteil zugunsten von individuellen Unterschieden. Dies kann so gedeutet werden, dass der Unterricht trotz Unterschiedlichkeiten (verschiedene Kantone, verschiedene Lehrmittel, andere Lehrperson, andere Klassenzusammensetzungen usw.) doch sehr einheitlich ist (dieselben Themen, vergleichbare Lehrpläne usw.). Insgesamt fielen bei verschiedenen Erhebungen und abhängigen Variablen die ICC-Werte in der Grössenordnung von rund 10% auf. Klassenunterschiede erklären also rund 10% der jeweils gemessenen Gesamtvarianz. Dies kann als ein Indiz dafür gesehen werden, dass der Unterricht in diesen Klassen nicht allzu unterschiedlich abläuft.

Sprachfähigkeiten und Mathematikleistung

In verschiedenen bisherigen Studien wurde auf die Bedeutung der Sprache für die Lernprozesse in der Mathematik hingewiesen. Auch in der vorliegenden Arbeit tritt dieser Sachverhalt hervor. Vom Einfluss des Sprachverständnisses sowie der Ausdrucksfähigkeit in der deutschen Sprache wird insbesondere in Kapitel 6.1.3 berichtet. Gute Sprachfähigkeiten gehen mit höherer Leistung in Mathematik einher und begünstigen die Lernprozesse in diesem Schulfach.

7.2 Ausblick

Im Anschluss an diese Arbeit sind verschiedene weiterführende Forschungsfelder denkbar:

Interessant wäre es, das eingesetzte Material (sowohl die Fördereinheiten als auch die Interventionen) unter «regulären Umständen» (ohne Covid-19-Nebeneffekte) noch einmal einzusetzen und die Auswertungen zu wiederholen. Mit «regulären Umständen» gemeint ist der Einsatz in einem Schuljahr, welches nicht von einer zweimonatigen Schulschliessung und allerlei weiteren Unsicherheiten und Ungleichheiten zwischen Kantonen und Schulhäusern geprägt ist. Die Ergebnisse könnten dann anders aussehen als in dieser Arbeit.

Sollte ein anschliessendes Forschungsprojekt in diese Richtung gehen, so wären Anpassungen am Material und am Studiendesign vorzusehen. Die Erhebungsinstrumente für das Operationsverständnis könnten basierend auf der Datenanalyse und den berichteten Erkenntnissen verfeinert werden. Die Items 5a und 5b des Vortests sollten entfernt oder ersetzt werden. Auch weitere Items könnten ersetzt werden, allenfalls um zusätzliche Items zur zeitlich-sukzessiven Grundvorstellung der Multiplikation aufnehmen zu können. Die vier Items mit den Punktefeldern in Vor- und Nachtest sollten angesichts des separaten Faktors, auf welchen sie laden, ebenfalls nochmals untersucht werden. Auch wäre es notwendig, den Deckeneffekt bei den Brückenitems im Auge zu behalten. Alternativ könnte ganz auf Brückenitems verzichtet werden.

Sofern eine Längsschnittanalyse wiederholt würde, könnte es – sofern es das Forschungsprojekt erlaubt – auch interessant sein, die Zeitdauer zwischen den Messungen zu verringern, um die Wirkung der Interventionen noch präziser zu erfassen. Die Anzahl der Klassen könnte gegebenenfalls auch reduziert werden, falls die Aufteilung in Kontroll- und Interventionsgruppe von Beginn weg ausgeglichener gewählt würde.

Ebenfalls wäre es interessant, den einzelnen Faktoren, in welche das Konstrukt Operationsverständnis zu «zerfallen» scheint, weiter nachzugehen. Die vorliegende Arbeit liefert quantitative Indizien dafür, dass Items zu den verschiedenen Grundvorstellungen der Multiplikation durch die Kinder unterschiedlich bearbeitet werden. Auch liefert die Arbeit ein weiteres Indiz dafür, dass die verschiedenen Grundoperationen auf unterschiedliche Faktoren laden. Beides sind bemerkenswerte Erkenntnisse, für die es von Interesse wäre, weitere wissenschaftliche Evidenz zu finden. Für eine solche Untersuchung wäre ein Querschnitt ausreichend, was natürlich eine deutliche Erleichterung des Umfangs der begleitenden Arbeiten für die Forschungsgruppe bedeuten würde.

Diese Arbeit ist des Weiteren eingeschränkt auf Untersuchungen im 2. Schuljahr und auf die Untersuchung des Lernzuwachses beim Multiplikationsverständnis. Eine natürliche Folgearbeit für das vorliegende Projekt ergäbe sich bei eingehender Betrachtung der Division. Diese vierte Grundrechenart war in der Vergangenheit bereits im Fokus mancher Studien (Burtscher, 2019; Schulz & Leuders, 2018). Eine Untersuchung der Wirksamkeit von Fördereinheiten wie in der vorliegenden Arbeit wäre aber auch für die Division eine Neuheit. Entsprechendes Material müsste noch erarbeitet werden.

Schulische Fördermöglichkeiten durch den Einsatz elektronischer Medien wurden in einigen Arbeiten bereits diskutiert (Focke, 2004; Ladel et al., 2018; D. Walter & Schulz, 2019). Dies ist ein Gebiet, in dem sicherlich auch in der kommenden Zeit weiter geforscht werden wird. Auch die Wirksamkeit des Computereinsatzes im Unterricht sollte in Zukunft vermehrt wissenschaftlich untersucht werden. In Bonow et al. (2020, S. 24 f.) beispielsweise werden digitale Medien als besondere Unterstützung bei der Differenzierung und der Ausgestaltung von inklusiven, kooperativen Settings angesehen. Inwiefern das Operationsverständnis bedient werden kann und soll, bleibt dabei offen. Zumindest werden aber die Themenfelder *Zahl* und *Operation* von einzelnen Apps abgedeckt. *Operationsverständnis* scheint ein Themenfeld zu sein, das sich für eine durch den Computer unterstützte Förderung eignen dürfte.

Die längerfristige Wirkung der eingesetzten Fördereinheiten konnte in der vorliegenden Arbeit nicht untersucht werden. Die Beobachtungen endeten mit der letzten Erhebungsrunde am Ende des zweiten Schuljahrs. Dies ermöglicht die Erfassung längerfristiger Effekte für das erste Schuljahr, aber nicht für das zweite. Wollte man solche Effekte auch für das zweite Schuljahr erfassen, so wäre ein allfälliges Folgeprojekt derart anzulegen, dass auch im dritten Schuljahr Messungen erfolgen. Zu den Nachteilen eines solchen Designs gehören natürlich der beachtliche Mehraufwand bei der Betreuung der Lehrpersonen sowie bei der Erfassung der zusätzlichen Zu- und Abgänge von Schulklassen und einzelnen Kindern.

Interessant könnte es auch sein, die Kopplung des kooperativen Settings mit der über einen längeren Zeitraum andauernden Behandlung des Operationsverständnisses aufzuheben. In der vorliegenden Arbeit waren diese beiden Aspekte kombiniert. Die Kombination scheint sich – wenn auch nicht signifikant – positiv auf den Lernzuwachs der schwächeren Kinder ausgewirkt zu haben. Denkbar wäre, dass einer der beiden methodischen Ansätze der gewichtigere ist, d.h. dass die Kinder entweder vom Setting oder aber von der länger andauernden Behandlung des Themas besonders profitierten. In den Daten dieser Arbeit sind die beiden Aspekte nicht zu trennen.

Im Querschnitt weiter untersucht werden könnte auch der Zusammenhang zwischen den verschiedenen Übersetzungsrichtungen beim Konstrukt Operationsverständnis. In dieser Arbeit wurde das Konstrukt aus Gründen der Praktikabilität eingeschränkt betrachtet. Wie an früherer Stelle ausgeführt, sind die bisherigen quantitativen Erkenntnisse zur Frage, ob im Rahmen von Operationsverständnis wirklich auch «verwandte» Fähigkeiten kombiniert betrachtet werden, nicht sehr zahlreich. Dieser Aspekt des Konstrukts Operationsverständnis könnte weiter untersucht werden. Dies selbst auf die Gefahr hin, dass am theoretischen Konstrukt gerüttelt werden müsste, weil sich ergäbe, dass sehr unterschiedliche Fähigkeiten der Kinder kombiniert zu einem möglicherweise unglücklichen Gesamtkonstrukt «verpackt» wurden.

Weitere Forschungsfelder gehen auf, wenn man den Zusammenhang der Leistungen der Kinder mit dem Lernstand der Lehrpersonen betrachtet. Eine kurze Lernstands-Erfassung während den Einführungstreffen mit den freiwillig teilnehmenden Lehrpersonen hat eine hohe Diversität zwischen den Lehrpersonen im Umgang mit dem Begriff Operationsverständnis ans Licht gebracht. Vermutlich stehen die Kenntnisse der Lehrpersonen, ihre Sensibilität für

die Schwierigkeiten der schwächeren Kinder, ihr Umgang mit der Thematik sowie die Art der Unterrichtsgestaltung mit dem Lernzuwachs der Kinder in Zusammenhang. Frühere Arbeiten, wie z.B. Safi (2009) deuten darauf hin, dass das Bewusstsein der Lehrkräfte für die Thematik unterschiedlich, aber dennoch für die Kinder von Relevanz ist. Diese Themenkreise wurden in der vorliegenden Arbeit nicht behandelt und würden sich ebenfalls als weitere Forschungsfelder anbieten.

Als Nebenprodukt von praktischem Wert stehen als Folge der Entwicklungsarbeit nun 16 Fördereinheiten für die Förderung des Operationsverständnisses im kooperativen Setting zur Verfügung (Florin et al., 2019). Das Material liegt im Moment in nicht publizierter Form vor. Sowohl die Rückmeldungen der Lehrpersonen als auch die Ergebnisse der vorliegenden Arbeit deuten darauf hin, dass das Material sich eher für die schwächeren als für die stärkeren Kinder in den regulären Klassen eignet. Gegebenenfalls könnte es durch weiterführende Unterlagen für die stärkeren Kinder angereichert und somit weiter optimiert werden.

8 Anhänge

8.1 Organisatorisches im Projekt MALKA

8.1.1 Strukturdaten

Deutschkenntnisse (Verständnis)

«Das Kind versteht die deutsche Sprache im Unterricht gut. »

Skala: 0: trifft gar nicht zu / 1: trifft eher nicht zu / 2: trifft eher zu / 3: trifft völlig zu

Deutschkenntnisse (Ausdrucksfähigkeit)

«Das Kind kann sich im Schulalltag in der deutschen Sprache gut ausdrücken. »

Skala: 0: trifft gar nicht zu / 1: trifft eher nicht zu / 2: trifft eher zu / 3: trifft völlig zu

Familiensprachen

«Im Familienalltag des Kindes werden eine oder mehrere Sprachen gesprochen. »

Skala: 1: nur deutsch oder schweizerdeutsch / 2: deutsch oder schweizerdeutsch und weitere Sprachen / 3: andere Sprache(n)

Lernzielanpassung

«Für das Kind wurden individuelle Lernziele / eine Lernziel-anpassung vereinbart. »

Skala: 0: keine Lernzielanpassung / 1: Lernzielanpassung

Geschlecht

Geburtsdatum

8.1.2 Testzeitpunkte

T1 Testdurchführung zu Beginn der 1. Klasse (September/Oktober 2018)

- allgemeine mathematische Fähigkeiten (BASIS-MATH 0+)
- MALKA-Test
- Arbeitsgedächtnis
- kognitive Fähigkeiten
- soziale Interaktionen

T2 Testdurchführung am Ende der 1. Klasse (Mai/Juni 2019)

- allgemeine mathematische Fähigkeiten (BASIS-MATH 1+)
- MALKA-Test
- soziale Interaktionen

T3 Testdurchführung zu Beginn der 2. Klasse (September 2019)

- Multiplikationsverständnis (Vortest)

T4 Testdurchführung am Ende der 2. Klasse (Juni/Juli 2020)

- allgemeine mathematische Fähigkeiten (BASIS-MATH 2+)
- Multiplikationsverständnis (Nachtest)
- soziale Interaktionen

Covid-19 bedingte Anpassungen zum Testzeitpunkt T4

Zum Testzeitpunkt T4 wurde aufgrund der Umstände und der Schutzmassnahmen gegen die Ausbreitung der Krankheit Covid-19 auf den ursprünglich geplanten erneuten Einsatz des MALKA-Tests verzichtet. Bei diesem Test hätten die testleitende Person und das Kind auf denselben Laptop-Bildschirm schauen müssen, womit die vom Schutzkonzept vorgeschriebenen Mindestabstände nicht hätten eingehalten werden können.

Auch wurde der Testzeitpunkt T4 um rund einen Monat nach hinten verschoben. Ursprünglich wären Erhebungen in den Monaten Mai und Juni geplant gewesen, die Erhebungen fanden dann aber aufgrund des Zeitpunkts der Schulöffnung erst ab dem 2. Juni 2020 statt.

8.2 Testmaterial

8.2.1 Vortest Operationsverständnis (im Projekt verwendete Version)

Eine Ente ist auf der Wiese. Acht Enten kommen dazu.

Am Baum hängen zwanzig Birnen.
Zwei davon fallen herunter.

Mia hat neun Stifte.
Drei davon gibt sie ihrem Bruder.

Maja und Lea haben zusammen sieben Autos.
Maja hat zwei Autos.

Wie viele Autos hat Lea? _____

Wie viele Stücke Schokolade hat die Tafel? _____

Alle vier Eierschachteln sind voll.

Wie viele Eier hat es? _____

Ben hat vier Bücher.
Manuel hat doppelt so viele Bücher.

Wie viele Bücher haben sie zusammen? _____

An jeden Tisch kommen vier Stühle.

Wie viele Stühle braucht es? _____

Maria steckt in jede Vase zwei Blumen.

Wie viele Blumen braucht sie? _____

Sven geht viermal in den Keller.
Er füllt den Korb jedes Mal mit Flaschen.

Wie viele Flaschen holt er? _____

1 Hand, 5 Finger

3 Hände, _____ Finger

1 Hund, 4 Beine

4 Hunde, _____ Beine

1 Kind, 2 Füsse

12 Kinder, _____ Füsse

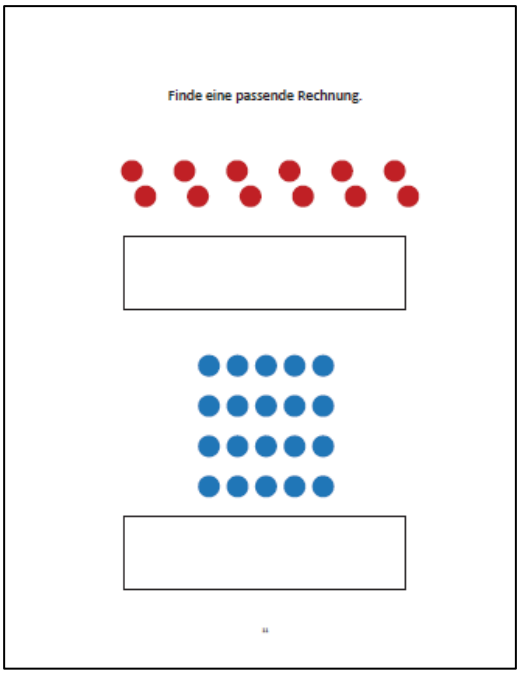
1 Teller, 3 Beeren

2 Teller, 6 Beeren

3 Teller, _____ Beeren

7 Teller, _____ Beeren

Finde eine passende Rechnung.



Quelle der Bilder: www.pixabay.com, www.unsplash.com, (Royar et al., 2014)

8.2.2 Codiermanual des Vortests Operationsverständnis

Codierung:
0 = falsch, ausgelassen oder aus anderen Gründen nicht bearbeitet
1 = richtig (bzw. erwünscht)
4 = wenn Zehner und Einer vertauscht werden (Bsp. 12 statt 21)
5 = wenn die 6 und die 9 verwechselt werden
9 = unleserlich, unverständlich

Korrekturregeln

a) für alle Items:

- Spiegelverkehrt geschriebene Zahlen werden mit 0 codiert (MALKA-Konvention für Zweitklässler).
- Zeichnungen oder Textteile werden mit 0 codiert. Werden Zeichnungen und Textteile mit einer Rechnung ergänzt, so wird nur die Rechnung codiert. Textteile, die eine Zahl bzw. eine Rechnung beschreiben (z.B. «drei» oder «drei plus vier») werden mit 1 codiert, sofern die beschriebene Zahl bzw. Rechnung mit 1 codiert wird.
- Unleserliche oder unverständliche Bearbeitungen werden mit 9 codiert.

b) bei Fragen nach der Rechnung (dem Term):

- Wenn nur das Ergebnis und keine Rechnung notiert ist, wird dies mit 0 codiert.
- Fehlen Operationszeichen im Term, so wird der Term mit 0 codiert.
- Bei der Multiplikation werden die zwei folgenden Schreibweisen mit 1 codiert: 2 · 7 oder 2 x 7. Der Ausdruck «2 7» wird mit 0 codiert.
- Bei der Multiplikation wird immer auch die Tauschaufgabe mit 1 codiert: falls 2 · 7 zur Codierung 1 führt, wird auch 7 · 2 mit 1 codiert.
- Wiederholungen des Operationszeichens werden ignoriert, d.h. «2 ++ 2» wird als «2 + 2» gelesen und entsprechend codiert.
- Operations- oder Gleichheitszeichen am Ende einer Rechnung werden ignoriert. Die Ausdrücke «6+6+6+6+», «6+6+6+6=» und auch «6+6+6+6+=» werden alle als «6+6+6+6» interpretiert und entsprechend codiert.
- Bei Gleichheitszeichen mit Ergebnis zusätzlich zum Rechenterm wird nur der Rechenterm bewertet. Das Ergebnis und das Gleichheitszeichen beeinflussen die Codierung nicht. Es wird also z.B. «5 + 5 = 10» mit 1 codiert, falls der Term «5 + 5» mit 1 codiert wird. Ebenso wird in diesem Falle der Ausdruck «5 + 5 = 12» mit 1 codiert.

c) bei Fragen nach der Rechnung und dem Term zur selben Aufgabe:

- Wird das richtige Ergebnis im Kästchen zur Rechnung dazu geschrieben (z.B. in der Form 4 + 4 = 8) und fehlt die Zahl auf der Linie, so wird auch das Ergebnis, das auf der Linie stehen sollte, mit 1 codiert. Steht hingegen ein falsches Ergebnis auf der Linie, so wird dieses auf jeden Fall mit 0 codiert.

Item 1: Enten

Rechnung: Additionsterm mit Ergebnis 9 unter Verwendung des Summanden 1 (die acht neuen Enten könnten auch einzeln dazu kommen), z.B. «1 + 8», «8+1», «1+1+1+1+1+1+1+1+1» → Code 1

z.B. «3+6» -> Code 0

Item 2: Stifte

Rechnung: Subtraktions- oder Additionsterm mit Gesamtmenge 9 und Teilmengen 6 und 3, «9 – 3», «6 + 3», «3 + 6», «9 – 6» → Code 1

z.B. «9+3», «6+0» - > Code 0

Item 3: Birnen

Rechnung: Subtraktions- oder Additionsterm mit Gesamtmenge 20 und Teilmengen 18 und 2, «20 – 2», «20 – 18», «18 + 2», «2 + 18» → Code 1

z.B. «20+2» -> Code 0

Item 4a & 4b: Autos

Zahl: 5 -> Code 1

Rechnung: Subtraktions- oder Additionsterm mit Gesamtmenge 7 und Teilmengen 2 und/oder 5,
«2 + 5», «5 + 2», «7 – 2» , «7 – 5» → Code 1

Item 5a & 5b: Schokolade

Zahl: 24 → (Code) 1

Rechnung: Additions- oder Multiplikationsterm mit Ergebnis 24 -> Code 1

Item 6a & 6b: Bücher

Zahl: 12 -> Code 1

Rechnung: Additions- oder Multiplikationsterm mit Ergebnis 12, falls Additionsterm dann nur mit Summanden, die Vielfache von 4 sind, also «4+4+4», «4+8», «8+4», «12+0» oder «3 · 4», «4 · 3» → Code 1

z.B. «6+6», «9+3», «10+2» usw. sowie «6 · 2», «2 · 6» -> Code 0

Item 7a & 7b: Eier

Zahl: 24 -> Code 1

Rechnung: Additions- oder Multiplikationsterm mit Ergebnis 24, falls Additionsterm dann mit Summanden, die Vielfache von 6 oder von 3 sind, z.B. «6 + 6 + 6 + 6», «12 + 12», «6 + 18», «24 + 0», «3 + 3 + 6 + 6 + 6», «3 + 3 + 18» oder «6 · 4», «4 · 6», «8 · 3», «3 · 8» → (Code) 1

z.B. «4+4+4+4+4+4», «17 + 7», «20 + 4» -> Code 0

Item 8a & 8b: Stühle

Zahl: 28 -> Code 1

Rechnung: Additions- oder Multiplikationsterm mit Ergebnis 28, falls Additionsterm dann mit Summanden, die Vielfache von 4 sind, z.B. «4 + 4 + 4 + 4 + 4 + 4 + 4», «24+4», «8 + 8 + 8 + 4», «8 + 8 + 4 + 4 + 4», «28 + 0», «12 + 12 + 4» oder «7 · 4», «4 · 7» → Code 1

z.B. «7+7+7+7» oder «14 + 14» -> Code 0

Item 9a & 9b: Blumen

Zahl: 6 -> Code 1

Rechnung: Additions- oder Multiplikationsterm mit Ergebnis 6, falls Additionsterm dann nur in Form von «3 + 3» oder mit Summanden, die Vielfache von 2 sind, also z.B. «6+0», «4+2», «2+2+2» oder Multiplikationsterme «2 · 3», «3 · 2», «1 · 6», «6 · 1» → Code 1

z.B. «5 + 1» und «1 + 5» -> Code 0

Item 10a & 10b: Korb

Zahl: 12 -> Code 1

Rechnung: Additions- oder Multiplikationsterm mit Ergebnis 12, falls Additionsterm dann mit Summanden, die Vielfache von 3 sind, z.B.«3 + 3 + 3 + 3», «9 + 3», «3 + 9», «12 + 0», «6+6», «6 + 3 + 3» oder «4 · 3» → (Code) 1

z.B. «7+5», «8+4» -> Code 0

Item 11: Hand

Zahl: 15 -> Code 1

Item 12: Hund

Zahl: 16 -> Code 1

Item 13: Füsse

Zahl: 24 -> Code 1

Item 14 & 15: Beeren

Zahl bei 3 Tellern: 9 -> Code 1

Zahl bei 7 Tellern: 21 -> Code 1

Item 16: drei 6er Würfel

Rechnung: Additions- oder Multiplikationsterm mit Ergebnis 18, falls Additionsterm, dann mit Summanden, die Vielfache von 6 sind, z.B. «6 + 6 + 6», «12 + 6», «6+12», «18 + 0». Falls Multiplikations-term, dann mit Faktoren, die Vielfache von 6 sind, z.B. «6 · 3», «3 · 6» oder «18 · 1» → (Code) 1

z.B. «9+9», «10+8», «16+2», «2 · 9» usw. -> Code 0

Item 17: sechs 5er Würfel

Rechnung: Additions- oder Multiplikationsterm mit Ergebnis 30, falls Additionsterm, dann mit Summanden, die Vielfache von 5 sind, z.B., «5 + 5 + 5 + 5 + 5 + 5», «15 + 15», «10 + 10 + 10», «20 + 10», «15 + 10 + 5», «25 + 5», «30 + 0». Alle Multiplikationsterme mit Ergebnis 30 (da sie immer einen Faktor haben, der ein Vielfaches von 5 ist), z.B. «6 · 5», «5 · 6», «3 · 10», «2 · 15», «2 · 3 · 5» usw. → (Code) 1

z.B. «28+2», «27+3» usw. -> Code 0

Item 18: 12 Punkte

Rechnung: Additions- oder Multiplikationsterm mit Ergebnis 12 → (Code) 1

Item 19: 20 Punkte

Rechnung: Additions- oder Multiplikationsterm mit Ergebnis 20 → (Code) 1

8.2.3 Nachtest Operationsverständnis (im Projekt verwendete Version)

Maja und Lea haben zusammen sieben Autos.
Maja hat zwei Autos.

Wie viele Autos hat Lea? _____

Nico und seine sechs Freunde pflücken Blumen.
Jedes Kind nimmt fünf Blumen.

Ben hat vier Bücher.
Manuel hat doppelt so viele Bücher.

Wie viele Bücher haben sie zusammen? _____

Sechs Buben und sieben Mädchen gehen wandern.
Jedes Kind hat zwei Flaschen.

Sven geht viermal in den Keller.
Er füllt den Korb jedes Mal mit Flaschen.

Wie viele Flaschen holt er? _____

Auf dem Parkplatz sind acht Autos.
Jedes Auto hat vier Räder.

Mia hat neun Stifte.
Drei davon gibt sie ihrem Bruder.

An jeden Tisch kommen vier Stühle.

Wie viele Stühle braucht es? _____

Hanna legt in jede Schale zwölf Nüsse.

Wie viele Nüsse sind es? _____

Auf jeder Pizza sind vier Scheiben Ananas.

Wie viele Scheiben Ananas hat es? _____

1 Teller, 3 Beeren

2 Teller, 6 Beeren

3 Teller, _____ Beeren

7 Teller, _____ Beeren

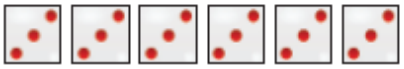
1 Kind, 2 Füsse

12 Kinder, _____ Füsse

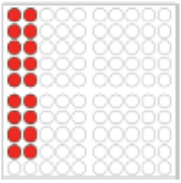
1 Marienkäfer, 6 Beine

5 Marienkäfer, _____ Beine

Finde eine passende **Malrechnung.**

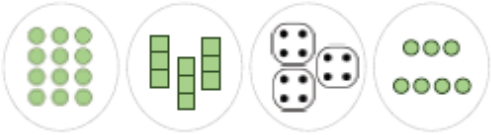


Finde zu den roten Punkten eine passende **Malrechnung.**



Umkreise alle Bilder, in denen du **2 mal 5** siehst.

Umkreise alle Bilder, in denen du **3 mal 4** siehst.



Quelle der Bilder: www.pixabay.com, www.unsplash.com, (Strasser, 2020), (Royar et al., 2014)

8.2.4 Codiermanual des Nachtests Operationsverständnis

Codierung:
0 = falsch oder ausgelassen, nicht bearbeitet, unerwünschte Struktur (siehe unten)
1 = richtig (bzw. erwünscht)
4 = Sondercode für triviale Lösungen (Addition oder Subtraktion von 0, Multiplikation mit 1)
9 = unleserlich, unverständlich

Korrekturregeln

a) für alle Items:

- Spiegelverkehrt geschriebene Zahlen werden mit 0 codiert.
- Verstauschte Ziffern (z.B. «12» statt «21») werden mit 0 codiert.
- Verwechslungen der 6 mit der 9 werden mit 0 codiert.
- Zeichnungen oder Textteile werden mit 0 codiert. Werden Zeichnungen und Textteile mit einer Rechnung ergänzt, so wird nur die Rechnung codiert. Textteile, die eine Zahl bzw. eine Rechnung beschreiben (z.B. «drei» oder «drei plus vier») werden mit 1 codiert, sofern die beschriebene Zahl bzw. Rechnung mit 1 codiert wird.

b) bei Fragen nach der Rechnung (dem Term):

- Wenn nur das Ergebnis und keine Rechnung notiert ist, wird die Rechnung mit 0 codiert.
- Fehlen Operationszeichen im Term, so wird der Term mit 0 codiert.
- Bei der Multiplikation werden die zwei folgenden Schreibweisen mit 1 codiert: 2 · 7 oder 2 x 7. Der Ausdruck «2 7» wird mit 0 codiert.
- Bei der Multiplikation und der Addition wird immer auch die Tauschaufgabe mit 1 codiert: d.h. falls 2 · 7 zur Codierung 1 führt, wird auch 7 · 2 mit 1 codiert. Oder falls 4 + 6 mit 1 codiert wird, führt auch 6 + 4 zur Codierung 1.
- Wiederholungen des Operationszeichens werden ignoriert, d.h. «2 ++ 2» wird als «2 + 2» gelesen und entsprechend codiert.
- Operations- oder Gleichheitszeichen am Ende einer Rechnung werden ignoriert. Die Ausdrücke «6+6+6+6+», «6+6+6+6=» und auch «6+6+6+6+=» werden alle als «6+6+6+6» interpretiert und entsprechend codiert.
- Bei Gleichheitszeichen mit Ergebnis zusätzlich zum Rechenterm wird nur der Rechenterm bewertet. Das Ergebnis und das Gleichheitszeichen beeinflussen die Codierung nicht. Es wird also z.B. «5 + 5 = 10» mit 1 codiert, falls der Term «5 + 5» mit 1 codiert wird. Ebenso wird in diesem Falle der Ausdruck «5 + 5 = 12» mit 1 codiert.
- Wird bei einer Rechnung nicht ein reiner Additions- oder Multiplikationsterm notiert, sondern ein Mischterm, wie z.B. «2 · 5 + 2 · 5» oder «3 · 5 + 5», so gilt: kann der Term so zusammengefasst werden, dass ein Term entsteht, der bei der Aufgabe mit 1 codiert wird, so wird der Mischterm mit 1 codiert. Andernfalls nicht.
- Additionen mit 0, Subtraktionen mit 0 und Multiplikationen mit 1 (falls passend) führen zu Code 4

c) bei Fragen nach der Rechnung und dem Term zur selben Aufgabe:

- Wird das richtige Ergebnis im Kästchen zur Rechnung dazu geschrieben (z.B. in der Form 4+4= 8) und fehlt die Zahl auf der Linie, so wird auch das Ergebnis, das auf der Linie stehen sollte, mit 1 codiert. Steht hingegen ein falsches Ergebnis auf der Linie, so wird dieses auf jeden Fall mit 0 codiert.

Item 1a & 1b: Autos

Zahl: 5 -> Code 1

Rechnung: Subtraktions- oder Additionsterm mit Gesamtmenge 7 und Teilmengen 2 und/oder 5, «2 + 5», «5 + 2», «7 – 2», «7 – 5» → Code 1

z.B. «2 – 7», «7+2», «2+3», «5+0» → Code 0

kein Code 4

Item 2: Blumen

Rechnung: Additionsterm mit Ergebnis 35 und Summanden, die Vielfache von 5 sind, also z.B. «5+30», «5+5+5+20» → Code 1
Multiplikationsterm mit Ergebnis 35 und Faktor 5, also «7 · 5» → Code 1

z.B. «6 + 1», «6 · 5» → Code 0

z.B. «35 · 1», «35+0», «35-0» -> Code 4 (wird bei den folgenden Items nicht immer wiederholt)

Item 3a & 3b: Bücher

Zahl: 12 -> Code 1

Rechnung: Additionsterm mit Ergebnis 12 und Summanden, die Vielfache von 4 sind, also «4+4+4», «4+8» → Code 1
Multiplikationsterm mit Ergebnis 12 und Faktor 4, also «3 · 4» → Code 1

z.B. «6+6», «10+2» sowie «6 · 2» → Code 0

Item 4: Flaschen

Rechnung: Additionsterm mit Ergebnis 26 und Summanden, die Vielfache von 2 sind, also z.B. «2+2+2+2+2+2+2+2+2+2+2+2+2», «12+14», «4+4+4+4+4+4+2», «12+8+6» → Code 1
Multiplikationsterm mit Ergebnis 26: «13 · 2» → Code 1

Mischterme: «6·2 + 7·2» -> Code 1

z.B. «11+15» oder «25+1» -> Code 0

Item 5a & 5b: Korb

Zahl: 12 → Code 1

Rechnung: Additionsterm mit Ergebnis 12 und Summanden, die Vielfache von 3 sind, also «3+3+3+3», «6+6», «6+3+3», «9+3» → Code 1
Multiplikationsterm mit Ergebnis 12 und einem Faktor, der ein Vielfaches von 3 ist, also «4 · 3» oder
«2 · 6» → Code 1

z.B. «2+10», «7+5», «8+4», «4+4+4» → Code 0

Item 6: Parkplatz

Rechnung: Additionsterm mit Ergebnis 32 und Summanden, die Vielfache von 4 sind, z.B. «8+8+8+8», «4+4+4+4+4+4+4+4», «28+4», «20+12», «16+16», → (Code) 1
Multiplikationsterm mit Ergebnis 32, also «8 · 4» oder «16 · 2» → Code 1

z.B. «30+2», «15+15+2», «40 – 8» → Code 0

Item 7: Stifte

Rechnung: Subtraktions- oder Additionsterm mit Gesamtmenge 9 und Teilmengen 6 und/oder 3, also
«9 – 3», «6+3», «9 – 6» → Code 1

z.B. «9+3», «3 – 9», «6 – 3», «6+0» → Code 0

kein Code 4

Item 8a & 8b: Stühle

Zahl: 28 -> Code 1

Rechnung: Additionsterm mit Ergebnis 28, und Summanden, die Vielfache von 4 sind, z.B. «8+8+8+4», «4+4+4+4+4+4+4», «24+4», «4+8+16», «16+12», «20+8» → Code 1
Multiplikationsterm mit Ergebnis 28 und einem Faktor, der ein Vielfaches von 4 ist, also «7 · 4» → Code 1

z.B. «7+7+7+7», «14+14», «10+18», «30-2»→ Code 0

Item 9a & 9b: Nüsse

Zahl: 60 -> Code 1

Rechnung: Additionsterm mit Ergebnis 60 und Summanden, die Vielfache von 12 sind, z.B. «12+12+12+12+12», «48+12», «24+36», «24+24+12» → Code 1
Multiplikationsterm mit Ergebnis 60 und einem Faktor 12, also «5 · 12» → Code 1

z.B. «30+30», «10+10+10+10+10+10» oder «2 · 30» → Code 0

Item 10a & 10b: Pizzas

Zahl: 48 -> Code 1

Rechnung: Additionsterm mit Ergebnis 48 und Summanden, die Vielfache von 4 sind, also z.B. «16+16+16», «12+12+12+12», «4+4+4+4+16+16» → Code 1
Multiplikationsterm mit Ergebnis 48 und einem Faktor, der ein Vielfaches von 4 ist, also «4 · 12», «3 ·16», «6 · 8», «2 · 24» → Code 1

z.B. «30+18», «47+1» → Code 0

Item 11 & 12: Beeren

Zahl bei 3 Tellern: 9 → Code 1 (sonst Code 0, kein Code 4)

Zahl bei 7 Tellern: 21 → Code 1 (sonst Code 0, kein Code 4)

Item 13: Füsse

Zahl: 24 → Code 1 (sonst Code 0, kein Code 4)

Item 14: Marienkäfer

Zahl: 30 → Code 1 (sonst Code 0, kein Code 4)

Item 15: sechs 3er Würfel (Frage nach Malrechnung)

Rechnung: Multiplikationsterm mit Ergebnis 18, also «3 · 6», «2 · 9» → Code 1

z.B. alle Additionsterme (auch «18 + 0», «18-0») → Code 0

«18 · 1» -> Code 4

Item 16: 20 Punkte (Frage nach Malrechnung)

Rechnung: Multiplikationsterm mit Ergebnis 20, also «4 · 5», «2 · 10» → Code 1

z.B. alle Additionsterme (auch «20+0», «20-0») → Code 0

«20 · 1» -> Code 4

Item 17: Hunderterfeld, 18 Punkte (Frage nach Malrechnung)

Rechnung: Multiplikationsterm mit Ergebnis 18, also «3 · 6», «2 · 9» → Code 1

z.B. «10 · 8» und alle Additionsterme, auch «18+0», «18-0» → Code 0

«18 · 1» -> Code 4

Item 18: Hunderterfeld, 49 Punkte (Frage nach Malrechnung)

Rechnung: Multiplikationsterm mit Ergebnis 49, also «7 · 7» → Code 1

z.B. alle Additionsterme (auch «49+0», «49-0») → Code 0

«1 · 49» -> Code 4

Item 19: Umkreise alle Bilder, in denen du 2 mal 5 siehst.

Sofern eindeutig die drei Bilder mit 10 Einheiten markiert sind (z.B. durch Umkreisen) -> Code 1

Sonst (auch wenn einzelne Bilder richtig umkreist sind) -> Code 0

Kein Code 4

Item 20: Umkreise alle Bilder, in denen du 3 mal 4 siehst.

Sofern eindeutig die beiden Bilder mit 12 Einheiten markiert sind (z.B. durch Umkreisen) -> Code 1

Sonst (auch wenn einzelne Bilder richtig umkreist sind) -> Code 0

Kein Code 4

8.3 Interventionsmaterial

8.3.1 Fördereinheit FE06 – Ablauf des Unterrichts

Fördereinheit Herbst MALKA – 2. Klasse FE06

Operationsverständnis Multiplikation: Umgruppieren von Punktebildern zu neuen Malaufgaben

Kompetenzen

- Die Kinder können zu einem vorgegebenen Punktebild (mindestens) eine passende Malaufgabe finden.
- Durch Umgruppieren der Punkte können sie neue Malaufgaben mit demselben Ergebnis finden.

Worum geht es?

- Die Kinder gruppieren ein gegebenes Punktebild durch Umlegen der Plättchen um und finden dadurch neue Malaufgaben mit demselben Ergebnis.
- Handelnd erfahren sie, dass das Ganze (das Produkt) sich nicht ändert, auch wenn die gleich grossen Teile zu neuen gleich grossen Teilen umgruppiert werden. Dies ist eine wichtige Voraussetzung für das flexible Rechnen.

Bezug zum Lehrmittel

- Mathematik 2 Primarstufe: Handbuch, S. 153ff, 159, 161, 259
- Schweizer Zahlenbuch 2: Schulbuch, S. 59ff

Literatur

Gaidoschik (2016), S. 48–50, 65–68

MALKA 1 / 4

Unterrichtsverlauf	Material
Einstiegsphase: Die Lehrperson präsentiert ein Punktebild in Form eines Würfelbildes (mit Magneten an der Wandtafel oder mit Plättchen am Visualizer). ⇨ *Welche Plusaufgabe passt zu diesem Punktebild? Weshalb passt diese Plusaufgabe zum Bild?* ⇨ *Gibt es weitere Plusaufgaben, die zum Bild passen?* ⇨ *Welche Malaufgabe passt zu diesem Punktebild? Kannst du beschreiben und uns zeigen, was du siehst?* Mögliche Kinderantwort: *Ich sehe 3 mal* (zeigt nacheinander auf die drei Vierergruppen) *4 Plättchen* (umkreist die Vierergruppen mit dem Finger). Die Lehrperson schreibt die Malaufgabe auf. ⇨ *Wir legen einzelne Plättchen so um, dass eine neue Malaufgabe entsteht. Wir legen dabei möglichst wenige Plättchen um.* Die Lehrperson beginnt von jeder Viergruppe ein Plättchen (die jeweils rot markierten) zu verschieben: Das Ergebnis der Verschiebung ist das Folgende: 	Tafelbild (zu Beginn): Magnete an der Wandtafel, Plättchen mit Visualizer oder Bilder aus den Arbeitsunterlagen mit Visualizer Tafelbild (am Ende):

⇔ *Welche Malaufgabe passt nun?*

Mögliche Kinderantwort: *Ich sehe 4 mal 3 Plättchen, also 4 mal 3.*

⇔ *Können wir noch andere Malaufgaben bilden?*

Die Lehrperson ergänzt:

⇔ *Versuche immer möglichst wenige Plättchen zu bewegen.*
⇔ *Versuche dir vorzustellen, welches Bild du haben möchtest, bevor du die Plättchen bewegst.*

Die Kinder beschreiben das Umgruppieren einzelner Punkte. Die Lehrperson führt die beschriebene Umgruppierung aus. Die Kinder nennen die entstehenden neuen Malaufgaben:

Mögliche Kinderantwort: *Hier kann man die Malaufgabe 2 · 6 sehen.*

Mögliche Kinderantwort: *Hier kann man die Malaufgabe 6 · 2 sehen.*

Die Malaufgaben 12 · 1 oder 1 · 12 werden nur thematisiert, wenn der Vorschlag von den Kindern kommt.

Die Lehrperson schliesst mit der Beobachtung:

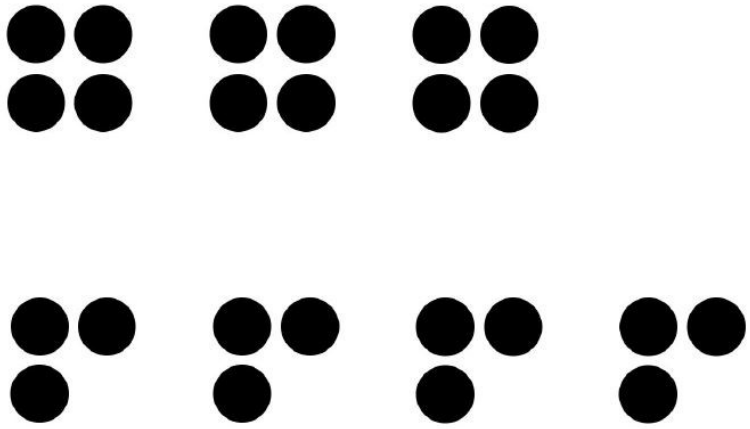
⇔ *Wir haben ganz unterschiedliche Malaufgaben gefunden. Die Anzahl der Plättchen, nämlich 12, bleibt immer gleich.*

Arbeitsphase: Beide Kinder bekommen dieselbe Arbeitsunterlage mit einem Punktebild in Form eines Würfelbildes, das zur Malaufgabe 4 · 5 passt. Kind A legt dieses Punktbild mit Plättchen nach. Kind B nennt die dazu passende Malaufgabe. Kind B bildet durch Umgruppieren der Plättchen eine neue Malaufgabe. Kind A nennt die Malaufgabe. Beide Kinder zeichnen das neue Punktebild auf ihrer Arbeitsunterlage ein. Nun sucht Kind A durch weiteres Umlegen nach einer weiteren Malaufgabe usw. Zusatzaufgabe: Kind A denkt sich eine neue Malaufgabe aus und legt die Plättchen entsprechend. Kind B fährt mit den Plättchen fort wie oben und sucht eine neue Malaufgabe usw. Methodische Form: Wippe	20 Plättchen für jedes einzelne Kind Arbeitsunterlagen zur Übertragung der Bilder mit den Plättchen
Reflexionsphase: Die Lehrperson präsentiert das Punktebild aus der Arbeitsunterlage z.B. mit Magneten an der Wandtafel oder mit Plättchen am Visualizer. Ein Kinderpaar beschreibt seine Umgruppierung. Die Lehrperson nimmt die beschriebene Umgruppierung vor. Andere Kinder ergänzen dies mit weiteren Umgruppierungen, die in der Arbeitsphase entstanden sind. Die Lehrperson nimmt jeweils die Umgruppierung der Magnete von 4 · 5 zu 5 · 4 bzw. zu 10 · 2 usw. vor. Sie ergänzt das Wandtafelbild jeweils auch mit dem Gleichheitszeichen: 4 · 5 = 5 · 4 usw. Zusätzlich werden die Varianten 20 · 1 oder 1 · 20 besprochen. Falls diese von den Kindern nicht selbst angesprochen werden, leitet die Lehrperson das Gespräch mit den folgenden Fragen: ⇨ *Haben wir bereits alle möglichen Malaufgaben gefunden?* ⇨ *Kannst du dir noch andere Malaufgaben mit genau diesen Plättchen vorstellen?* ⇨ *Wir versuchen, alle möglichen Malaufgaben mit diesen Plättchen zu finden.*	Tafelbild (zu Beginn): 4 · 5 Tafelbild (später): je nach Verlauf z.B. 4 · 5 = 5 · 4 oder 4 · 5 = 10 · 2

8.3.2 Fördereinheit FE06 – Arbeitsunterlagen

FE06

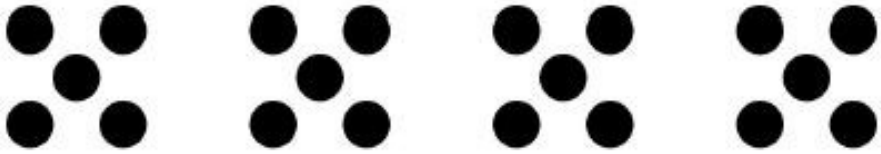
Einstieg FE06*



* Diese Vorlage dient zur Illustration des Tafelbildes. Magnete oder Wendeplättchen eignen sich für die Gestaltung der Einstiegsphase aber besser als diese Vorlage, da sie leicht verschoben werden können.

© MALKA

Name: FE06



© MALKA

8.4 Auswertungen (weitere Daten)

8.4.1 Inter-Item-Korrelationsmatrizen

Inter-Item-Korrelationsmatrix des Vortests (alle Werte)

Item	1	2	3	4a	4b	5a	5b	6a	6b	7a	7b	8a	8b	9a	9b	10a	10b	11	12	13	14	15	16	17	18	19
1	1	.24**	.24**	.13**	.17**	.08**	.12**	.14**	.14**	.11**	.20**	.08**	.09**	.20**	.22**	.10**	.10**	.12**	.09**	.08**	.12**	.09**	.16**	.15**	.17**	.16**
2	.24**	1	.53**	.21**	.30**	.17**	.21**	.20**	.23**	.20**	.24**	.23**	.21**	.25**	.31**	.17**	.25**	.24**	.22**	.22**	.22**	.16**	.20**	.31**	.27**	.23**
3	.24**	.53**	1	.25**	.36**	.14**	.23**	.20**	.20**	.22**	.27**	.24**	.24**	.26**	.33**	.17**	.22**	.25**	.19**	.16**	.21**	.18**	.22**	.24**	.25**	.22**
4a	.13**	.21**	.25**	1	.49**	.16**	.18**	.28**	.25**	.21**	.23**	.20**	.22**	.25**	.21**	.23**	.22**	.29**	.23**	.22**	.20**	.20**	.15**	.22**	.19**	.14**
4b	.17**	.30**	.36**	.49**	1	.16**	.22**	.29**	.30**	.21**	.27**	.22**	.24**	.24**	.26**	.25**	.28**	.29**	.27**	.23**	.21**	.19**	.20**	.27**	.22**	.23**
5a	.08**	.17**	.14**	.16**	.16**	1	.50**	.14**	.08**	.27**	.14**	.16**	.15**	.14**	.17**	.11**	.09**	.12**	.14**	.17**	.15**	.16**	.09**	.14**	.15**	.18**
5b	.12**	.21**	.23**	.18**	.22**	.50**	1	.24**	.19**	.25**	.31**	.21**	.27**	.25**	.32**	.23**	.23**	.27**	.25**	.22**	.22**	.22**	.14**	.24**	.27**	.29**
6a	.14**	.20**	.20**	.28**	.29**	.14**	.24**	1	.76**	.23**	.29**	.26**	.24**	.28**	.24**	.33**	.31**	.26**	.30**	.27**	.22**	.21**	.18**	.24**	.22**	.22**
6b	.14**	.23**	.20**	.25**	.30**	.08**	.19**	.76**	1	.19**	.34**	.21**	.24**	.26**	.25**	.30**	.33**	.24**	.26**	.25**	.19**	.21**	.21**	.23**	.21**	.20**
7a	.11**	.20**	.22**	.21**	.21**	.27**	.25**	.23**	.19**	1	.32**	.26**	.24**	.27**	.25**	.27**	.24**	.21**	.26**	.24**	.17**	.19**	.11**	.23**	.20**	.20**
7b	.20**	.24**	.27**	.23**	.27**	.14**	.31**	.29**	.34**	.32**	1	.23**	.29**	.29**	.35**	.23**	.37**	.30**	.26**	.24**	.21**	.20**	.28**	.29**	.25**	.24**
8a	.08**	.23**	.24**	.20**	.22**	.16**	.21**	.26**	.21**	.26**	.23**	1	.55**	.31**	.28**	.28**	.22**	.27**	.26**	.22**	.24**	.26**	.12**	.25**	.22**	.22**
8b	.09**	.21**	.24**	.22**	.24**	.15**	.27**	.24**	.24**	.24**	.29**	.55**	1	.26**	.34**	.25**	.25**	.21**	.27**	.23**	.21**	.23**	.17**	.24**	.20**	.22**
9a	.20**	.25**	.26**	.25**	.24**	.14**	.25**	.28**	.26**	.27**	.29**	.31**	.26**	1	.55**	.28**	.24**	.33**	.27**	.23**	.26**	.25**	.23**	.28**	.27**	.28**
9b	.22**	.31**	.33**	.21**	.26**	.17**	.32**	.24**	.25**	.25**	.35**	.28**	.34**	.55**	1	.25**	.31**	.32**	.26**	.26**	.26**	.20**	.21**	.32**	.26**	.25**
10a	.10**	.17**	.17**	.23**	.25**	.11**	.23**	.33**	.30**	.27**	.23**	.28**	.25**	.28**	.25**	1	.70**	.25**	.35**	.29**	.20**	.28**	.15**	.21**	.18**	.20**
10b	.10**	.25**	.22**	.22**	.28**	.09**	.23**	.31**	.33**	.24**	.37**	.22**	.25**	.24**	.31**	.70**	1	.25**	.32**	.29**	.21**	.25**	.22**	.23**	.21**	.21**
11	.12**	.24**	.25**	.29**	.29**	.12**	.27**	.26**	.24**	.21**	.30**	.27**	.21**	.33**	.32**	.25**	.25**	1	.41**	.32**	.34**	.28**	.21**	.30**	.27**	.22**
12	.09**	.22**	.19**	.23**	.27**	.14**	.25**	.30**	.26**	.26**	.26**	.26**	.27**	.27**	.26**	.35**	.32**	.41**	1	.33**	.31**	.31**	.23**	.30**	.23**	.26**
13	.08**	.22**	.16**	.22**	.23**	.17**	.22**	.27**	.25**	.24**	.24**	.22**	.23**	.23**	.26**	.29**	.29**	.32**	.33**	1	.32**	.31**	.16**	.28**	.21**	.20**
14	.12**	.22**	.21**	.20**	.21**	.15**	.22**	.22**	.19**	.17**	.21**	.24**	.21**	.26**	.26**	.20**	.21**	.34**	.31**	.32**	1	.40**	.17**	.20**	.22**	.18**
15	.09**	.16**	.18**	.20**	.19**	.16**	.22**	.21**	.21**	.19**	.20**	.26**	.23**	.25**	.20**	.28**	.25**	.28**	.31**	.31**	.40**	1	.13**	.20**	.19**	.21**
16	.16**	.20**	.22**	.15**	.20**	.09**	.14**	.18**	.21**	.11**	.28**	.12**	.17**	.23**	.21**	.15**	.22**	.21**	.23**	.16**	.17**	.13**	1	.42**	.38**	.32**
17	.15**	.31**	.24**	.22**	.27**	.14**	.24**	.24**	.23**	.23**	.29**	.25**	.24**	.28**	.32**	.21**	.23**	.30**	.30**	.28**	.20**	.19**	.42**	1	.53**	.50**
18	.17**	.27**	.25**	.19**	.22**	.15**	.27**	.22**	.21**	.20**	.25**	.22**	.20**	.27**	.26**	.18**	.21**	.27**	.23**	.21**	.22**	.19**	.38**	.53**	1	.54**
19	.16**	.23**	.22**	.14**	.23**	.18**	.29**	.22**	.20**	.20**	.24**	.22**	.22**	.28**	.25**	.20**	.21**	.22**	.26**	.20**	.18**	.21**	.32**	.50**	.54**	1

**. Die Korrelation ist auf dem Niveau von 0,01 (2-seitig) signifikant.

Inter-Item-Korrelationsmatrix des Vortests (nur Werte über .30)

Item	1	2	3	4a	4b	5a	5b	6a	6b	7a	7b	8a	8b	9a	9b	10a	10b	11	12	13	14	15	16	17	18	19
1	1																									
2		1	.53**		.30**										.31**									.31**		
3		.53**	1		.36**										.33**											
4a				1	.49**																					
4b		.30**	.36**	.49**	1				.30**																	
5a						1	.50**																			
5b						.50**	1				.31**				.32**											
6a								1	.76**							.33**	.31**									
6b					.30**			.76**	1		.34**						.33**									
7a										1	.32**															
7b							.31**		.34**	.32**	1				.35**		.37**									
8a												1	.55**	.31**												
8b												.55**	1		.34**											
9a												.31**		1	.55**			.33**								
9b		.31**	.33**				.32**				.35**		.34**	.55**	1		.31**	.32**						.32**		
10a								.33**								1	.70**		.35**							
10b								.31**	.33**		.37**				.31**	.70**	1		.32**							
11														.33**	.32**			1	.41**	.32**	.34**					
12																.35**	.32**	.41**	1	.33**	.31**	.31**		.30**		
13																		.32**	.33**	1	.32**	.31**				
14																		.34**	.31**	.32**	1	.40**				
15																			.31**	.31**	.40**	1				
16																							1	.42**	.38**	.32**
17		.31**													.32**				.30**				.42**	1	.53**	.50**
18																							.38**	.53**	1	.54**
19																							.32**	.50**	.54**	1

**. Die Korrelation ist auf dem Niveau von 0,01 (2-seitig) signifikant.

Inter-Item-Korrelationsmatrix des Vortests (nur Werte über .35)

Item	1	2	3	4a	4b	5a	5b	6a	6b	7a	7b	8a	8b	9a	9b	10a	10b	11	12	13	14	15	16	17	18	19
1	1																									
2		1	.53**																							
3		.53**	1		.36**																					
4a				1	.49**																					
4b			.36**	.49**	1																					
5a						1	.50**																			
5b						.50**	1																			
6a								1	.76**																	
6b								.76**	1																	
7a										1																
7b											1				.35**		.37**									
8a												1	.55**													
8b												.55**	1													
9a														1	.55**											
9b											.35**			.55**	1											
10a																1	.70**									
10b											.37**					.70**	1									
11																		1	.41**							
12																		.41**	1							
13																				1						
14																					1	.40**				
15																					.40**	1				
16																							1	.42**	.38**	
17																							.42**	1	.53**	.50**
18																							.38**	.53**	1	.54**
19																								.50**	.54**	1

**. Die Korrelation ist auf dem Niveau von 0,01 (2-seitig) signifikant.

Inter-Item-Korrelationsmatrix des Nachtests (alle Werte)

Item	1a	1b	2	3a	3b	4	5a	5b	6	7	8a	8b	9a	9b	10a	10b	11	12	13	14	15	16	17	18	19	20
1a	1	.63**	.10**	.21**	.21**	.20**	.22**	.21**	.20**	.22**	.17**	.20**	.13**	.20**	.16**	.20**	.14**	.20**	.25**	.13**	.24**	.24**	.22**	.17**	.16**	.13**
1b	.63**	1	.14**	.21**	.23**	.19**	.24**	.22**	.26**	.29**	.20**	.25**	.23**	.25**	.18**	.22**	.21**	.23**	.26**	.18**	.20**	.22**	.20**	.16**	.17**	.16**
2	.10**	.14**	1	.20**	.18**	.29**	.19**	.26**	.23**	.16**	.13**	.22**	.20**	.23**	.15**	.19**	.13**	.20**	.21**	.14**	.14**	.11**	.14**	.14**	.20**	.10**
3a	.21**	.21**	.20**	1	.86**	.31**	.30**	.25**	.28**	.20**	.13**	.29**	.22**	.25**	.24**	.28**	.19**	.24**	.30**	.24**	.17**	.16**	.21**	.16**	.18**	.14**
3b	.21**	.23**	.18**	.86**	1	.31**	.30**	.24**	.25**	.17**	.12**	.26**	.23**	.24**	.22**	.27**	.17**	.23**	.28**	.24**	.19**	.19**	.25**	.17**	.18**	.16**
4	.20**	.19**	.29**	.31**	.31**	1	.27**	.30**	.36**	.16**	.17**	.31**	.29**	.32**	.28**	.32**	.13**	.23**	.30**	.21**	.22**	.21**	.24**	.20**	.23**	.17**
5a	.22**	.24**	.19**	.30**	.30**	.27**	1	.67**	.30**	.20**	.21**	.31**	.29**	.33**	.20**	.28**	.24**	.29**	.36**	.29**	.20**	.21**	.20**	.22**	.25**	.20**
5b	.21**	.22**	.26**	.25**	.24**	.30**	.67**	1	.46**	.23**	.13**	.35**	.31**	.37**	.17**	.31**	.18**	.24**	.32**	.28**	.26**	.26**	.25**	.24**	.24**	.17**
6	.20**	.26**	.23**	.28**	.25**	.36**	.30**	.46**	1	.30**	.22**	.46**	.38**	.48**	.27**	.38**	.17**	.26**	.36**	.28**	.29**	.28**	.27**	.25**	.27**	.19**
7	.22**	.29**	.16**	.20**	.17**	.16**	.20**	.23**	.30**	1	.18**	.24**	.27**	.27**	.20**	.19**	.19**	.21**	.24**	.18**	.19**	.15**	.17**	.11**	.20**	.13**
8a	.17**	.20**	.13**	.13**	.12**	.17**	.21**	.13**	.22**	.18**	1	.38**	.17**	.18**	.19**	.20**	.18**	.19**	.22**	.14**	.13**	.13**	.16**	.10**	.17**	.15**
8b	.20**	.25**	.22**	.29**	.26**	.31**	.31**	.35**	.46**	.24**	.38**	1	.33**	.51**	.26**	.41**	.15**	.25**	.27**	.26**	.29**	.24**	.23**	.13**	.20**	.15**
9a	.13**	.23**	.20**	.22**	.23**	.29**	.29**	.31**	.38**	.27**	.17**	.33**	1	.47**	.33**	.34**	.17**	.29**	.34**	.31**	.20**	.18**	.19**	.18**	.18**	.18**
9b	.20**	.25**	.23**	.25**	.24**	.32**	.33**	.37**	.48**	.27**	.18**	.51**	.47**	1	.25**	.44**	.18**	.25**	.34**	.27**	.31**	.29**	.25**	.24**	.18**	.17**
10a	.16**	.18**	.15**	.24**	.22**	.28**	.20**	.17**	.27**	.20**	.19**	.26**	.33**	.25**	1	.48**	.18**	.21**	.29**	.21**	.12**	.13**	.18**	.14**	.15**	.17**
10b	.20**	.22**	.19**	.28**	.27**	.32**	.28**	.31**	.38**	.19**	.20**	.41**	.34**	.44**	.48**	1	.19**	.23**	.31**	.25**	.15**	.21**	.26**	.20**	.19**	.17**
11	.14**	.21**	.13**	.19**	.17**	.13**	.24**	.18**	.17**	.19**	.18**	.15**	.17**	.18**	.18**	.19**	1	.37**	.25**	.20**	.13**	.08*	.17**	.15**	.16**	.19**
12	.20**	.23**	.20**	.24**	.23**	.23**	.29**	.24**	.26**	.21**	.19**	.25**	.29**	.25**	.21**	.23**	.37**	1	.35**	.32**	.20**	.17**	.15**	.11**	.22**	.21**
13	.25**	.26**	.21**	.30**	.28**	.30**	.36**	.32**	.36**	.24**	.22**	.27**	.34**	.34**	.29**	.31**	.25**	.35**	1	.38**	.23**	.20**	.19**	.18**	.19**	.19**
14	.13**	.18**	.14**	.24**	.24**	.21**	.29**	.28**	.28**	.18**	.14**	.26**	.31**	.27**	.21**	.25**	.20**	.32**	.38**	1	.16**	.15**	.11**	.12**	.17**	.12**
15	.24**	.20**	.14**	.17**	.19**	.22**	.20**	.26**	.29**	.19**	.13**	.29**	.20**	.31**	.12**	.15**	.13**	.20**	.23**	.16**	1	.51**	.37**	.34**	.23**	.22**
16	.24**	.22**	.11**	.16**	.19**	.21**	.21**	.26**	.28**	.15**	.13**	.24**	.18**	.29**	.13**	.21**	.08*	.17**	.20**	.15**	.51**	1	.42**	.50**	.19**	.15**
17	.22**	.20**	.14**	.21**	.25**	.24**	.20**	.25**	.27**	.17**	.16**	.23**	.19**	.25**	.18**	.26**	.17**	.15**	.19**	.11**	.37**	.42**	1	.57**	.21**	.17**
18	.17**	.16**	.14**	.16**	.17**	.20**	.22**	.24**	.25**	.11**	.10**	.13**	.18**	.24**	.14**	.20**	.15**	.11**	.18**	.12**	.34**	.50**	.57**	1	.17**	.13**
19	.16**	.17**	.20**	.18**	.18**	.23**	.25**	.24**	.27**	.20**	.17**	.20**	.18**	.18**	.15**	.19**	.16**	.22**	.19**	.17**	.23**	.19**	.21**	.17**	1	.43**
20	.13**	.16**	.10**	.14**	.16**	.17**	.20**	.17**	.19**	.13**	.15**	.15**	.18**	.17**	.17**	.17**	.19**	.21**	.19**	.12**	.22**	.15**	.17**	.13**	.43**	1

**. Die Korrelation ist auf dem Niveau von 0,01 (2-seitig) signifikant.

8.4.2 Kommunalitäten der Faktoranalyse

Werte für den Vortest (26 Items und 5 Faktoren)

Item	Anfänglich	Extraktion
1	1.000	.474
2	1.000	.532
3	1.000	.615
4a	1.000	.601
4b	1.000	.658
5a	1.000	.767
5b	1.000	.678
6a	1.000	.850
6b	1.000	.863
7a	1.000	.357
7b	1.000	.396
8a	1.000	.729
8b	1.000	.713
9a	1.000	.530
9b	1.000	.606
10a	1.000	.800
10b	1.000	.834
11	1.000	.472
12	1.000	.460
13	1.000	.426
14	1.000	.588
15	1.000	.496
16	1.000	.483
17	1.000	.649
18	1.000	.654
19	1.000	.620

Anmerkung. Extraktionsmethode: Hauptkomponentenanalyse.

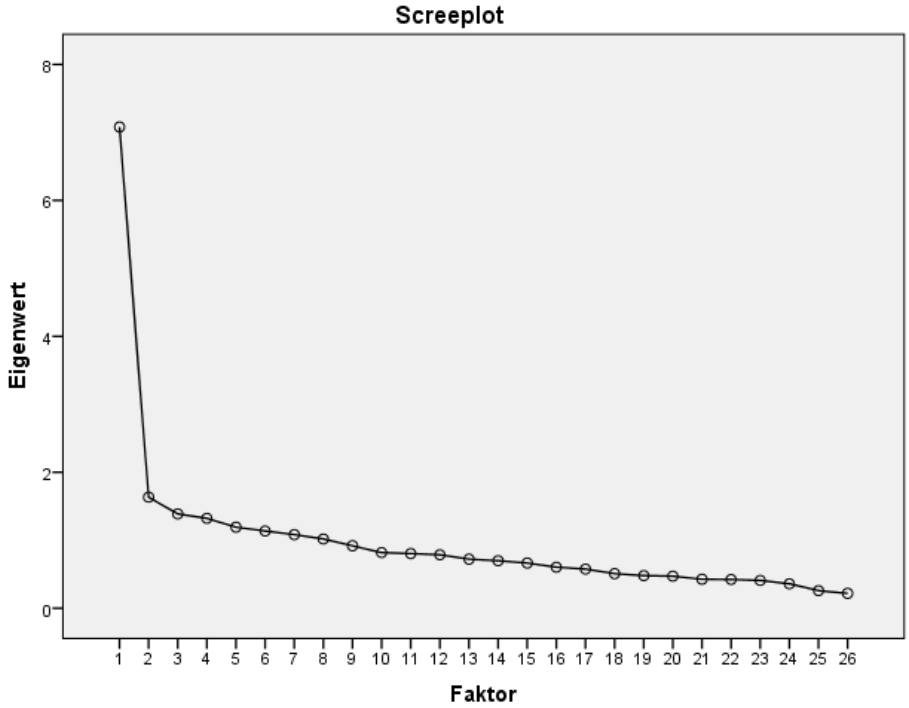
Werte für den Nachtest (26 Items und 5 Faktoren)

Item	Anfänglich	Extraktion
1a	1.000	.762
1b	1.000	.762
2	1.000	.231
3a	1.000	.898
3b	1.000	.902
4	1.000	.379
5a	1.000	.660
5b	1.000	.754
6	1.000	.533
7	1.000	.289
8a	1.000	.308
8b	1.000	.553
9a	1.000	.458
9b	1.000	.578
10a	1.000	.553
10b	1.000	.561
11	1.000	.563
12	1.000	.540
13	1.000	.459
14	1.000	.438
15	1.000	.498
16	1.000	.629
17	1.000	.620
18	1.000	.683
19	1.000	.681
20	1.000	.660

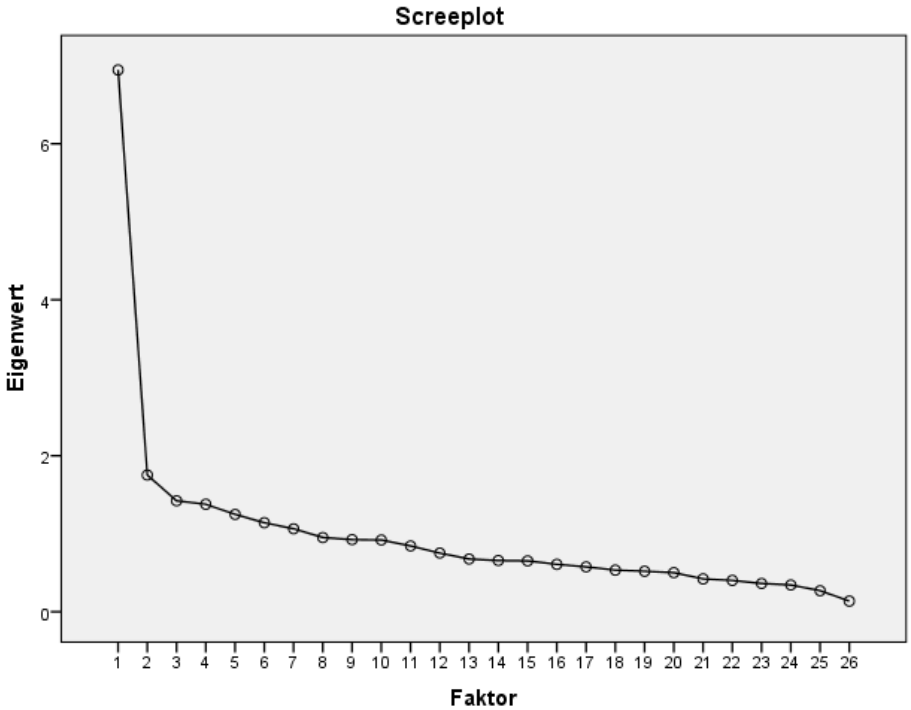
Anmerkung. Extraktionsmethode: Hauptkomponentenanalyse.

8.4.3 Screeplots für die Faktoranalyse

Screeplot der Eigenwerte für den Vortest

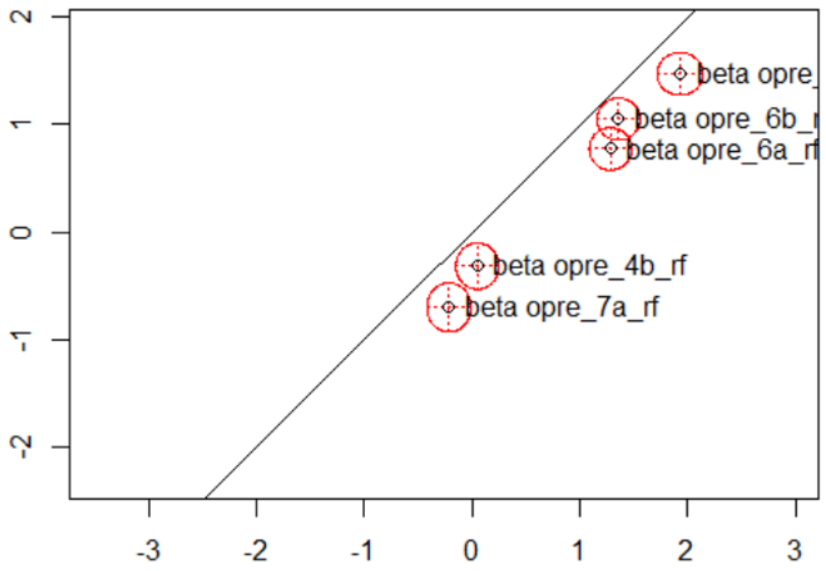


Screeplot der Eigenwerte für den Nachtest

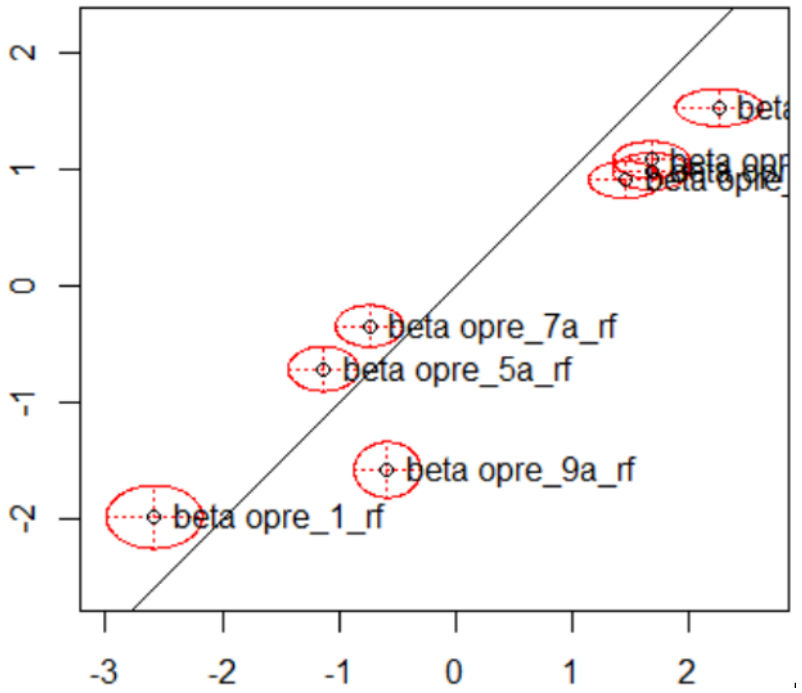


8.4.4 Graphische DIF-Analyse

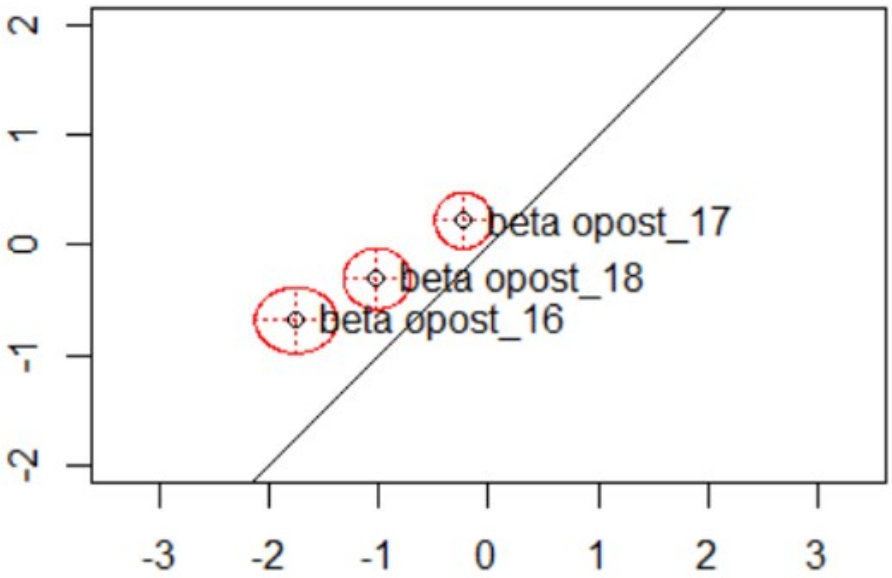
Signifikantes Item-DIF für das Geschlecht mit Konfidenzellipse im Vortest



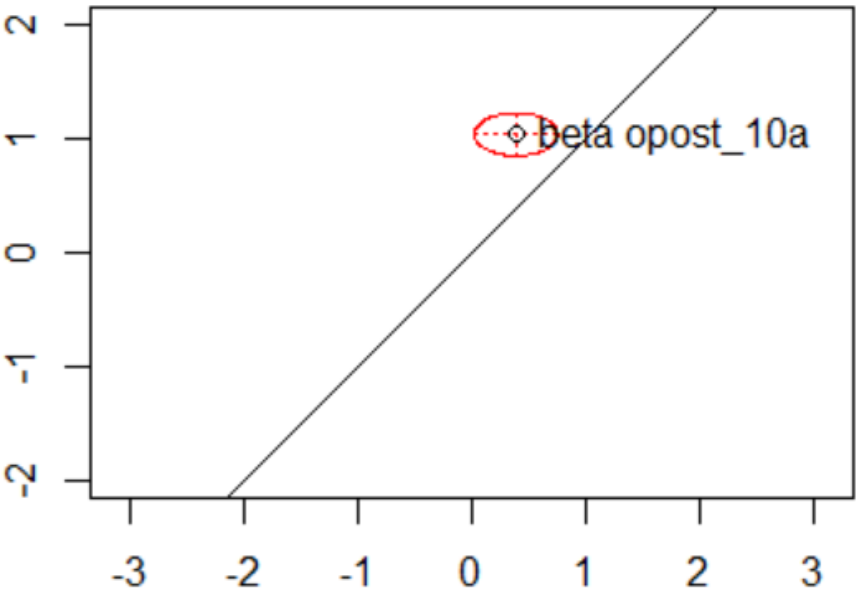
Signifikantes Item-DIF für das Sprachverständnis mit Konfidenzellipse im Vortest



Signifikantes Item-DIF für das Geschlecht mit Konfidenzellipse im Nachtest



Signifikantes Item-DIF für das Sprachverständnis mit Konfidenzellipse im Nachtest



9 Abbildungs-, Tabellen- und Literaturverzeichnis

9.1 Abbildungsverzeichnis

Abbildung 1: Veranschaulichung einer zweistelligen Operation 6

Abbildung 2: Vergrösserung eines 2 · 6 Punktefelds durch Hinzunahme neuer Punkte 17

Abbildung 3: Punktefeld zu den Malaufgaben 3 · 10 sowie 2 · 15 18

Abbildung 4: Klassisches Modell mit drei Darstellungsarten und sechs Übersetzungsrichtungen 22

Abbildung 5: Beispiel eines Modells mit vier Darstellungsarten 23

Abbildung 6: Beispiel eines Modells mit fünf Darstellungsarten und 20 Übersetzungsrichtungen 23

Abbildung 7: Beispiel eines Modells mit verschiedenen Darstellungsebenen 24

Abbildung 8: Modell von Bruner 24

Abbildung 9: Für das Operationsverständnis wesentliche Darstellungswechsel 27

Abbildung 10: Aspekte von Operationsverständnis gemäss PIKASkompakt 38

Abbildung 11: Drei Aspekte von umfassendem Operationsverständnis 39

Abbildung 12: Die zehn Aspekte von «operation sense» 43

Abbildung 13: Operationsverständnis und Schülergruppen (zählend/nicht-zählend) 52

Abbildung 14: Einteilung der Klassen in Gruppen (I, II, III und IV) und Testzeitpunkte T1, T2, T3 und T4 73

Abbildung 15: Für die vorliegende Arbeit zentraler Teil des Forschungsdesigns 82

Abbildung 16: Tafelbild zu Beginn der Einstiegsphase der Fördereinheit 6 (FE06) 105

Abbildung 17: Tafelbild am Ende der Einstiegsphase der Fördereinheit 6 (FE06) 105

Abbildung 18: Skript zur Einstiegsphase der Fördereinheit 6 (FE06) 106

Abbildung 19: Beschreibung der Arbeitsphase der Fördereinheit 6 (FE06) 106

Abbildung 20: Arbeitsunterlage der FE06 für die Kinder (im Format A4) 107

Abbildung 21: Beispiel eines Items aus der Erprobung, welches sprachlich verbessert werden musste. 112

Abbildung 22: Beispiel eines Items aus der Erprobung, welches unterschiedlich interpretiert wurde. 113

Abbildung 23: Beispiel eines Items aus der Erprobung, welches später ergänzt wurde. 114

Abbildung 24: Beispiel eines Items, das überraschende Kinderlösungen hervorbrachte. 122

Abbildung 25: Drei ICCs aus dem Rasch-Modell für verschiedene Item-Schwierigkeiten 130

Abbildung 26: Personenfähigkeit und Testscore; gerechnet mit den Vortest-Daten 132

Abbildung 27: Drei ICCs aus dem Birnbaum-Modell für verschiedene Item-Schwierigkeiten 132

Abbildung 28: Schülerinnen und Schüler (erste Ebene) in einzelnen Schulen (zweite Ebene) 139

Abbildung 29: Verteilung des Fortschritts der 669 Kinder im zweiten Schuljahr (Differenz der Testscores) 149

Abbildung 30: Testscore der 1'084 Kinder (Vortest für das Multiplikationsverständnis) 154

Abbildung 31: Testscore der 1'084 Kinder nach Ausgangssituation (Vortest Multiplikationsverständnis) 155

Abbildung 32: Testscore der 738 Kinder (Nachtest für das Multiplikationsverständnis) 156

Abbildung 33: Anzahl Kinder pro Testscore im Nachtest Multiplikationsverständnis (12 Brückenitems) 157

Abbildungsverweise

Abbildung 2: Ausschnitt aus Fördereinheit 14 (Florin et al., 2019)

Abbildung 3: aus der Reflexionsphase von Fördereinheit 10 (Florin et al., 2019)

Abbildung 4: aus Gerster & Schultz (2004, S. 387)

Abbildung 5: aus Freesemann & Breucker (2014) bzw. aus Kaufmann & Wessolowski (2017)

Abbildung 6: aus Lesh, Post & Behr (1987), Übersetzung aus Freesemann (2014, S. 114)

Abbildung 7: aus Prediger et al. (2011)

Abbildung 8: Modell von Bruner (1964) verwendet in Gleissberg & Eichler (2019)

Abbildung 10: Deutsches Zentrum für Lehrerbildung Mathematik, https://pikas-kompakt.dzlm.de/themenh%C3%A4user/operationsverst%C3%A4ndnis-multiplikation; 7.01.2020

Abbildung 11: https://primakom.dzlm.de/inhalte/zahlen-und-operationen/operationsverst%C3%A4ndnis-aufbauen/hintergrund; 23.01.2020 («Berücksichtung [sic]»)

Abbildung 12: nach Slavit (1999)

Abbildung 13: aus Schäfer (2005, S. 207)

Abbildung 14: aus dem Projekt MALKA, www.malka1und2.ch

Abbildung 15: aus dem Projekt MALKA. www.malka1und2.ch

Abb. 16 bis 20: aus der Fördereinheit 06 aus Florin et al. (2019) nach Gaidoschik (2016a)

Abbildung 22: Ausschnitt aus Testheft, Projekt MALKA; Quelle des Bilds: www.unsplash.com

Abbildung 22: Ausschnitt aus Testheft, Projekt MALKA; Quelle des Bilds: (Scherer, 2016)

Abbildung 23: Ausschnitt aus Testheft, Projekt MALKA; Quelle des Bilds: (Royar et al., 2014)

Abbildung 25: aus Lechner & Niehaus (2010, S. 14)

Abbildung 27: aus Lechner & Niehaus M. (2010, S.61, 69)

Abbildung 28: aus https://www.cmm.bris.ac.uk/lemma; 17.02.2020

Abbildungen in Anhang 8.2.1: www.pixabay.com, www.unsplash.com, (Royar et al., 2014)

Abbildungen in Anhang 8.2.3: www.pixabay.com, www.unsplash.com, (Strasser, 2020), (Royar et al., 2014)

9.2 Tabellenverzeichnis

Tabelle 1: Grösse der Gesamtstichprobe (Anzahl Schulklassen) 74

Tabelle 2: Stichprobe für das 2. Schuljahr (Anzahl Schulklassen) 84

Tabelle 3: Stichprobe für das 2. Schuljahr (Anzahl Kinder) 86

Tabelle 4: Vorgeschlagener Zeitplan zur Behandlung der 16 Fördereinheiten 89

Tabelle 5: Operationsverständnis Addition und Subtraktion & Hinführung zur Multiplikation 101

Tabelle 6: Multiplikationsverständnis 102

Tabelle 7: Liste der Brückenitems 126

Tabelle 8: Item-Schwierigkeit im Vortest 158

Tabelle 9: Item-Schwierigkeit im Nachtest 159

Tabelle 10: Item-Schwierigkeit der Brückenitems (Vor- und Nachtest) 160

Tabelle 11: Trennschärfen («Item-Skala-Korrelationen») im Vortest 161

Tabelle 12: Trennschärfen («Item-Skala-Korrelationen») im Nachtest 161

Tabelle 13: Die grössten zehn Eigenwerte für den Vortest, von denen acht über 1 liegen. 164

Tabelle 14: Faktorladungen (>.3) für den Vortest (mit 5 Faktoren; nach Rotation) 167

Tabelle 15: Die grössten zehn Eigenwerte für den Nachtest, von denen sieben über 1 liegen. 169

Tabelle 16: Faktorladungen (>.3) für den Nachtest (mit 5 Faktoren; nach Rotation) 171

Tabelle 17: Item-Schwierigkeiten β im Rasch-Modell (gerechnet mit eRm) 173

Tabelle 18: Personenfähigkeits-Parameter im Rasch-Modell (gerechnet mit eRm) 174

Tabelle 19: Reliabilitätsmasse im Raschmodell und Werte für den Vortest Multiplikationsverständnis 174

Tabelle 20: Item-Schwierigkeiten β im Rasch-Modell (gerechnet mit TAM) 179

Tabelle 21: Personenfähigkeits-Parameter im Rasch-Modell (gerechnet mit eRm) 180

Tabelle 22: Reliabilitätsmasse im Raschmodell und Werte für den Nachtest Multiplikationsverständnis 180

Tabelle 23: Korrelationen des Vortests (OVpre) mit anderen Erhebungsinstrumenten 183

Tabelle 24: ICC-Werte für verschiedene Erhebungen 185

Tabelle 25: Übersicht Lern-Fortschritt im BASIS-MATH (alle 669 Kinder) 189

Tabelle 26: Ergebnisse der Mehrebenenanalyse (BASIS-MATH, Gesamtbestand) 189

Tabelle 27: Übersicht Lern-Fortschritt im BASIS-MATH (unteres Quartil des Gesamtbestands) 190

Tabelle 28: Ergebnisse der Mehrebenenanalyse (BASIS-MATH, unteres Quartil des Gesamtbestands) 190

Tabelle 29: Übersicht Lern-Fortschritt im BASIS-MATH (unteres Quartil jeder Interventionsgruppe) 191

Tabelle 30: Ergebnisse der Mehrebenenanalyse (BASIS-MATH, unteres Quartil jeder Interventionsgruppe) 191

Tabelle 31: Übersicht Lern-Fortschritt im BASIS-MATH (stärkere 75% des Gesamtbestands) 192

Tabelle 32: Ergebnisse der Mehrebenenanalyse (BASIS-MATH, stärkere 75% des Gesamtbestands) 192

Tabelle 33: Übersicht Lern-Fortschritt im BASIS-MATH (stärkere 75% jeder Interventionsgruppe) 193

Tabelle 34: Ergebnisse der Mehrebenenanalyse (BASIS-MATH, stärkere 75% jeder Interventionsgruppe) 193

Tabelle 35: Übersicht Lern-Fortschritt im BASIS-MATH (schwächere Hälfte, Split bei 21) 194

Tabelle 36: Ergebnisse der Mehrebenenanalyse (BASIS-MATH, schwächere Hälfte, Split bei 21) 195

Tabelle 37: Übersicht Lern-Fortschritt im BASIS-MATH (stärkere Hälfte, Split bei 21) 195

Tabelle 38: Ergebnisse der Mehrebenenanalyse (BASIS-MATH, stärkere Hälfte, Split bei 21) 195

Tabelle 39: Übersicht Lern-Fortschritt im BASIS-MATH (schwächere Hälfte, Split bei 22) 196

Tabelle 40: Ergebnisse der Mehrebenenanalyse (BASIS-MATH, schwächere Hälfte, Split bei 22) 196

Tabelle 41: Übersicht Lern-Fortschritt im BASIS-MATH (stärkere Hälfte, Split bei 22) 197

Tabelle 42: Ergebnisse der Mehrebenenanalyse (BASIS-MATH, stärkere Hälfte, Split bei 22) 197

Tabelle 43: Änderung einzelner Parameter-Werte bei Änderung der abhängigen Variablen 198

Tabelle 44: Modellgüte-Masse „random intercept“-Modell und „random slope“-Modell 199

Tabelle 45: erklärte Varianzanteile („random intercept“ und „random slope“-Modelle) 199

Tabelle 46: Modellgüte-Masse ohne bzw. mit Intervention auf Ebene 1 200

Tabelle 47: Veränderung der erklärten Varianzanteile (Gesamtbestand der 669 Kinder) 200

Tabelle 48: Veränderung der erklärten Varianzanteile (stärkere Hälfte, Split bei 21) 200

Tabelle 49: Parameterschätzungen der β-Werte bei Hinzunahme des Arbeitsgedächtnisses als Prädiktor 201

Tabelle 50: Parameterschätzungen der β-Werte bei Hinzunahme der kognitiven Fähigkeiten als Prädiktor 201

Tabelle 51: Parameterschätzungen der β-Werte bei Hinzunahme des Sprachverständnisses als Prädiktor 202

Tabelle 52: Parameterschätzungen der β-Werte bei Hinzunahme der Ausdrucksfähigkeit als Prädiktor 202

Tabelle 53: Parameterschätzungen der β-Werte bei Hinzunahme der Familiensprache als Prädiktor 202

Tabelle 54: Parameterschätzungen der β-Werte bei Hinzunahme des Geschlechts als Prädiktor 203

Tabelle 55: Parameterschätzungen der β-Werte bei Hinzunahme des Alters als Prädiktor 203

Tabelle 56: Parameterschätzungen der β-Werte bei Hinzunahme des Lehrmittels als Prädiktor 203

Tabelle 57: Veränderung der Modellgüte-Masse AIC und BIC durch Einfügen von Prädiktoren 204

Tabelle 58: Schätzung für β mit «lasso»-Regression (λ gemäss Kreuzvalidierung) 206

Tabelle 59: Schätzung für β mit «lasso»-Regression (λ mit 10 multipliziert) 207

Tabelle 60: Nicht-standardisierte und standardisierte β-Werte für das Modell mit allen 10 Prädiktoren 208

Tabelle 61: Modellgüte-Masse bei Hinzunahme einer dritten Ebene 209

Tabelle 62: Übersicht Lern-Fortschritt im Multiplikationsverständnis (alle 669 Kinder) 210

Tabelle 63: Ergebnisse der Mehrebenenanalyse (Multiplikationsverständnis, Gesamtbestand) 210

Tabelle 64: Übersicht Lern-Fortschritt im Multiplikationsverständnis (unteres Quartil des Gesamtbestands) 211

Tabelle 65: Ergebnisse (Multiplikationsverständnis, unteres Quartil des Gesamtbestands) 211

Tabelle 66: Übersicht Lern-Fortschritt Multiplikationsverständnis (unteres Quartil jeder Int.gruppe) 212

Tabelle 67: Ergebnisse (Multiplikationsverständnis, unteres Quartil jeder Interventionsgruppe) 212

Tabelle 68: Übersicht Lern-Fortschritt im Multiplikationsverständnis (stärkere 75% des Gesamtbestands) 213

Tabelle 69: Ergebnisse (Multiplikationsverständnis, stärkere 75% des Gesamtbestands) 213

Tabelle 70: Übersicht Lern-Fortschritt Multiplikationsverständnis (stärkere 75% jeder Int.gruppe) 214

Tabelle 71: Ergebnisse (Multiplikationsverständnis, stärkere 75% jeder Interventionsgruppe) 214

Tabelle 72: Ergebnisübersicht Multiplikationsverständnis bei Einteilung in zwei „Hälften“ 215

Tabelle 73: Übersicht Lern-Fortschritt bei den Brückenitems im Operationsverständnis (Gesamtbestand) 218

Tabelle 74: Ergebnisse (Multiplikationsverständnis, nur Brückenitems, Gesamtbestand) 219

Tabellenverweise

Tabelle 4: Auszug aus Florin et al. (2019)

Tabelle 5: nach Florin et al. (2019)

Tabelle 6: nach Florin et al. (2019)

9.3 Literaturverzeichnis

Acevedo Nistal, A., Dooren, W., Clarebout, G., Elen, J., & Verschaffel, L. (2009). Conceptualising, investigating and stimulating representational flexibility in mathematical problem solving and learning: A critical review. *ZDM*, *41*(5), 627–636.

Aebli, H. (1994). *Denkprozesse* (2. Auflage, Bd. 2). Klett-Cotta.

Aebli, H. (2011). *Zwölf Grundformen des Lehrens: Eine allgemeine Didaktik auf psychologischer Grundlage; Medien und Inhalte didaktischer Kommunikation, der Lernzyklus* (14. Auflage). Klett-Cotta.

Ainsworth, S., Bibby, P., & Wood, D. (2002). Examining the Effects of Different Multiplie Representational Systems in Learning Primary Mathematics. *The Journal of Learning Science*, *11(1)*, 25–61.

Akinwunmi, K., & Deutscher, T. (2014). 5: 5 = 0 5 Bonbons verteilt an 5 Kinder, da bleibt keins übrig. *PM; Thema: Schwierigkeiten in Mathematik begegnen*, *Heft 56* (56. Jahrgang), 9–15.

Ames, A. J., & Penfield, R. D. (2015). An NCME Instructional Module on Item-Fit Statistics for Item Response Theory Models. *Educational Measurement: Issues and Practice*, *34*(3), 39–48.

Andersen, E. B. (1973). *Conditional inference and models for measuring*. Mentalhygiejnisk forlag.

Anderson, D. (2012). *Hierarchical Linear Modeling (HLM): An Introduction to Key Concepts Within Cross-Sectional and Growth Modeling Frameworks*. Behavioral Research and Teaching - University of Oregon.

Anghileri, J. (1989). An Investigation of Young Children's Understanding of Multiplication. *Educational Studies in Mathematics*, *Vol. 20*(No 4), 367–385.

Anghileri, J., Beishuizen, M., & van Putten, K. (2002). From Informal Strategies to Structured Procedures: Mind the Gap. *Educational Studies in Mathematics*, *49*, 149–170.

Aunio, P., & Räsänen, P. (2016). Core numerical skills for learning mathematics in children aged five to eight years – a working model for educators. *European Early Childhood Education Research Journal*, *24*(5), 684–704.

Ayub, A., Ghazali, M., & Othman, A. R. (2013). Preschool Children's Understanding of Numbers from the Multiple Representation Perspective. *IOSR Journal Of Humanities And Social Science (IOSR-JHSS)*, *14*(6), 93–100.

Baireuther, P. (1999). *Mathematikunterricht in den Klassen 1 und 2*. Auer.

Baker, F. B. (2017). *The basics of item response theory using R*. Springer Science+Business Media.

Ball, D. L., Hill, H. C., & Bass, H. (2005). Knowing Mathematics for Teaching: Who Knows Mathematics Well Enough To Teach Third Grade, and How Can We Decide? *American Educator - the quarterly journal of the American Federation of Teachers, AFL-CIO*.

Barmby, P., Harries, T., Higgins, S., & Suggate, J. (2009). The array representation and primary children's understanding and reasoning in multiplication. *Educational Studies in Mathematics*, *70*(3), 217–241.

Baroody, A. J., Feil, Y., & Johnson, A. R. (2007). An Alternative Reconceptualization of Procedural and Conceptual Knowledge. *Journal for Research in Mathematics*, *Vol. 38*(No. 2), 115–131.

Baroody, A., Lai, M.-L., & Mix, K. (2006). *The Development of Young Children's Early Number and Operation Sense and its Implications for Early Childhood Education.*

Barton, K. (2020). *Package „MuMIn" Multi-Model-Inference*. Comprehensive R Archive Network CRAN. cran.r-project.org/web/packages/MuMIn/MuMIn.pdf

Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R. H. B., Singmann, H., Dai, B., Scheipl, F., Grothendieck, G., Green, P., & Fox, J. (2020). *Package „lme4"*. Comprehensive R Archive Network CRAN. https://cran.r-project.org/web/packages/TAM/TAM.pdf

Baumert, J., Kunter, M., Blum, W., Brunner, M., Voss, T., Jordan, A., Klusmann, U., Krauss, S., Neubrand, M., & Tsai, Y.-M. (2010). Teachers' Mathematical Knowledge, Cognitive Activation in the Classroom and Student Progress. *American Educational Research Journal*, *47*(1), 133–180.

Bender, P. (1991). Ausbildung von Grundvorstellungen und Grundverständnissen—Ein tragendes didaktisches Konzept für den Mathematikunterricht—Erläutert an Beispielen aus dem Sekundarstufenbereich. *Mathematik lehren und lernen*, *Festschrift für Heinz Griesel*, 48–60.

Benz, C. (2018). Den Blick schärfen: Grundlage für arithmetische Kompetenzen. In *Inhalte im Fokus—Mathematische Strategien entwickeln* (Bd. 8, S. 9–24). University of Bamberg Press.

Berch, D. B. (2005). Making Sense of Number Sense: Implications for Children With Mathematical Disabilities. *Journal of Learning Disabilities*, *38*(4), 333–339.

Beutler, B. (2013). Zerlegen und Zusammensetzen. *mahematica didactica*, *36*, 242–271.

Bigalke, H.-G. (1974). Sinn und Bedeutung der Mathematikdidaktik. *Zentralblatt für Didaktik der Mathematik*, *6*, 109–115.

Bönig, D. (1995). *Multiplikation und Division: Empirische Untersuchungen zum Operationsverständnis bei Grundschülern*. Waxmann.

Bonow, J., Schreiber, C., Leinigen, A., Greisbach, M., Steinfeld, L., & Reinert, M. (2020). Digitale Medien im Mathematikunterricht inklusiv gedacht—Eine Kooperation von Ma-

thematikdidaktik und Förderpädagogik. *Mitteilungen der Gesellschaft für Mathematik*, *108*(Februar 2020), 23–28.

Bortz, J., & Döring, N. (2016). *Forschungsmethoden und Evaluation: Für Human- und Sozialwissenschaftler*. Springer.

Bortz, J., & Schuster, C. (2016). *Statistik für Human- und Sozialwissenschaftler: Extras online* (7. Auflage). Springer.

Brandenberg, M., Diener, M., Good, B., Grünigen Mota Campos, S. von, Schärli Wächsler, M., Keller, B., Keller, R., & Walser, M. (2015). *Mathematik: Primarstufe. 1, Handbuch*.

Brandenberg, M., Diener, M., Grünigen Mota Campos, S. von, Höhtker, B., Keller, B., Keller, R., & Müller, B. N. (2013). *Mathematik: Primarstufe. 2, Handbuch*.

Brandt, B. (2019). Mathematik lernen zwischen actio und interactio: Darstellen und Kommunizieren aus interaktionistisch-konstruktiver Perspektive. In *Darstellen und Kommunizieren* (Bd. 9, S. 9–22). University of Bamberg Press.

Bräuning, K., & Schülke, C. (2010). Mathematische Diagnosegespräche mit Kindern. In *Mathematik im Denken der Kinder* (S. 138–168). Klett Kallmeyer.

Breit, S., & Schreiner, C. (Hrsg.). (2016). *Large-Scale Assessment mit R: Methodische Grundlagen der österreichischen Bildungsstandardüberprüfung*. facultas.

Bronštejn, I. N., & Semendjaev, K. A. (Hrsg.). (2006). *Taschenbuch der Mathematik* (6. Auflage, Nachdruck). Deutsch.

Brosius, F. (2013). *SPSS 21: Fundierte Einführung in SPSS und in die Statistik; alle statistischen Verfahren mit praxisnahen Beispielen ; inklusive CD-ROM*. Mitp, Verl.-Gruppe Hüthig, Jehle, Rehm.

Bruner, J. S. (1964). The course of cognitive growth. *American Psychologist*, *19*(1), 1–15.

Büchter, A., Padberg, F., & Springer-Verlag GmbH. (2019). *Einführung in die Arithmetik Primarstufe und Sekundarstufe*.

Bühner, M., & Ziegler, M. (2017). *Statistik für Psychologen und Sozialwissenschaftler* (2., aktualisierte und erweiterte Auflage). Pearson.

Burkhardt, M. (2008). *Operationsverständnis des Addierens und Subtrahierens bei Förderschülern im 3. Schuljahr: Kindliche Übersetzungen zwischen ikonischer und symbolischer Darstellung von Operationen*.

Burtscher, M. (2019). Arbeitsgruppe Arithmetik: Entwicklungen von Grundvorstellungen zur Division. In *Darstellen und Kommunizieren* (Bd. 9, S. 69–72). University of Bamberg Press.

Buschmeier, G., Buttermann, E., Eidt, H., Hacker, J., Lack, C., Lammel, R., Wichmann, M., Link, M., & Buhl, G. (2017). *Mathematik einundzwanzig 2*.

Carpenter, T. P., Ansell, E., Franke, M. L., Fennema, E., & Weisbeck, L. (1993). Models of Problem Solving: A Study of Kindergarten Children's Problem-Solving Processes. *Journal for Research in Mathematics Education*, *24*(5), 428.

Centre for Multilevel Modelling. (2019). *Multilevel Modelling Course. LEMMA VLE.* University of Bristol, Centre for Multilevel Modelling. https://www.cmm.bris.ac.uk/lemma

Confrey, J., & Smith, E. (1995). Splitting, Covariation, and their Role in the Development of Exponential Functions. *Journal for Research in Mathematics Education*, *Vol. 26*(No. 1), 66–86.

Diener, M. (in Vorbereitung). *Die Ablösung vom zählenden Rechnen durch geeignete Förderung im Klassenunterricht anregen. Die Qualität der Erarbeitung von mathematischen Grundlagen im inklusiven Unterricht und deren Zusammenhang zur Ablösung vom zählenden Rechnen. Eine Interventionsstudie zum zählenden Rechnen in der ersten Klasse im Rahmen der MALKA -Studie.* Universität Zürich.

Ditton, H. (1998). *Mehrebenenanalyse: Grundlagen und Anwendungen des hierarchisch linearen Modells*. Juventa-Verl.

Dörfler, W. (1988). Die Genese mathematischer Objekte und Operationen aus Handlungen als kognitive Konstruktion. In *Kognitive Aspekte mathematischer Begriffsentwicklung* (Bd. 16, S. 55–125). Hölder-Pichler-Tempsky.

Dörfler, W. (2015). Abstrakte Objekte in der Mathematik. In G. Kadunz (Hrsg.), *Semiotische Perspektiven auf das Lernen von Mathematik* (S. 33–49). Springer Berlin Heidelberg.

Dörfler, W. (2019). Peirce und Wittgenstein: Ideen für die Grundschule. In *Darstellen und Kommunizieren* (Bd. 9, S. 23–38). University of Bamberg Press.

Dreher, A. (2013). Den Wechsel von Darstellungsebenen fördern und fordern oder vermeiden? In *Mathematik lernen, darstellen, deuten, verstehen* (S. 215–225). Springer Spektrum.

Drollinger-Vetter, B. (2011). *Verstehenselemente und strukturelle Klarheit: Fachdidaktische Qualität der Anleitung von mathematischen Verstehensprozessen im Unterricht*. Waxmann.

Durkin, K., Star, J. R., & Rittle-Johnson, B. (2017). Using comparison of multiple strategies in the mathematics classroom: Lessons learned and next steps. *ZDM*, *49*(4), 585–597.

Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., & Schellinger, K. B. (2011). The Impact of Enhancing Students' Social and Emotional Learning: A Meta-Analysis of School-Based Universal Interventions: Social and Emotional Learning. *Child Development*, *82*(1), 405–432.

Duval, R. (2006). A Cognitive Analysis of Problems of Comprehension in a Learning of Mathematics. *Educational Studies in Mathematics*, *Vol 61, No 1/2*, 103–131.

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.

Eid, M., Gollwitzer, M., & Schmitt, M. (2017). *Statistik und Forschungsmethoden: Mit Online-Materialien* (5., korrigierte Auflage). Beltz.

Eid, M., & Schmidt, K. (2014). *Testtheorie und Testkonstruktion*. Hogrefe.

Enders, C. K., & Tofighi, D. (2007). Centering predictor variables in cross-sectional multilevel models: A new look at an old issue. *Psychological Methods*, *12*(2), 121–138.

Engel, U. (1998). *Einführung in die Mehrebenenanalyse: Grundlagen, Auswertungsverfahren und praktische Beispiele*. Westdt. Verl.

Fetzer, M. (2015). Argumentieren—Prozesse verstehen und Fähigkeiten fördern. In *Entwicklung mathematischer Fähigkeiten von Kindern im Grundschulalter* (Bd. 5, S. 9–24). University of Bamberg Press.

Field, A. P. (2009). *Discovering statistics using IBM SPSS* (3rd edition). Sage.

Field, A. P. (2013). *Discovering statistics using IBM SPSS statistics: And sex and drugs and rock n roll* (4th edition). Sage.

Finch, W. H., Bolin, J. E., & Kelley, K. (2019). *Multilevel modeling using R* (2nd edition). CRC Press.

Fischbein, E., Deri, M., Nello, M. S., & Marino, M. S. (1985). The Role of Implicit Models in Solving Verbal Problems in Multiplication and Division. *Journal for Research in Mathematics Education*, *16*(1), 3.

Flanders, S. T. (2014). *Investigating Flexibility, Reversibility and Multiple Representations in a Calculus Environment: Bd. Graduate Faculty of the School of Education, University of Pittsburg*.

Florin, J., Strasser, C., & Streit, C. (2020). Entwicklung eines Testinstruments zum Multiplikationsverständnis. *Tagungsband der GDM*, 277–280.

Florin, J., Streit, C., & Strasser, C. (2019). *Kooperative Fördereinheiten Operationsverständnis / 2. Klasse*. PH FHNW Muttenz.

Focke, E. (2004). *Computerunterstützte Prävention und Frühförderung bei Rechenschwäche*.

Freesemann, O. (2014). *Schwache Rechnerinnen und Rechner fördern*. Springer Fachmedien Wiesbaden.

Freesemann, O., & Breucker, T. (2014). Förderung flexibler Übersetzungsprozesse. *Grundschulunterricht Mathematik*, *1*, 8–12.

Friedman, J., Hastie, T., Tibshirani, R., Tay, K., Narasimhan, B., Noah, S., & Qian, J. (2020). *Package 'glmnet'—Lasso and Elastic-Net Regularized Generalized Linear Models*. Comprehensive R Archive Network CRAN.

Fritz, A., Ehlert, A., & Leutner, D. (2018). Arithmetische Konzepte aus kognitiv-entwicklungspsychologischer Sicht. *Journal für Mathematik-Didaktik*, *39*(1), 7–41.

Gaidoschik, M. (2010). *Wie Kinder rechnen lernen--oder auch nicht: Eine empirische Studie zur Entwicklung von Rechenstrategien im ersten Schuljahr*. Peter Lang.

Gaidoschik, M. (2015a). *Rechenschwäche-Dyskalkulie: Eine unterrichtspraktische Einführung für LehrerInnen und Eltern* (9. Auflage). Persen.

Gaidoschik, M. (2015b). Vermeidbare und unvermeidbare Hürden beim Rechnenlernen. In *Entwicklung mathematischer Fähigkeiten von Kindern im Grundschulalter* (Bd. 5, S. 25–38). University of Bamberg Press.

Gaidoschik, M. (2016a). *Einmaleins verstehen, vernetzen, merken: Strategien gegen Lernschwierigkeiten* (3. Auflage). Klett/Kallmeyer.

Gaidoschik, M. (2016b). *Rechenschwäche verstehen - Kinder gezielt fördern: Ein Leitfaden für die Unterrichtspraxis* (9. Auflage). Persen.

Garrote, A., Gloor, N., Nesme, C., & Moser Opitz, E. (in Vorbereitung). Längsschnittanalyse der sozialen Partizipation im Forschungsprojekt MALKA. *Institut für Erziehungswissenschaft. Universität Zürich. Schweiz*.

Geary, D. C. (1994). *Children's mathematical development: Research and practical applications*. American Psychological Association.

Gerlach, M. (2007). *Entwicklungsaspekte des Rechnenlernens: Bd. Fachbereich Bildungswissenschaften*.

Gerster, H.-D. (1994). Arithmetik im Anfangsunterricht. In *Handbuch zur Grundschulmathematik* (S. 35–102). Klett.

Gerster, H.-D., & Schultz, R. (2004). *Schwierigkeiten beim Erwerb mathematischer Konzepte im Anfangsunterricht*. Institut für Mathematik und Informatik und ihre Didaktiken.

Gleissberg, S. (2018). „Man muss es ja doch auswendig wissen"—Die Behandlung der Multiplikation zwischen Anspruch und Wirklichkeit. *Beiträge zum Mathematikunterricht*. GDM-Tagung, Münster.

Gleissberg, S., & Eichler, K.-P. (2019). *The offer of tasks to work on multiplication in grades 2 and 3*. Eleventh Congress of the European Society for Research in Mathematics Education, Utrecht, Netherlands.

Gloor, N. (2018). *Überprüfung der spontanen Fokussierung auf Anzahlen (SFON) und der Anzahlerfassung in der Schuleingangsstufe: Entwicklung eines Instruments*. Universität Zürich.

Götze, D. (2007). *Mathematische Gespräche unter Kindern: Zum Einfluss sozialer Interaktion von Grundschulkindern beim Lösen komplexer Aufgaben*. Franzbecker.

Götze, D. (2015). *Sprachförderung im Mathematikunterricht*. Cornelsen.

Götze, D. (2018). *Multiplikationsverständnis sprachsensibel fördern—Ergebnisse einer Interventionsstudie im inklusiven Setting*. *Beiträge zum Mathematikunterricht 2018*(Münster, WTM-Verlag), 639–642.

Greefrath, G., Oldenburg, R., Siller, H.-S., Ulm, V., & Weigand, H.-G. (2016). Aspects and "Grundvorstellungen" of the Concepts of Derivative and Integral: Subject Matter-related Didactical Perspectives of Concept Formation. *Journal Für Mathematik-Didaktik*, *37*(S1), 99–129.

Griesel, H., vom Hofe, R., & Blum, W. (2019). Das Konzept der Grundvorstellungen im Rahmen der mathematischen und kognitionspsychologischen Begrifflichkeit in der Mathematikdidaktik. *Journal für Mathematik-Didaktik*, *40*(1), 123–133.

Grob, U. (2018). *Einführung in die probabilistische Testtheorie—Skript „Doktoratsstudiengang Erziehungswissenschaften—HS18"*. Universität Zürich.

Gysin, B. (2013). Kinder erkennen Strukturen. In *Mathematik lernen, darstellen, deuten, verstehen* (S. 125–138). Springer Spektrum.

Harries, T., & Barmby, P. (2007). Representing and Understanding Multiplication. *Research in Mathematics Education*, *9*(1), 33–45.

Hartig, J., & Bechtoldt, M. (2020). *Hierarchische Lineare Modelle*. Uni Frankfurt. https://user.uni-frankfurt.de/~johartig/hlm/HLM_Muenster.pdf

Hartig, J., & Frey, A. (2013). Sind Modelle der Item-Response-Theorie (IRT) das „Mittel der Wahl" für die Modellierung von Kompetenzen? *Zeitschrift für Erziehungswissenschaft*, *16*(S1), 47–51.

Hartig, J., & Kühnbach, O. (2006). Schätzung von Veränderung mit „plausible values" in mehrdimensionalen Rasch-Modellen. In *Veränderungsmessung und Längsschnittstudien in der empirischen Erziehungswissenschaft* (S. 27–44). VS Verlag für Sozialwissenschaften.

Häsel-Weide, U. (2016). Mathematik gemeinsam lernen—Lernumgebungen für den inklusiven Mathematikunterricht. In *Inklusiver Mathematikunterricht—Mathematiklernen in ausgewählten Förderschwerpunkten* (Bd. 6, S. 9–24). University of Bamberg Press.

Häsel-Weide, U., & Nührenbörger, M. (2016). *Vom Zählen zum Rechnen: Strukturfokussierende Deutungen in kooperativen Lernumgebungen*. Springer Spektrum.

Häsel-Weide, U., & Nührenbörger, M. (2018). 4. Aufgabenformate für einen inklusiven Arithmetikunterricht. In *Inklusiver Mathematikunterricht in der Grundschule* (4. Auflage, S. 58–74). Mildenberger.

Häsel-Weide, U., Nührenbörger, M., Moser Opitz, E., & Wittich, C. (2017). *Ablösung vom zählenden Rechnen: Fördereinheiten für heterogene Lerngruppen* (4. Auflage). Klett, Kallmeyer.

Hasemann, K., & Stern, E. (2002). Die Förderung des mathematischen Verständnisses anhand von Textaufgaben—Ergebnisse einer Interventionsstudie in Klassen des 2. Schuljahres. *Journal für Mathematik-Didaktik*, *23*(3–4), 222–242.

Hasselhorn, M., Schumann-Hengsteler, R., Gronauer, J., Grube, D., Mähler, C., Schmid, I., Seitz-Stein, K., & Zoelch, C. (2012). *AGTB 5-12 Arbeitsgedächtnistestbatterie für Kinder von 5 bis 12 Jahren*. Hogrefe.

Hayton, J. C., Allen, D. G., & Scarpello, V. (2004). Factor Retention Decisions in Exploratory Factor Analysis: A Tutorial on Parallel Analysis. *Organizational Research Methods*, *7*(2), 191–205.

Hess, K. (2018). *MATHWELT 1*. Schulverlag plus AG. www.mathwelt.ch

Hiebert, J., & Carpenter, T. P. (1992). Learning and teaching with understanding. In *Handbook of research on mathematics teaching and learning.* (S. 65–97). Macmillan.

Hino, K., & Kato, H. (2019). Teaching whole-number multiplication to promote children's proportional reasoning: A practice-based perspective from Japan. *ZDM*, *51*(1), 125–137.

Hofmann-Villiger, A. (in Vorbereitung). *Mikro- und makroadaptives Handeln von Schulischen Heilpädagoginnen und Heilpädagogen in Kleingruppen. Mathematische Förderung von Schülerinnen und Schülern der 1.Klasse zur Ablösung vom zählenden Rechnen (Unveröffentlichtes Manuskript)*. Pädagogische Hochschule St. Gallen.

Hox, J. J., Moerbeek, M., & Schoot, R. van de. (2018). *Multilevel Analysis: Techniques and Applications* (Third edition). Routledge.

Huinker, D. M. (1993). *Interviews: A Window to Students' Conceptual Knowledge of Operations*.

Hurst, C. (2015). The multiplicative situation. *Australian Primary Mathematics Classroom*.

Hurst, C., & Hurrel, D. (2016). Multiplicative thinking: Much more than knowing multiplication facts and procedures. *Australian Primary Mathematics Classroom*.

Hurst, C., & Hurrel, D. (2017). Where we were...where we are heading: One mutiplicative journey. *Australian Primary Mathematics Classroom*.

Ise, E., Dolle, K., Pixner, S., & Schulte-Körne, G. (2012). Effektive Förderung rechenschwacher Kinder: Eine Metaanalyse. *Kindheit und Entwicklung*, *21*(3), 181–192.

Jacobs, C., & Petermann, F. (2003). Übersichten / Dyskalkulie—Forschungsstand und Perspektiven. *Kindheit und Entwicklung*, *12*(4), 197–211.

Jordan, A., Krauss, S., Löwen, K., Blum, W., Neubrand, M., Brunner, M., Kunter, M., & Baumert, J. (2008). Aufgaben im COACTIV-Projekt: Zeugnisse des kognitiven Aktivierungspotentials im deutschen Mathematikunterricht. *JMD*, *29 H. 2*, 83–107.

Kamata, A., & Bauer, D. J. (2008). A Note on the Relation Between Factor Analytic and Item Response Theory Models. *Structural Equation Modeling: A Multidisciplinary Journal*, *15*(1), 136–153.

Käpnick, F. (2014). *Mathematiklernen in der Grundschule*. Springer Spektrum.

Kaufmann, S., & Wessolowski, S. (2017). *Rechenstörungen: Diagnose und Förderbausteine* (6., erweiterte, überarbeitete Auflage). Klett Kallmeyer.

Kaur, B. (2019). The why, what and how of the 'Model' method: A tool for representing and visualising relationships when solving whole number arithmetic word problems. *ZDM*, *51*(1), 151–168.

Kish, L. (1965). *Survey Sampling*. Wiley.

Koller, I., Alexandrowicz, R., & Hatzinger, R. (2012). *Das Rasch-Modell in der Praxis: Eine Einführung mit eRm*. facultas.wuv.

Korff, N. (2016). *Inklusiver Mathematikunterricht in der Primarstufe: Erfahrungen, Perspektiven und Herausforderungen* (2. Auflage). Schneider Verlag Hohengehren GmbH.

Krähenmann, H., Labhart, D., Schnepel, S., Stöckli, M., & Moser Opitz, E. (2018). 3. Gemeinsam lernen—Individuell fördern: Differenzierung im inklusiven Mathematikunterricht. In *Inklusiver Mathematikunterricht in der Grundschule* (4. Auflage, S. 43–57). Mildenberger.

Krauss, S., Neubrand, M., Blum, W., Baumert, J., Brunner, M., Kunter, M., & Jordan, A. (2008). Die Untersuchung des professionellen Wissens deutsscher Mathematik-Lehrerinnen und—Lehrer im Rahmen der COACTIV-Studie. *Journal für Mathematik-Didaktik*, *JMD 29*(3/4), 223–258.

Krauthausen, G., & Scherer, P. (2014). *Einführung in die Mathematikdidaktik* (3. Aufl., Nachdr). Springer Spektrum.

Krauthausen, G., & Scherer, P. (2016). *Natürliche Differenzierung im Mathematikunterricht: Konzepte und Praxisbeispiele aus der Grundschule* (2. Auflage). Klett/Kallmeyer.

Kreis, A., Wick, J., Kosorok Labhart, C., & Staub, F. C. (2012). Professionsrelavantes Wissen in Unterrichtsbesprechungen—Eine Interventionsstudie zu Fachspezifischem Unterrichtscoaching. In *Lehrerprofessionalisierung wissenschaftlich begleiten—Strategien und Methoden* (S. 35–50). Waxmann.

Kretschmann, R. (2006). „Pädagnostik"—Optimierung pädagogischer Angebote durch differenzierte Lernstandsdiagnosen, unter besonderer Berücksichtigung mathematischer Kompetenzen. In *Die Entwicklung mathematischen Denkens in Kindergarten und Grundschule: Beobachten—Fördern—Dokumentieren* (S. 29–54). Mildenberger.

Kubinger, K. D., Rasch, D., & Yanagida, T. (2011). *Statistik in der Psychologie: Vom Einführungskurs bis zur Dissertation*. Hogrefe.

Kuhnke, K. (2012). *Unterrichtsanregungen zur Förderung des Darstellungswechsels*. PIK AS Uni Dortmund.

Kuhnke, K. (2013). *Vorgehensweisen von Grundschulkindern beim Darstellungswechsel: Eine Untersuchung am Beispiel der Multiplikation im 2. Schuljahr*. Springer Spektrum.

Kunter, M., Baumert, J., Blum, W., Klusmann, U., Krauss, S., & Neubrand, M. (Hrsg.). (2013). *Cognitive Activation in the Mathematics Classroom and Professional Competence of Teachers: Results from the COACTIV Project*. Springer US.

Kuntze, S. (2013). Sichtweisen von Lehramtsstudierenden zur Bedeutung des Nutzens vielfältiger Darstellungen im Mathematikunterricht. In *Mathematik lernen, darstellen, deuten, verstehen* (S. 263–272). Springer Spektrum.

Laakmann, H. (2013). *Darstellungen und Darstellungswechsel als Mittel zur Begriffsbildung: Eine Untersuchung in rechnerunterstützten Lernumgebungen*. Springer Spektrum.

Ladel, S. (2011a). Multiple externe Repräsentationen (MERs)—Gestaltungsprinzipien und deren Umsetzung bei Software für den Anfangsunterricht. *Forum Forschung das Wissenschaftsmagazin der Pädagogischen Hochschule Schwäbisch Gmünd*, *2*, 38–41.

Ladel, S. (2011b). Multiplex-R: Zum Wechsel zwischen verschiedenen Darstellungsformen von Zahlen und Operationen bei 5- bis 8-jährigen Kindern. *Beiträge zum Mathematikunterricht*, *BzMU*, 527–530.

Ladel, S. (2018). Sinnvolle Kombination virtueller und physischer Materialien. In *Digitalisierung und Bildung* (S. 3–22). Springer VS.

Ladel, S., Knopf, J., & Weinberger, A. (Hrsg.). (2018). *Digitalisierung und Bildung*. Springer Fachmedien Wiesbaden.

Lamprecht, X. (2017). „Sich Multiplikation vorstellen"—Individuelle Grundvorstellungen von Kindern mit und ohne Förderbedarf. *Beiträge zum Mathematikunterricht*, 593–596.

Lamprecht, X. (2020). *Multiplikatives Verständnis fördern Entwicklung und Evaluation eines Förderkonzepts in differenten Rahmenbedingungen Volume 35. Volume 35.*

Langer, W. (2009). *Mehrebenenanalyse: Eine Einführung für Forschung und Praxis* (2. Auflage). VS, Verlag für Sozialwissenschaften.

Langer-Osuna, J. M. (2018). Exploring the central role of student authority relations in collaborative mathematics. *ZDM*, *50*(6), 1077–1087.

Langhorst, P., Ehlert, A., & Fritz, A. (2011). Das Teil-Teil-Ganze-Konzept. *MNU PRIMAR 3*, *1*, 10–17.

Lavie, Steiner, A., & Sfard, A. (2018). Routines we live by: From ritual to exploration. *Educational Studies in Mathematics*, 1–24.

Lechner, C., & Niehaus, M. (2010). *Skript zur Vorlesung „Theorien psychometrischer Tests II" WS2009/2010*. Uni Jena.

Leisen, J. (2005). Wechsel der Darstellungsformen, Ein Unterrichtsprinzip für alle Fächer. *Der fremdsprachliche Unterricht Englisch*, *78*, 9–11.

Leiss, D., & Plath, J. (2020). „Im Mathematikunterricht muss man auch mit Sprache rechnen!“ – Sprachbezogene Fachleistung und Unterrichtswahrnehmung im Rahmen mathematischer Sprachförderung. *Journal für Mathematik-Didaktik*, *41*(1), 191–236.

Lesh, R., Post, T., & Behr, M. (1987). Representations and Translations among Representations in Mathematics and Problem Solving. *C. Janvier, (Ed), Problems of Representations in the Teaching and Learning of Mathematics*, *Hillsdale, NJ: Lawrence Erlbaum*, 33–40.

Leuenberger, D. (in Vorbereitung). *Erfassung und Lernzuwachs von nicht-zählenden Kopfrechenkompetenzen im ersten Schuljahr: Eine Interventionsstudie zur Ablösung vom zählenden Rechnen im inklusiven Unterricht. Dissertation.* Universität Zürich.

Long, J. D. (2012). *Longitudinal data analysis for the behavioral sciences using R*. SAGE.

Lorenz, J. H. (2008). *Lernschwache Rechner fördern: Ursachen der Rechenschwäche, Frühhinweise auf Rechenschwäche, diagnostisches Vorgehen* (5. Aufl). Cornelsen Scriptor.

Lorenz, J. H. (2013). Zahlen und Rechenoperationen. In *Mathematik lernen, darstellen, deuten, verstehen* (S. 3–15). Springer Spektrum.

Lorenz, J. H. (2017). Einige Anmerkungen zur Repräsentation von Wissen über Zahlen. *Journal für Mathematik-Didaktik*, *38*(1), 125–139.

Lorenz, J. H., & Radatz, H. (2008). *Handbuch des Förderns im Mathematikunterricht* (Dr. A 8). Schroedel.

Lüdtke, O., Robitzsch, A., Trautwein, U., & Kunter, M. (2009). Assessing the impact of learning environments: How to use student ratings of classroom or school characteristics in multilevel modeling. *Contemporary Educational Psychology*, *34*(2), 120–131.

Luhmann, M. (2015). *R für Einsteiger: Einführung in die Statistiksoftware für die Sozialwissenschaften: mit Online-Material* (4. Auflage). Beltz.

Magis, D., Beland, S., & Raiche, G. (2018). *Package „difR“: Collection of Methods to Detect Dichotomous Differential Item Functioning (DIF)*. https://cran.r-project.org/web/packages/difR/difR.pdf.

Mair, P., Hatzinger, R., Maier, M. J., Rusch, T., & Debelak, R. (2020). *Package „eRm“ Extended Rasch Modelling*. Comprehensive R Archive Network CRAN.

Malle, G. (2003). Vorstellungen vom Differenzenquotienten fördern. *mathematik lehren*, *Heft 118*, 57–62.

Malle, G., & Malle, S. (2003). Was soll man sich unter einer Wahrscheinlichkeit vorstellen? *mathematik lehren*, *Heft 118*, 52–56.

Mantel, N., & Haenszel, W. (1959). Statistical Aspects of the Analysis of Data From Retrospective Studies of Disease. *JNCI: Journal of the National Cancer Institute*.

Moosbrugger, H., & Kelava, A. (Hrsg.). (2012). *Testtheorie und Fragebogenkonstruktion: Mit 66 Abbildungen und 41 Tabellen* (2., aktualisierte und überarbeitete Auflage). Springer.

Morgan, C., Craig, T., Schuette, M., & Wagner, D. (2014). Language and communication in mathematics education: An overview of research in the field. *ZDM*, *46*(6), 843–853.

Moser Opitz, E. (2005). Lernschwierigkeiten Mathematik in Klasse 5 und 8. *VHN*, *74 Jg.*, 113–128.

Moser Opitz, E. (2009). Erwerb grundlegender Konzepte der Grundschulmathematik als Voraussetzung für das Mathematiklernen in der Sekundarstufe I. In *Fördernder Mathematikunterricht in der Sek. I* (S. 29–43). Beltz.

Moser Opitz, E. (2013). *Rechenschwäche/Dyskalkulie: Theoretische Klärungen und empirische Studien an betroffenen Schülerinnen und Schülern* (2. Auflage). Haupt Verlag.

Moser Opitz, E. (2016). Inklusiver Mathematikunterricht—Auch für Schülerinnen und Schüler mit dem Förderschwerpunkt Geistige Entwicklung (FGE). In *Inklusiver Mathematikunterricht—Mathematiklernen in ausgewählten Förderschwerpunkten* (Bd. 6, S. 57–60). University of Bamberg Press.

Moser Opitz, E., Freesemann, O., Grob, U., & Prediger, S. (2016). *BASIS-MATH-G 4+-5*. Hogrefe.

Moser Opitz, E., Grob, U., Wittich, C., Häsel-Weide, U., & Nührenbörger, M. (2018). Fostering the Computation Competence of Low Achievers Through Cooperative Learning in Inclusive Classrooms: A Longitudinal Study. *Learning Disabilities: A Contemporary Journal*, *16*(1), 19–35.

Moser Opitz, E., & Schmassmann, M. (2007). Grundoperationen. In *Didaktik des Unterrichts im Förderschwerpunkt Lernen* (S. 266–279). W. Kohlhammer.

Moser Opitz, E., Stöckli, M., Grob, U., Nührenbörger, M., & Reusser, L. (2019). *BASIS-MATH-G 3+*. Hogrefe.

Moser Opitz, E., Stöckli, M., Grob, U., Reusser, L., & Nührenbörger, M. (2020). *BASIS-MATH-G 2+*. hogrefe.

Mulligan, J. T., & Mitchelmore, M. C. (1997). Young Children's Intuitive Models of Multiplication and Division. *Journal for Research in Mathematics Education*, *28*(3), 309–330.

Nakagawa, S., Johnson, P. C. D., & Schielzeth, H. (2017). The coefficient of determination R^2 and intra-class correlation coefficient from generalized linear mixed-effects models revisited and expanded. *Journal of The Royal Society Interface*, *14*(134), 20170213.

Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining R^2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, *4*(2), 133–142.

Nesme, C. (in Vorbereitung). *Mit Kindern kooperative Aktivitäten reflektieren. Eine rekonstruktive Videoanalyse von Lehrpersonen-zentrierten Klassengesprächen.* Institut für Erziehungswissenschaft. Universität Zürich. Schweiz.

Nührenbörger, M. (2009). Interaktive Konstruktionen mathematischen Wissens. *JMD*, *30*(2), 147–172.

Obersteiner, A., Reiss, K., & Heinze, A. (2018). Psychological Theories in Mathematics Education: Introduction to the Special Issue. *Journal Für Mathematik-Didaktik*, *39*(1), 1–6.

Oehl, W. (1962). *Der Rechenunterricht in der Grundschule*. Hermann Schroedel.

Padberg, F. (2009). *Didaktik der Arithmetik: Für Lehrerausbildung und Lehrerfortbildung* (3., erw.völlig überarb. Aufl., Nachdr). Spektrum, Akad. Verl.

Padberg, F., & Büchter, A. (2015). *Einführung Mathematik Primarstufe—Arithmetik* (2. Auflage). Springer Spektrum.

Paek, I., & Wilson, M. (2011). Formulating the Rasch Differential Item Functioning Model Under the Marginal Maximum Likelihood Estimation Context and Its Comparison With Mantel–Haenszel Procedure in Short Test and Small Sample Conditions. *Educational and Psychological Measurement*, *71*(6), 1023–1046.

Peter-Koop, A., Rottmann, T., & Lüken, M. M. (Hrsg.). (2018). *Inklusiver Mathematikunterricht in der Grundschule* (4. Auflage). Mildenberger.

Pfister, M. (2016). *Adaptive Lernunterstützung im integrativen Mathematikunterricht: Eine Videostudie*. Universität Zürich.

Pfister, M., Stöckli, M., Moser Opitz, E., & Pauli, C. (2015). Inklusiven Mathematikunterricht erforschen: Herausforderungen und erste Ergebnisse aus einer Längsschnittstudie. *Unterrichtswissenschaft*, *01*, 53–66.

Pospeschill, M. (2010). *Testtheorie, Testkonstruktion, Testevaluation: Mit 77 Fragen zur Wiederholung*. Ernst Reinhardt Verlag.

Prediger, S. (2009). Inhaltliches Denken vor Kalkül. In *Fördernder Mathematikunterricht in der Sek. I* (S. 213–234). Beltz.

Prediger, S. (2010). „Aber wie sag ich es mathematisch?“ – Empirische Befunde und Konsequenzen zum Lernen von Mathematik als Mittel zur Beschreibung von Welt. In *Entwicklung naturwissenschaftlichen Denkens zwischen Phänomen und Systematik.* (S. 6–20). LIT-Verlag.

Prediger, S., & Götze, D. (2017). Sprachbildung als langfristige Entwicklungsaufgabe—Praktische Ansätze und ihre empirische Fundierung am Beispiel Algebra. In *Mathematik und Sprache* (Bd. 7, S. 9–24). University of Bamberg Press.

Prediger, S., & Wessel, L. (2013). Fostering German-language learners’ constructions of meanings for fractions—Design and effects of a language- and mathematics-integrated intervention. *Mathematics Education Research Journal*, *25*(3), 435–456.

Prediger, S., Wessel, L., & Özdil, E. (Hrsg.). (2011). *Mathematiklernen unter Bedingungen der Mehrsprachigkeit—Stand und Perspektiven der Formschung und Entwicklung in Deutschland* (S. 163–184). Waxmann.

Prediger, S., Wilhelm, N., Büchter, A., Gürsoy, E., & Benholz, C. (2015). Sprachkompetenz und Mathematikleistung – Empirische Untersuchung sprachlich bedingter Hürden in den Zentralen Prüfungen 10. *Journal für Mathematik-Didaktik*, *36*(1), 77–104.

Purpura, D. J., & Lonigan, C. J. (2013). Informal Numeracy Skills: The Structure and Relations Among Numbering, Relations, and Arithmetic Operations in Preschool. *American Educational Research Journal*, *50*(1), 178–209.

Radatz, H. (1991). Einige Beobachtungen bei rechenschwachen Grundschülern. *Untersuchungen zum Mathematikunterricht / Störungen beim Mathematiklernen*, *16*, 74–89.

Rasbash, J. (2019). *Module 4: Multilevel structures and classifications. LEMMA VLE*. University of Bristol, Centre for Multilevel Modelling. https://www.cmm.bris.ac.uk/lemma

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed). Sage Publications.

Rechsteiner-Merz, C. (2015). Rechnen entwickeln—Flexibilität fördern. In *Entwicklung mathematischer Fähigkeiten von Kindern im Grundschulalter* (Bd. 5, S. 55–70). University of Bamberg Press.

Revelle, W. (2020). *Package „psych“ Procedures for Psychological, Psychometric, and Personality Research*. Comprehensive R Archive Network CRAN. https://cran.r-project.org/web/packages/psych/psych.pdf

Rider, R. (2007). Shifting from Traditional to Nontraditional Teaching Practices Using Multiple Representations. *Mathematics Teacher*, *100*(7), 494–500.

Rittle-Johnson, B. (2017). Developing Mathematics Knowledge. *Child Development Perspectives*, *11*(3), 184–190.

Rittle-Johnson, B., & Star, J. R. (2007). Does comparing solution methods facilitate conceptual and procedural knowledge? An experimental study on learning to solve equations. *Journal of Educational Psychology*, *99*(3), 561–574.

Rittle-Johnson, B., & Star, J. R. (2011). The Power of Comparison in Learning and Instruction: Learning Outcomes Supported by Different Types of Comparisons. In *Psychology of Learning and Motivation* (Bd. 55, S. 199–225). Elsevier.

Robitzsch, A., Kiefer, T., & Wu, M. (2020). *Package „TAM“ Test Analysis Modules*. Comprehensive R Archive Network CRAN. https://cran.r-project.org/web/packages/TAM/TAM.pdf

Rödiger, K. (2017). Methodische Grundlagen. In K. Rödiger, *Empathie im Kundenkontakt* (S. 101–143). Springer Fachmedien Wiesbaden.

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.

Rosch, J. (2015). Eine Fallstudie zum Verstehen von Algebra im Mathematikunterricht. In G. Kadunz (Hrsg.), *Semiotische Perspektiven auf das Lernen von Mathematik* (S. 111–133). Springer Berlin Heidelberg.

Rost, D. H. (2007). *Interpretation und Bewertung pädagogisch-psychologischer Studien: Eine Einführung* (2., überarb. und erw. Aufl). Beltz.

Rost, J. (2004). *Lehrbuch Testtheorie—Testkonstruktion* (2., vollst. überarb. und erw. Aufl). Huber.

Rottmann, T. (2018). 5. Aufbau von Grundvorstellungen im inklusiven Mathematikunterricht—Erfahrungen aus der Bielefelder Beratungsstelle für Kinder mit Rechenstörungen. In *Inklusiver Mathematikunterricht in der Grundschule* (4. Auflage, S. 75–89). Mildenberger.

Royar, T. (2011). Zum Operationsverständnis der Grundrechenarten. *Beiträge zum Mathematikunterricht*, *Tagungsband BzMU* (45. Jahrestagung der GDM).

Royar, T. (2013). *Handlung - Vorstellung - Formalisierung: Entwicklung und Evaluation einer Aufgabenreihe zur Überprüfung des Operationsverständnisses für Regel- und Förderklassen*. Kovač.

Royar, T., Streit, C., & Ziska, S. (2014). Entwicklung eines Instruments zur Erfassung des Operationsverständnisses der Multiplikation. *Beiträge zum Mathematikunterricht*, 1019–1022.

Royar, T., Streit, C., & Ziska, S. (2015). *Ein Instrument zur Erfassung des Multiplikationsverständnisses bei Kindern*. PH FHNW, unpubliziertes Dokument.

Ruwisch, S. (1999). *Angewandte Multiplikation: Klassenfest, Puppenhaus, und Kinderbowle: eine qualitative empirische Studie zum Lösungsverhalten von Grundschulkindern beim Bearbeiten multiplikativer Sachsituationen*. P. Lang.

Ruwisch, S. (2001). Multiplikative Vorstellungen von Viert- und Sechstklässlern im Bereich natürlicher sowie Bruchzahlen. *Beiträge zur Didaktik der Mathematik für die Primärstufe*, *31*, 173–187.

Safi, F. (2009). *Exploring the Understanding of Whole Number Concepts and Operations: A case study analysis of prospective elementary school teachers*. University of Central Florida.

Schäfer, J. (2005). *Rechenschwäche in der Eingangsstufe der Hauptschule: Lernstand, Einstellungen und Wahrnehmungsleistungen; eine empirische Studie*. Kovač.

Scheid, H. (Hrsg.). (1981). *Die Mathematik. 1: 5. bis 10. Schuljahr* (4., völl. neu bearb. u. erw. Aufl).

Scherer, P. (2004). *Addition und Subtraktion im Hunderterraum* (2. Auflage). Persen.

Scherer, P. (2016). *Multiplikation und Division im Hunderterraum* (7. Auflage). Persen.

Scherer, P., & Moser Opitz, E. (2010). *Fördern im Mathematikunterricht der Primarstufe*. Spektrum Akademischer Verlag.

Schipper, W. (2002). Thesen und Empfehlungen zum schulischen und außerschulischen Umgang mit Rechenstörungen. *Journal für Mathematik-Didaktik*, *23*(3–4), 243–261.

Schipper, W. (2005). *Lernschwierigkeiten erkennen—Verständnisvolles Lerrnen fördern*. SINUS-Transfer Grundschule.

Schnepel, S. (2019). *Mathematische Förderung von Kindern mit einer intellektuellen Beeinträchtigung: Eine Längsschnittstudie in inklusiven Klassen*.

Schoppek, W. (2015). Mehrebenenanalyse oder Varianzanalyse?: Ein simulationsbasierter Vergleich von Verfahren zur Auswertung pädagogisch-psychologischer Experimente. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, *47*(4), 199–209.

Schulz, A., & Leuders, T. (2018). Learning trajectories towards strategy proficiency in multidigit division – A latent transition analysis of strategy and error profiles. *Learning and Individual Differences*, *66*, 54–69.

Schulz, A., Leuders, T., & Rangel, U. (2019). The Use of a Diagnostic Competence Model About Children's Operation Sense for Criterion-Referenced Individual Feedback in a Large-Scale Formative Assessment. *Journal of Psychoeducational Assessment*.

Schütte, S. (2008). *Qualität im Mathematikunterricht der Grundschule sichern: Für eine zeitgemässe Unterrichts- und Aufgabenkultur*. Oldenbourg.

Schwarzkopf, R. (2019). Produktive Kommunikationsanlässe im Mathematikunterricht der Grundschule: Zur lerntheoretischen Funktion des Argumentierens. In *Darstellen und Kommunizieren* (Bd. 9, S. 55–68). University of Bamberg Press.

Schweinberger, K. (2018). *Statistik II - Theorie und Anwendungen, Sitzung 08: Faktoranalytische Verfahren*. Universität Basel.

Selter, C., & Akinwunmi, K. (Hrsg.). (2014). *Mathe sicher können: Diagnose- und Förderkonzept zur Sicherung mathematischer Basiskompetenzen. 2: Handreichungen für eine Diagnose- und Förderkonzept zur Sicherung mathematischer Basiskompetenzen: Natürliche Zahlen: für Lehrerinnen und Lehrer*. Cornelsen.

Sfard, A. (2012a). Glossary of communicational (commognitive) terms. *International Journal of Educational Research*, *51–52*, 155–158.

Sfard, A. (2012b). Introduction: Developing mathematical discourse—Some insights from communicational research. *International Journal of Educational Research*, *51–52*, 1–9.

Sherin, B., & Fuson, K. (2005). Multiplication Strategies and the Appropriation of Computational Resources. *Journal for Research in Mathematics Education*, *36*(4), 347–395.

Sikora, S., & Voß, S. (2018). *Mathematikunterricht in der inklusiven Grundschule*. Verlag W. Kohlhammer.

Slavin, R. E. (2008). Perspectives on Evidence-Based Research in Education—What Works? Issues in Synthesizing Educational Program Evaluations. *Educational Researcher*, *37*(1), 5–14.

Slavit, D. (1999). The role of operation sense in transitions from arithmetic to algebraic thought. *Educational Studies in Mathematics*, 251–274.

Smith, S. Z., & Smith, M. E. (2006). Assessing Elementary Understanding of Multiplication Concepts. *School Science and Mathematics*, *Volume 106(3)*, 140–149.

Snijders, T. A. B., & Bosker, R. J. (2012). *Multilevel analysis: An introduction to basic and advanced multilevel modeling* (2nd ed). Sage.

Söbbeke, E. (2005). *Zur visuellen Strukturierungsfähigkeit von Grundschulkindern: Epistemologische Grundlagen und empirische Fallstudien zu kindlichen Strukturierungsprozessen mathematischer Anschauungsmittel*. Franzbecker.

Souvignier, E. (2007). Kooperatives Lernen. In *Didaktik des Unterrichts im Förderschwerpunkt Lernen* (S. 138–148). W. Kohlhammer.

Spiegel, H., & Selter, C. (2013). *Kinder & Mathematik: Was Erwachsene wissen sollten* (8. Aufl). Kallmeyer.

Stanat, P., Schipolowski, S., Rjosk, C., Weirich, S., & Haag, N. (Hrsg.). (2017). *IQB-Bildungstrend 2016: Kompetenzen in den Fächern Deutsch und Mathematik am Ende der 4. Jahrgangsstufe im zweiten Ländervergleich*. Waxmann.

stats.stackexchange.com. (2011). https://stats.stackexchange.com/questions/16331/doing-principal-component-analysis-or-factor-analysis-on-binary-data/16335#16335

Steenpass, A. (2014). *Grundschulkinder deuten Anschauungsmittel—Eine epistemologische Kontext- und Rahmenanalyse zu den Bedingungen der visuellen Strukturierungskompetenz*. Universität Duisburg-Essen.

Steenpass, A., & Söbbeke, E. (2009). Ich sehe das so. Und du? *Die Grunschulzeitschrift*, *23. Jahrgang*(Heft 222.223), 42–45.

Steiner, G. (1997). A V - Erwerb mathemtischer Kompetenzen—Kommentar. In *Entwicklung im Grundschulalter* (S. 171–179). Beltz Psychologie Verlags Union.

Stern, E. (1997). Kapitel V - Erwerb mathematischer Kompetenzen—Ergebnisse aus dem SCHOLASTIK - Projekt. In *Entwickung im Grundschulalter* (S. 157–170). Beltz Psychologie Verlags Union.

Stern, E. (2002). Wie abstrakt lernt das Grunschulkind? Neuere Ergebnisse der entwicklungspsychologischen Forschung. *Individuelles und soziales Lernen in der Grundschule - Kinderperspektive und pädagogische Konzepte*, *Hans Petillon (Hrsg.)*, 27–42.

Stingl, P. (2004). *Mathematik für Fachhochschulen: Technik und Informatik ; mit über 1000 Aufgaben und Lösungen* (7., durchges. Aufl). Hanser.

Stöckli, M. (2019). *Unterrichtsintegrierte Förderung im Mathematikunterricht: Eine empirische Studie in der Primarschule*. Universität Zürich.

Strasser, C. (2020). *Multiplikationsverständnis von Kindern in der zweiten Primarklasse: Entwicklung eines Testinstruments*. Universität Basel, Masterarbeit.

Strobl, C. (2015). *Das Rasch-Modell: Eine verständliche Einführung für Studium und Praxis* (3. erweiterte Auflage). Reiner Hampp Verlag.

Sun, X. H., Xin, Y. P., & Huang, R. (2019). A complementary survey on the current state of teaching and learning of Whole Number Arithmetic and connections to later mathematical content. *ZDM*, *51*(1), 1–12.

Takane, Y., & De Leeuw, J. (1987). On the Relationship between Item Response Theory and Factor Analysis of Discretized Variables. *Psychometrika*, *52/no 3*, 393–408.

Tarim, K., & Akdeniz, F. (2008). The effects of cooperative learning on Turkish elementary students' mathematics achievement and attitude towards mathematics using TAI and STAD methods. *Educational Studies in Mathematics*, *67*(1), 77–91.

Thomas, M. (2008). *Conceptual representations and versatile mathematical thinking*.

Tiedemann, K. (2017). Sprache trifft Material. In *Mathematik und Sprache* (Bd. 7, S. 41–56). University of Bamberg Press.

Tiedemann, K. (2018). 7. Sprache im inklusiven Mathematikunterricht der Grundschule. In *Inklusiver Mathematikunterricht in der Grundschule* (4. Auflage, S. 107–121). Mildenberger.

Tiedemann, K. (2020). Praktiken des Beschreibens – Zu Funktionen der Sprache bei der Erarbeitung des Teilschrittverfahrens im Zahlenraum bis 100. *Journal für Mathematik-Didaktik*, *41*(1), 11–41.

Till, C., & Sproesser, U. (2013). Eine Grafik sagt mehr als tausend Worte?! In *Mathematik lernen, darstellen, deuten, verstehen* (S. 205–213). Springer Spektrum.

Trendtel, M., Pham, G., & Yanagida, T. (2016). Skalierung und Linking. In *Large-Scale Assessment mit R* (S. 185–224). facultas.

Trendtel, M., Schwabe, F., & Fellinger, R. (2016). Differenzielles Itemfunktionieren in Subgruppen. In *Large-Scale Assessment mit R* (S. 111–148). facultas.

Ufer, S., & Bochnik, K. (2020). The Role of General and Subject-specific Language Skills when Learning Mathematics in Elementary School. *Journal Für Mathematik-Didaktik*, *41*(1), 81–117.

Ufer, S., Leiss, D., Stanat, P., & Gasteiger, H. (2020). Sprache und Mathematik – theoretische Analysen und empirische Ergebnisse zum Einfluss sprachlicher Fähigkeiten in mathematischen Lern- und Leistungssituationen. *Journal für Mathematik-Didaktik*, *41*(1), 1–9.

Venkat, H., & Mathews, C. (2019). Improving multiplicative reasoning in a context of low performance. *ZDM*, *51*(1), 95–108.

Verboom, L. (2017). Wege durch den Sprachförderschungel. In *Mathematik und Sprache* (Bd. 7, S. 57–72). University of Bamberg Press.

Vergnaud, G. (1983). Multiplicative Structures. In *Acquisition of mathematics concepts and processes* (S. 127–174). ACADEMIC Press.

Vergnaud, G. (1996). The Theory fo Conceptual Fields. In *Theories of mathematical learning* (S. 219–239). L. Erlbaum Associates.

Verschaffel, L., Greer, B., & De Corte, E. (2007). Whole Numer Concepts and Operations. In *Second Handbook of Research on Mathematics Teaching and Learning* (S. 557–628). Information Age Publishing Inc.

Vickers, A. J., & Altman, D. G. (2001). Statistics Notes: Analysing controlled trials with baseline and follow up measurements. *BMJ*, *323*, 1123–1124.

vom Hofe, R. (1995). *Grundvorstellungen mathematischer Inhalte*. Spektrum, Akad. Verlag.

Votta, F. (2017). *Multilevel Modelling mit R*. R-Group Stuttgart. https://rpubs.com/favstats/multilevel

Wagner, A., & Wörn, C. (2013). Veranschaulichungs- und Erklärmodelle zum Rechnen mit negativen Zahlen. In *Mathematik lernen, darstellen, deuten, verstehen* (S. 191–203). Springer Spektrum.

Walter, D., & Schulz, A. (2019). Darstellungen im Mathematikunterricht—Real, mental, digital. In *Darstellen und Kommunizieren* (Bd. 9, S. 39–54). University of Bamberg Press.

Walter, J., Suhr, K., & Werner, B. (2001). Experimentell beobachtete Effekte zweier Formen von Mathematikunterricht in der Förderschule. *Zeitschrift für Heilpädagogik*, *4*, 143.151.

Warm, T. A. (1989). Weighted likelihood estimation of ability in item response theory. *Psychometrika*, *54*(3), 427–450.

Wartha, S. (2010). Aufbau von Grundvorstellungen: Ein Förderkonzept. *Beiträge zum Mathematikunterricht*, *BzMU*.

Wartha, S. (2011). Handeln und Verstehen / Förderbaustein: Grundvorstellungen aufbauen. *mathematik lehren*, *166*, 8–14.

Wartha, S., & Schulz, A. (2013). *Rechenproblemen vorbeugen* (2. Auflage). Cornelsen.

Weigand, H.-G. (2015). Begriffsbildung. In *Handbuch der Mathematikdidaktik* (S. 255–278). Springer.

Weinhold Zulauf, M., Schweiter, M., & von Aster, M. (2003). Das Kindergartenalter: Sensitive Periode für die Entwicklung numerischer Fertigkeiten. *Diagnostik*, *Kindheit und Entwicklung 12 (4)*(4), 222–230.

Weiss, R. H., & Osterland, J. (2012). *CFT 1-R Grundintelligenztest Skala 1*. Hogrefe.

Weisshaupt, S., & Peucker, S. (2009). Entwicklung arithmetischen Vorwissens. In *Handbuch Rechenschwäche: Lernwege, Schwierigkeiten und Hilfen bei Dyskalkulie* (2. Auflage, erweitert und aktualisiert, S. 52–76). Beltz.

Werner, B. (2019). *Mathematik inklusive: Grundriss einer inklusiven Fachdidaktik*. Verlag W. Kohlhammer.

Wieseke, J., Geigenmüller, A., & Kraus, F. (2012). On the Role of Empathy in Customer-Employee Interactions. *Journal of Service Research*, *15*(3), 316–331.

Wittich, C. (2017). *Mathematische Förderung durch kooperativ-strukturiertes Lernen*. Springer Fachmedien Wiesbaden.

Wittmann, E. C. (Hrsg.). (2011). *Schweizer Zahlenbuch. 1, Schulbuch* (1. Aufl., 6, unveränd. Nachdr). Klett und Balmer.

Wittmann, E. C., & Müller, G. N. (2008). *Schweizer Zahlenbuch. 2, Schulbuch* (1. Aufl., [Nachdr.]). Klett und Balmer.

Wittmann, E. Ch. (2015). Strukturgenetsiche didaktische Analysen- empirische Forschung „erster Art“. *mathematica didactica*, *38*, 239–255.

Woltman, H., Feldstain, A., MacKay, J. C., & Rocchi, M. (2012). An introduction to hierarchical linear modeling. *Tutorials in Quantitative Methods for Psychology*, *8*(1), 52–69.

Zhang, S., Cao, Y., Wang, L., & Li, X. (2019). Characteristics of teaching and learning single-digit whole number multiplication in china: The case of the nine-times table. *ZDM*, *51*(1), 81–94.

Zhao, Y., Huen, J. M. Y., & Chan, Y. W. (2017). Measuring Longitudinal Gains in Student Learning: A Comparison of Rasch Scoring and Summative Scoring Approaches. *Research in Higher Education*, *58*(6), 605–616.

Zimmermann, M., & Bescherer, C. (2013). Repräsentationen „on demand“ bei mathematischen Beweisen in der Hochschule. In *Mathematik lernen, darstellen, deuten, verstehen* (S. 241–252). Springer Spektrum.

Zwick, R., Thayer, D. T., & Lewis, C. (1999). An Empirical Bayes Approach to Mantel-Haenszel DIF Analysis. *Journal of Educational Measurement*, *36*(1), 1–28.