Stand: April 2020

Mathes 1

Screening zur Erfassung der Mathematikleistungen in Klasse 1

Manual

www.lernlinie.de/to/mathes1

Inhalt

Abbildungsverzeichnis

Tabellenverzeichnis

Kurzinformation

Titel, Autoren, Jahr	Mathes 1, Simon Sikora & Stefan Voß, 2016
Diagnostische Zielsetzung	Erfassung der arithmetischen Kompetenzen der ersten Klassenstufe bei allen Kindern
Anwendungsbereiche	Grundschulmathematikunterricht in der ersten Klasse
Aufbau	44 Aufgaben aus den Bereichen Zahlbegriffsverständnis sowie den Grundrechenarten Addition und Subtraktion
Anwendungszeitraum	Mitte Klasse 1 (20. / 21. Schulwoche) Anfang Klasse 2 (3. / 4. Schulwoche)
Durchführung	Gruppentestung im Klassenkontext durch die Lehrkraft bei geführter Testbearbeitung innerhalb einer Unterrichtsstunde
Auswertung	Auswertung mithilfe von Auswertungsvorlagen
Normen	Totalerhebung in ländlicher und städtischer Region in Mecklenburg-Vorpommern (N = 1648)
Reliabilität	Interne Konsistenz: α = .94 (Mitte Klasse 1; N = 1207) α = .95 (Anfang Klasse 2; N = 1218) Retest-Reliabilität: r_{tt} = .74** (N = 911)
Validität	Konstruktvalidität: konstruktnahe Korrelation: mit DEMAT 1+: r = .82** (N = 1082) konstruktferne Korrelationen: mit KFT 1-2 R: r = .56** (N = 1155); mit WLLP: r = .46** (N = 1156) prognostische Validität: Korrelation Mathes 1 (Mitte Klasse 1) zum DEMAT 1+ (Ende Klasse 1): r = .73** (N = 872) RATZ-Index Mathes 1 (Mitte Klasse 1) zum DEMAT 1+ (Ende Klasse 2): 0.71 (N = 870)

1. Theoretische Grundlagen

1.1 Zielstellung der Mathes-Testreihe

Täglich stellen sich Lehrkräfte Fragen wie die folgenden:

> *„Lernen alle Kinder in meinem Unterricht erfolgreich oder kommt jemand nicht mit?“*
> *„Kann ich das aktuelle Themengebiet abschließen und im Stoff weitermachen oder brauchen die Schülerinnen und Schüler noch mehr Lernzeit?“*
> *„Was genau hat die Schülerin bzw. der Schüler noch nicht verstanden?“*

Solche Fragen haben das Ziel, den Unterricht bestmöglich an die individuellen Lernausgangslagen und Förderbedürfnisse der Kinder anzupassen. Die Mathes-Testreihe soll Lehrkräfte bei vertretbarem Aufwand dabei unterstützen, zu einer präzisen Einschätzung der aktuellen Lernstände sowie der Leistungsentwicklung der Schülerinnen und Schüler zu gelangen. Die Testverfahren können in dreierlei Hinsicht helfen, nämlich

- beim genauen Einschätzen des Spektrums der Leistungen der Schülerinnen und Schüler, um den Unterricht daran bestmöglich anpassen zu können,
- beim rechtzeitigen Erkennen derjenigen Schülerinnen und Schüler mit Risiken bzw. bereits ausgeprägten Schwierigkeiten im Kompetenzerwerb sowie
- bei der Planung effektiver Fördermaßnahmen, insbesondere für diejenigen Schülerinnen und Schüler mit besonderen Unterstützungsbedarfen.

Zu diesem Zweck wurden Mathes-Tests für jede Klassenstufe entwickelt, welche jeweils zum Beginn und in der Mitte des Schuljahres eingesetzt werden können.

1.2 Exkurs: Klassifikation und Epidemiologie von mathematischen Lernschwierigkeiten - Was soll verhindert werden?

Schwierigkeiten beim Rechnen- bzw. mathematischen Lernen gehören für viele Schülerinnen und Schüler (und deren Lehrerinnen und Lehrer) zum schulischen Alltag. Allerdings erhält nicht jedes Kind mit schwachen Leistungen in Mathematik die Diagnose „Dyskalkulie“ bzw. „Rechenstörung“. Nach den Kriterien der Weltgesundheitsorganisation leidet ein Kind nur dann unter einer Dyskalkulie, wenn seine Beeinträchtigung der Rechenfertigkeiten im Gegensatz sowohl zur allgemeinen Intelligenz als auch zu anderen schulischen Leistungen, z. B. dem Lesen und der Rechtschreibung, steht (sog. doppeltes Diskrepanzkriterium):

> *„Diese Störung beinhaltet eine umschriebene Beeinträchtigung von Rechenfertigkeiten, die nicht allein durch eine allgemeine Intelligenzminderung oder eine unangemessene Beschulung erklärbar ist. Das Defizit betrifft die Beherrschung grundlegender Rechenfertigkeiten wie Addition, Subtraktion, Multiplikation und Division, weniger die höheren mathematischen Fertigkeiten, die für Algebra, Trigonometrie, Geometrie oder Differential- und Integralrechnungen benötigt werden."* (ICD-10, Dilling, Mombour & Schmidt, 2011, S. 338)

Das Verwenden des Diskrepanzkriteriums zur Bestimmung von Kindern mit Rechenschwierigkeiten sowie eine daran gebundene Zuweisung von Fördermaßnahmen wird in der Fachliteratur aus verschiedenen Gründen kritisiert (u. a. Gaidoschik, 2011; Hartke & Diehl, 2013; Krajewski, 2003; Lorenz, 2005; Moser Opitz, 2004; Koch, 2005; Koch & Knopp, 2010). Gaidoschik stellt berechtigterweise die Frage: „Verdient denn ein Kind, das nicht nur im Rechnen, sondern auch beim Lesen Probleme hat, weniger Förderung in Mathematik als jenes, welches dem ‚Diskrepanz-Kriterium' genügt?" (2011, S. 12). „Es erscheint hingegen sinnvoller, all jene Kinder in die Förderung aufzunehmen, deren Lernfortschritte, durch welche Gründe auch immer, als unzureichend angesehen werden" (Lorenz, 2005, S. 15).

Während nur etwa 4 % bis 8 % aller Schülerinnen und Schüler den Kriterien einer Dyskalkulie bzw. Rechenstörung entsprechen (von Aster, Schweiter & Weinhold Zulauf, 2007; Lorenz, 2014), gehen Hasselhorn, Marx und Schneider (2005) vor dem Hintergrund der Befunde einschlägiger Prävalenz- und Schulleistungsstudien wie IGLU (Bos et al., 2003) oder TIMSS (Bos et al., 2008; Bos, Wendt, Köller & Selter, 2012; Selter, Walter, Walther & Wendt, 2016) davon aus, dass etwa 20 % aller Viertklässlerinnen und Viertklässler im Fach Mathematik Leistungsrückstände im Umfang von zwei Schuljahren aufweisen.

Schwierigkeiten im Fach Mathematik kommen bei Jungen und Mädchen etwa gleich häufig vor (Jacobs & Petermann, 2012; Landerl & Kaufmann, 2008). Bei vielen Schülerinnen und Schülern treten mathematische Lernschwierigkeiten nicht isoliert auf, sondern in Kombination mit Lese-Rechtschreibschwächen und psychischen Auffälligkeiten, insbesondere ADHS, Ängste und Depressionen (zusammenfassend Lambert, 2015; Sikora & Voß, 2018).

Die berichteten Befunde machen deutlich, dass statistisch gesehen etwa jedes fünfte Kind besondere Unterstützung im Fach Mathematik benötigt. „Frühzeitig zu erkennen, wenn Kinder Schwierigkeiten beim Erlernen mathematischer Begriffe haben, ist vermutlich der wichtigste Schritt auf dem Weg zur Förderung" (Hasemann & Gasteiger, 2014, S. 151). Die Verfahren der Mathes-Testreihe sollen Lehrkräfte dabei unterstützen.

2 Theoriebasierte Testentwicklung

Aufgrund der Relevanz der Arithmetik für den weiteren mathematischen Kompetenzerwerb wurde beschlossen, die Mathes-Testreihe in den ersten beiden Grundschuljahren auf den Inhaltsbereich *Zahlen und Operationen* zu begrenzen, wenngleich auch im Anfangsunterricht bereits geometrische und stochastische Fragestellungen sowie der Umgang mit Größen thematisiert werden sollten (KMK, 2005). Diese Entscheidung war notwendig, um dem Anspruch der Verfahren einer zeitökonomischen Leistungsmessung ganzer Klassen gerecht werden zu können.

2.1 Welche Kompetenzen erwerben Schülerinnen und Schüler in der ersten Klasse?

Die meisten Kinder kommen mit einem enormen Wissen über Zahlen und Operationen in die Schule (zusammenfassend Sikora, 2017). „Die verschiedenen Untersuchungen zeigen […], dass Schulanfänger zu unterschiedlichen mathematischen Themengebieten äußerst beachtliche, im Detail aber sehr unterschiedliche Vorerfahrungen und Vorkenntnisse besitzen“ (Käpnick, 2014, S. 73). So sitzen den Lehrkräften in vielen Klassen Kinder gegenüber, die bereits sicher rechnen können, und möglicherweise auch solche, die noch nicht mal alle Zahlwörter kennen. Eine zentrale Aufgabe des Mathematikunterrichts in der ersten Klassenstufe ist es daher, eine gemeinsame Wissensbasis für den weiteren mathematischen Lernprozess zu schaffen. Dazu gilt es zunächst, aspektreiche Vorstellungen von Zahlen aufzubauen, was typischerweise im Zahlenraum bis 20 geschieht (Krauthausen, 2018). Grundlegend sind die *kardinale* und die *ordinale* Sichtweise auf Zahlen, welche nach Wittmann und Müller (2012) die erste fundamentale Idee der Arithmetik darstellen.

Dass Zahlen eine Reihe bilden, wissen bereits die meisten dreijährigen Kinder, zumindest implizit. Dieser sogenannte *Ordinalzahlaspekt* wird typischerweise noch einmal unterteilt, und zwar in den Zählaspekt, also die Folge der Zahlen, welche beim Zählen durchlaufen wird, und in den Ordnungszahlaspekt, welcher den Platz einer Zahl in der Zahlenreihe angibt (z. B. 7 als Vorgänger von 8 und als Nachfolger von 6, aber auch der drittgrößte Junge der Klasse, etc.; Padberg & Benz, 2011).

Beim *Kardinalzahlaspekt* (auch Anzahlaspekt bezeichnet) wird eine Menge von Elementen in den Blick genommen (ebd.). Für die Zähl- und Rechenentwicklung ist die Einsicht von fundamentaler Bedeutung, dass strukturierte, d. h. regelhafte, Mengendarstellungen gegenüber unstrukturiert bzw. unübersichtlich präsentierten effektiver erfasst werden können. Die Ausnutzung von Strukturen wird umso wichtiger, je größer die Mengen werden, mit denen im Unterricht gearbeitet wird. In den einschlägigen didaktischen Lehrbüchern zum mathematischen Anfangsunterricht wird deshalb empfohlen, „gezielt zu einer *strukturierten, nicht-zählenden* Anzahl*erfassung* hinzuführen (Gaidoschik, 2010, S. 216; Hervorhebungen im Original). Diese erfordert „ein *bewusstes Wahrnehmen* der Teilanzahlen und ein *wissendes Zusammensetzen* dieser Teilanzahlen zum Zahlganzen“ (ebd., S. 217; Hervorhebungen im Original). Damit die Kinder Strukturen erkennen und nutzen können, müssen diese natürlich in den Mengendarstellungen auch vorhanden sein. Als günstig wurde die sogenannte Kraft der Fünf (Flexer, 1986) herausgearbeitet, d. h. die Bündelung der Einzelelemente zu Fünfern. Unser Stellenwertsystem gibt zudem eine weitere grundlegende Struktur vor, die sogenannte Kraft der Zehn, also die Zusammenfassung von zehn Einern zu einem Zehner.

Eine zentrale Voraussetzung für das Verständnis der Rechenoperationen ist der Erwerb des sogenannten Teil-Ganzes-Konzepts, welches in renommierten Entwicklungsmodellen früher mathematischer Kompetenzen (Fritz & Ricken, 2008; Krajewski, 2013) als wichtiger Meilenstein herausgestellt wird. Dieses Teil-Ganzes-Konzept beschreibt die Idee, dass eine Menge bzw. eine Zahl (das „Ganze“) auf verschiedene Weisen (in „Teile“) zerlegt und wieder zusammengesetzt werden kann, beispielsweise die Zahl 7 als Menge von 7 Objekten in Teilmengen von 3 und 4 Objekten. Bei der Addition werden zwei Teile zu einem Ganzen zusammengefügt (die Teile 3 und 4 zu 7), bei der Subtraktion wird ein Teil aus dem Ganzen herausgelöst, sodass der andere Teil des Ganzen übrig bleibt (von 7 werden 3 abgezogen, 4 bleiben übrig). Auch für die Anwendung tragfähiger Rechenstrategien ist das

Verständnis des Teil-Ganzes-Konzepts grundlegend. So wird beispielsweise bei der Nutzung des Teilschrittverfahrens zur Lösung der Aufgabe 8 + 7 = 15 der zweite Summand in zwei Teile zerlegt, nämlich in 2 und 5. Aufgrund der großen Bedeutung für den weiteren mathematischen Kompetenzerwerb verwundert es nicht, dass Winter (2001) die Teil-Ganzes-Relation zu den sieben Grundideen der Grundschulmathematik zählt.
Neben den verschiedenen Zahlaspekten werden auch die Rechenoperationen *Addition* und *Subtraktion* in der ersten Klassenstufe ausgiebig behandelt (Krauthausen, 2018). Während Schulanfängerinnen und Schulanfänger Additions- und Subtraktionsaufgaben in der Regel durch Vorwärts- bzw. Rückwärtszählen lösen (zusammenfassend Sikora, 2017), ist das zentrale Ziel des mathematischen Anfangsunterrichts, zukunftsfähigere Rechenstrategien zu entwickeln. Dazu gilt es zunächst, ein solides Verständnis der neuen Rechenoperationen aufzubauen (Hasemann & Gasteiger, 2014), d. h. es geht darum, Einsichten zu gewinnen, was die Addition und die Subtraktion als mathematische Ideen ausmacht. In der mathematikdidaktischen Literatur werden verschiedene *Grundvorstellungen* für die Addition und die Subtraktion unterschieden (Padberg & Benz, 2011), welche für den weiteren Kompetenzerwerb von zentraler Bedeutung sind. Dies sind in der Addition insbesondere das Hinzufügen und das Zusammenfassen (Wartha & Schulz, 2013) und in der Subtraktion das Abziehen, das Vergleichen sowie das Ergänzen (Fromme, Wartha & Benz, 2011). Kinder sollten in der ersten Klassenstufe ausreichend Gelegenheiten haben, sich mit mathematisch unterschiedlich strukturierten (Sach)Situationen im Sinne der aufgeführten Grundvorstellungen auseinanderzusetzen, um ein tiefes Verständnis für beide Rechenoperationen aufzubauen. Anschließend ist es das Ziel, Zählstrategien abzulösen und höherwertige, sogenannte heuristische, Strategien zur Lösung der Aufgaben zu erwerben (Häsel-Weide, Nührenbörger, Moser Opitz & Wittich, 2019). Diese Ableitungsstrategien machen sich die Strukturen des Zahlenraumes und die Beziehungen zwischen Aufgaben zunutze. Unter Rückgriff auf einfache oder bereits bekannte Aufgaben wird so das Rechnen erleichtert, beispielsweise indem Verdopplungen bzw. Halbierungen, Analogien oder Nachbaraufgaben für die Lösung einer unbekannten Aufgabe genutzt werden. Solche Strategien sind effizienter und eleganter als zählendes Rechnen, das zwar bei kleineren Zahlen noch gut funktionieren mag, in den immer größer werdenden Zahlenräumen der nachfolgenden Schuljahre jedoch versagt (Padberg & Benz, 2011).

2.2 Konstruktion inhaltsvalider Testanforderungen

Die im vorherigen Abschnitt 2.1 durchgeführte Analyse der mathematikdidaktischen Fachliteratur hat im Bereich der Arithmetik die folgenden inhaltlichen Schwerpunkte in der ersten Klassenstufe aufgezeigt:

1. kardinales und ordinales Zahlbegriffsverständnis
2. Teile-Ganzes-Konzept
3. Grundvorstellungen zur Addition und Subtraktion
4. Kopfrechnen in der Addition und Subtraktion

Zu 1) Der *kardinale Zahlaspekt* wird durch Aufgaben zur Anzahlerfassung abgeprüft. Die Mengen werden gebündelt angeordnet, wodurch eine strukturierte Erfassung begünstigt wird.

Um das Wissen über den Aufbau des Zahlenraumes abzubilden, müssen Zahlen auf dem Zahlenstrahl positioniert werden. Dieses Darstellungsmittel ist im Anfangsunterricht gebräuchlich (Käpnick, 2014), zudem können die zentralen Strukturen Kraft der Fünf und Kraft der Zehn daran veranschaulicht werden. Zusätzlich müssen Zahlen miteinander verglichen werden. Beide Anforderungen entsprechen dem *Ordnungsaspekt*. Der *Zählaspekt* wird durch Aufgaben erfasst, in denen Zahlenreihen weitergeführt werden müssen. Der Schwierigkeitsgrad der Aufgaben wird durch verschiedene Zählanforderungen variiert (vollständiges Auszählen, Weiterzählen, Zählen in Zählschritten; Padberg & Benz, 2011).

Zu 2) Das *Teil-Ganzes-Konzept* wird durch Zahlzerlegungen operationalisiert. Zur Erzeugung verschiedener Anforderungsniveaus sind gerade und ungerade Zahlen zu zerlegen, auch eine Aufteilung in vier Teile wird gefordert.

Zu 3) Ob die Kinder über *grundlegende Operationsvorstellungen bezüglich der Addition und Subtraktion* verfügen, wird durch Sachsituationen überprüft, welche in Situationsbildern (Käpnick, 2014) sowie in Textform vorgegeben werden, welche von der Lehrkraft vorgelesen werden. Die Aufgaben sprechen die Vorstellungen des Hinzufügens, Abziehens und Vergleichens an.

Zu 4) Die Fähigkeit, *Additions- und Subtraktionsaufgaben im Kopf* zu lösen, wird durch eine Vielzahl von Aufgaben verschiedener Schwierigkeitsgrade erfasst. Dabei wird die gesamte Bandbreite der möglichen Gleichungen in diesem Zahlenraum abgebildet. So werden sowohl sehr einfache Aufgaben vom Typ E ± E ohne Zehnerübergang dargeboten, welche sicherstellen, dass „Mathes 1“ auch bei Kindern mit Lernrückstanden im Fach Mathematik hinreichend differenziert. Außerdem wurden auch schwierige Aufgaben in das Verfahren aufgenommen. Am schwierigsten sind Rechnungen mit Zehnerüberschreitung. Die Aufgaben werden zudem nicht nur in der gewohnten Form präsentiert, bei denen die Summe bzw. Differenz zu ermitteln ist, sondern auch als Gleichungen mit Platzhaltern. Bei diesen müssen verschiedene Grundvorstellungen aktiviert und vernetzt werden, insbesondere das Gleichungsverständnis und das Teil-Ganzes-Konzept, was viele Kinder vor besondere Herausforderungen stellt, insbesondere zählend rechnende (Gaidoschik, 2011).

Insgesamt wurden 44 Aufgaben konstruiert. Die Tabelle 1 fasst die inhaltlichen Anforderungen des „Mathes 1“ zusammen.

Tabelle 1: inhaltliche Anforderungen des "Mathes 1"

Aufgabengruppe bzw. Anforderung	**Anzahl der Items**
Zahlen-Mengen-Zuordnung	3
Zahlenreihen ergänzen	4
Zahlen positionieren	2
Zahlvergleiche	3
Zahlzerlegungen	4
Operationsvorstellungen	3
Kopfrechnen Addition	7
Kopfrechnen Subtraktion	7
Gleichungen mit Platzhaltern	8
Textaufgaben	3

2.3 Überlegungen zur Steigerung der Testökonomie

Zur Erreichung einer hohen Testökonomie und Zumutbarkeit der Durchführung des „Mathes 1“ wurden folgende Maßnahmen getroffen:

- geführte Testbearbeitung innerhalb einer Unterrichtsstunde
- kurze und prägnante Aufgabenstellungen bzw. Operatoren
- weitgehend sprachfreie Gestaltung der Aufgaben, soweit zweckmäßig mit einem Beispiel und/oder grafischer Visualisierung
- kindlich angemessene, ansprechende Gestaltung des Testhefts mit Identifikationsfigur („Mathes – der Matheaffe“)
- freie Verfügbarkeit des Verfahrens durch Veröffentlichung unter Creative-Commons-Lizenz
- Testheftformat A5 mit Ermöglichung eines schwarz/weiß-Ausdrucks
- computergestützte Aufbereitung der Testergebnisse auf Kind- und Klassenebene

3 Testanwendung

Hinweis: Alle für die Durchführung und Auswertung des „Mathes 1“ benötigten Informationen sowie die Testinstruktionen finden sich in übersichtlicher Form in den Durchführungshinweisen.

3.1 Anwendungszeitraum und Zielgruppe

Vor dem Hintergrund der in Abschnitt 1.1 dargestellten Zielstellungen der Mathes-Testreihe, sind alle Mathes-Verfahren für die zeitökonomische Erfassung der Mathematikleistungen aller Schülerinnen und Schüler einer (inklusiven) Grundschulklasse konzipiert.
„Mathes 1“ kann zum Halbjahr der ersten und zu Beginn der zweiten Klassenstufe eingesetzt werden. Die nachfolgende Tabelle 2 stellt die Durchführungszeiträume dar.

Tabelle 2: Durchführungszeiträume des "Mathes 1"

	Schulwoche	Kommentar
MZP 1	20./21.	vor dem Halbjahreszeugnis
MZP 2	3./4.	Anfang Klasse 2

3.2 Testmaterial

Für die Durchführung des „Mathes 1“ werden folgende Materialien benötigt:

- 1 Testheft pro Kind,
- 1 Füller, 1 Bleistift,

- 1 Testheft für die Lehrkraft zur Demonstration,
- 1 Durchführungsanleitung für die Lehrkraft,
- Symbolkarten 1-3 (A 4-Format) für die Lehrkraft zur Demonstration.

3.3 Hinweise zur Testdurchführung

Die Durchführung des „Mathes 1“ erfolgt in Gruppen (Klassenverband). Um eine objektive Testanwendung zu gewährleisten, müssen folgende Punkte beachtet werden:

- Es muss gewährleistet sein, dass die Schülerinnen und Schüler in einer ruhigen, störungsfreien Atmosphäre die Aufgaben bearbeiten.
- Es muss sichergestellt werden, dass die Schülerinnen und Schülern ausreichend Zeit für die Aufgabenbearbeitung haben. Insgesamt werden etwa 40 Minuten für die Testdurchführung benötigt.
- „Mathes 1“ sollte möglichst ohne Pause durchgeführt werden.
- Die Durchführungshinweise sind zu berücksichtigen und die Testinstruktionen wörtlich vorzulesen.
- Die zu lösenden Aufgaben dürfen vorab nicht mit den Schülerinnen und Schülern geübt werden.
- Die Testhefte werden erst ausgeteilt, nachdem der erste Abschnitt der wörtlichen Instruktionen vorgelesen wurde.
- Die Schülerinnen und Schüler dürfen während der Durchführung keine Hinweise und Hilfestellungen erhalten. Ermutigungen sind erlaubt.
- Es ist darauf zu achten, dass die Kinder nicht voneinander abschreiben.
- „Mathes 1“ sollte möglichst nicht in der letzten Unterrichtsstunde und nicht im Nachmittagsunterricht durchgeführt werden.

3.4 Auswertung und Interpretation

3.4.1 Manuelle Auswertung

Die Auswertungsvorlage unterstützt eine objektive und ökonomische Auswertung des „Mathes 1“. Alle richtig gelösten Aufgaben werden mit einem Punkt, falsch gelöste mit null Punkten bewertet. Die erreichten Punkte werden aufsummiert. Mithilfe der Normtabelle im Anhang kann die Testleistung des Kindes mit denen gleichaltriger Schülerinnen und Schüler verglichen werden. Dazu stehen die für statusdiagnostische Einschätzungen gängigen Standardwerte (Prozentrang und T-Wert) für beide Messzeitpunkte zur Verfügung (s. Anhang C, S. 24).
Differenzierte Informationen zur Auswertung des „Mathes 1“ liefern die Durchführungshinweise des Verfahrens.

3.4.2 Automatisierte Auswertung

Für Lehrkräfte aus Mecklenburg-Vorpommern wird über die Homepage www.lernlinie.de eine internetgestützte Auswertung des „Mathes 1“ angeboten. Bei dieser Variante müssen lediglich die erreichten Rohwerte der Kinder mithilfe der Auswertungsvorlage wie in Abschnitt 3.4.1 beschrieben ermittelt und auf der Internetplattform eingetragen werden. Anschließend werden automatisiert Ergebnisübersichten für jedes Kind erstellt, sodass auf einen Blick ersichtlich ist, wie seine Leistungen im Vergleich zu gleichaltrigen Schülerinnen und Schülern einzuschätzen sind. Bei mehrmaligem Einsatz des „Mathes 1“ stellt die Internetplattform ebenfalls die Lernentwicklung des Kindes graphisch dar. Zudem besteht die Möglichkeit, die Ergebnisse aller Schülerinnen und Schüler einer Klasse im Überblick anzuzeigen.

Lehrkräfte außerhalb Mecklenburg-Vorpommerns können die Testergebnisse ihrer Schülerinnen und Schüler in die vorbereitete Klassenübersicht eintragen, welche automatisch den erreichten Rohwerten die Prozentränge und T-Werte zuordnet und die in Abschnitt 3.4.3 aufgeführten Referenzniveaus graphisch veranschaulicht.

3.4.3 Interpretation der Ergebnisse

Bei der Einschätzung der Testleistung eines Kindes helfen sogenannte Referenzniveaus, welche auf den Prozentrang Bezug nehmen und diesen vereinfachend interpretieren, indem die Testleistung des Kindes im Vergleich zur Referenzgruppe in fünf Stufen eingeordnet wird. Referenzniveaus stellen Marker an empirisch bedeutsamen Grenzen dar (Prozentrang 10, 25, 75 und 90). Ein Prozentrang von 10 bedeutet, dass 10 Prozent der gleichaltrigen Schülerinnen und Schüler gleiche oder schlechtere Leistungen aufweisen, ein Prozentrang von 25, dass ein Viertel der Kinder ein gleiches oder schlechteres Testergebnis erzielt, usw. Auf diese Weise entstehen fünf Leistungsbereiche, sodass einfach ersichtlich ist, wie das Kind im Vergleich zu Gleichaltrigen abgeschnitten hat.

	weit überdurchschnittlich		**> 90.**
	überdurchschnittlich		**≥ 75.**
Schülerleistungen gelten als	**durchschnittlich**	bei einem Prozentrang	**< 75 und > 25.**
	unterdurchschnittlich		**≤ 25.**
	weit unterdurchschnittlich		**≤ 10.**

Abbildung 1: Referenzniveaus als Interpretationshilfen für die erzielte Testleistung

4 Testgütekriterien

„Mathes 1" ist eine Kurzversion des im Rahmen der Dissertation von Knopp (2010) entwickelten Inventars „Rechenfische" (Knopp, 2010). Die Daten aus der Dissertationsstudie wurden reanalysiert, um Informationen zur Testgüte des „Mathes 1" zu gewinnen.

Die Untersuchung war in die Mecklenburger Längsschnittstudie (Koch, Hartke & Blumenthal, 2009; Blumenthal, Hartke & Koch, 2010) eingebettet, welche von 2006 bis 2010 in allen Grundschulen der Hansestadt Rostock und der Insel Rügen durchgeführt wurde. In deren Rahmen wurde die Schulleistungsentwicklung der Kinder, die im Schuljahr 2006 / 2007 eingeschult wurden, über die Grundschulzeit verfolgt. Dabei kamen am Ende jeder Klassenstufe ein Intelligenz-, ein Lese- und ein Mathematiktest zum Einsatz, wodurch Einschätzungen zur Validität des „Mathes 1" möglich werden (s. Abschnitt 4.4).

Für die Evaluation der psychometrischen Güte des „Mathes 1" werden die Daten von zwei Messzeitpunkten in der ersten Klassenstufe herangezogen, welche im Januar 2007 (ca. 20. Schulwoche) sowie im Juni und Juli 2007 (ca. 40. Schulwoche) stattfanden. Zum zweiten Erhebungszeitpunkt wurden mit dem „Deutscher Mathematiktest für erste Klassen" (DEMAT 1+; Krajewski, Küspert & Schneider, 2002), dem „Kognitiver Fähigkeitstest für 1. und 2. Klassen" (KFT 1-2 R; Kawthar & Perleth, 2005) sowie der „Würzburger Leise Leseprobe" (WLLP; Küspert & Schneider, 1998) drei weitere Messverfahren durchgeführt.

Die Stichprobe verteilte sich auf 37 Schulen mit 87 Klassen. Insgesamt wurden 1648 Schülerinnen und Schüler in die Mecklenburger Längsschnittstudie einbezogen. Mit 869 Jungen (53 %) und 779 Mädchen (47 %) waren die Geschlechter annähernd gleichverteilt. Zum Zeitpunkt der Einschulung waren die Kinder im Durchschnitt 6;8 Jahre alt.

4.1 Itemkennwerte

Für eine optimale Differenzierungsfähigkeit bei hoher Messeffizienz ist eine gute Passung zwischen den Aufgabenschwierigkeiten und den Personenfähigkeiten wünschenswert (Prenzel & Blum, 2007). Dazu sollte der Test über das gesamte Schuljahr hinweg alle Leistungsbereiche abdecken und möglichst zu keinem Messzeitpunkt Boden- bzw. Deckeneffekte aufweisen. Ob die Schwierigkeit der Aufgaben des „Mathes 1" für Schülerinnen und Schüler erster Klassen angemessen ist, wurde mittels der Übereinstimmung der im Rasch-Modell (s. Abschnitt 4.2) geschätzten Item- und Personenparameter geprüft (Carstensen & Taskinen, 2007). Die nachfolgende Abbildung 2 stellt die Verteilungen der Item- und Personenparameter des „Mathes 1" zu den beiden Messzeitpunkten gegenüber.

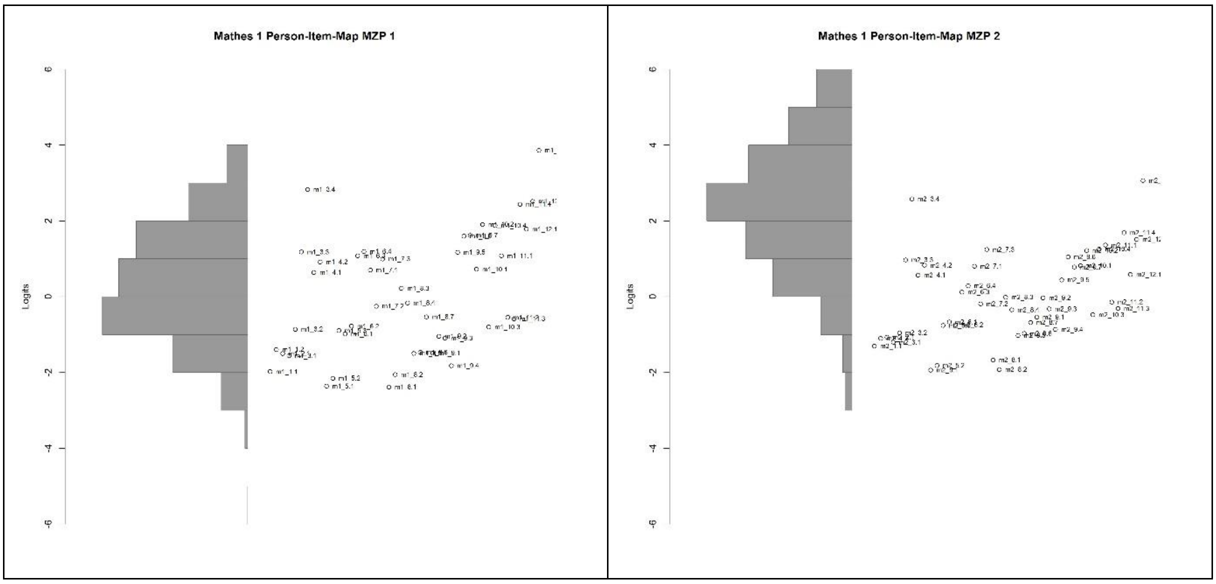

Abbildung 2: Verteilungen von Aufgabenschwierigkeiten und Schülerfähigkeiten zu den beiden Messzeitpunkten

Aus den Darstellungen geht hervor, dass „Mathes 1“ den unteren und mittleren Leistungsbereich gut bis sehr gut abdeckt. Zum ersten Messzeitpunkt ist auch noch eine ausreichende Anzahl an schwierigen Items vertreten, zum zweiten ist die Differenzierungsfähigkeit für leistungsstarke Schülerinnen und Schüler eingeschränkt.

Aus Gründen der Lesbarkeit werden die Trennschärfekoeffizienten an dieser Stelle nicht berichtet. Eine tabellarische Übersicht mit Angaben zur Schwierigkeit und Trennschärfe aller Aufgaben des „Mathes 1“ zu beiden Messzeitpunkten befindet sich im Anhang A auf Seite 22.

4.2 Gültigkeit des Messmodells

Dem Verfahren wurde das dichotome Rasch-Modell (Rasch, 1960) als Messmodell zugrunde gelegt. Dabei wurde die Technik virtueller Personen genutzt (Rost, 2004). Die Gültigkeit des Rasch-Modells wurde durch eine Analyse lokaler Modellverletzungen eingeschätzt. Die Rasch-Analysen wurden mit dem Statistikprogramm R (R Core Team, 2013) mithilfe des Pakets pairwise (Heine, 2020) durchgeführt. Die Modellpassung der Items wurde anhand ihrer geschätzten Infit-Werte beurteilt. Da die Outfit-Statistiken deutlich durch Ausreißerwerte beeinflusst werden, Infit-Werte hingegen sensitiver im Bereich mittlerer Fähigkeitsausprägungen ausfallen (Linacre, 2002), wurden für den „Mathes 1“ in erster Linie die Infit-Statistiken auf Abweichungen vom Erwartungswert 1 untersucht. In Anlehnung an Bond und Fox (2015) zeigen bei gewöhnlichen Stichprobengrößen Werte von $0.7 \leq \text{Infit} \leq 1.3$ an, dass das Item den Annahmen des Rasch-Modells entspricht.

In der Abbildung 3 werden die Ergebnisse der Raschskalierung präsentiert. Die Infit-Werte wurden aus Gründen der Übersichtlichkeit mittels Boxplots graphisch aufbereitet, sodass die Streuung der Werte des „Mathes 1“ schnell ersichtlich ist.

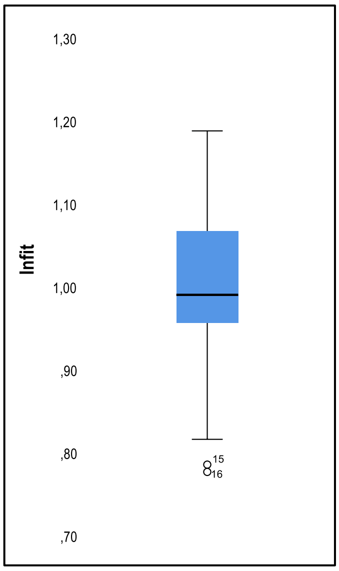

Abbildung 3: Verteilung der Itemfit-Statistiken des „Mathes 1“

Es zeigt sich, dass alle Werte innerhalb des zulässigen Wertebereiches liegen, der Großteil der Items weist einen Infit in der Nähe des Erwartungswertes von 1 auf. Lediglich die beiden letzten Items in der sechsten Aufgabengruppe (Zerlegung der 17 in 4 Teile) weichen etwas stärker vom Erwartungswert ab.
Die Passung der Items auf das (eindimensionale) Rasch-Modell ist ein Beleg dafür, dass die Aufgaben des „Mathes 1“ dasselbe Merkmal messen, sie also eine gemeinsame Skala bilden (Bühner, 2011).
Eine Auflistung der im Rasch-Modell geschätzten Fitstatistiken zu jedem Item kann im Anhang B auf S. 23 eingesehen werden.

4.3 Reliabilität

Zur Schätzung der Zuverlässigkeit des Verfahrens im Sinne der Internen Konsistenz wurde Cronbachs α (Cronbach, 1951) zu beiden Messzeitpunkten ermittelt (Rost, 2004). Die Alphakoeffizienten liegen mit α = .94 (N = 1207) zum ersten Messzeitpunkt und α = .95 (N = 1218) zum zweiten Messzeitpunkt im hohen Bereich.
Zudem wurde die Retest-Reliabilität als Maß für die Stabilität der Messungen berechnet (Schermelleh-Engel & Werner, 2012). Dazu wurden die Testergebnisse der Schülerinnen und Schüler, welche „Mathes 1“ zu beiden Messzeitpunkten bearbeiteten (s. Kapitel 4), miteinander korreliert (Rasch, Friese, Hofmann & Naumann, 2014). Die Korrelation wurde einem Signifikanztest unterzogen. Gemäß der gängigen Konvention sollten die Zusammenhänge mindestens auf dem Niveau α = .05 signifikant sein (ebd.). Wenn diese Bedingung erfüllt war, wurde die Höhe der Korrelation anhand der Klassifikation von Cohen (1988) eingeschätzt.
Der Zusammenhang zwischen den Testleistungen zur Mitte sowie zum Ende der zweiten Klassenstufe fällt signifikant aus und liegt mit r_{tt} = .74** (N = 911) im hohen Bereich. Die Retest-Reliabilität ist als geringer zu bewerten (Rost, 2013) als die Interne Konsistenz, was einen, wenn auch unspezifischen, Hinweis auf die Änderungssensitivität des „Mathes 1“ liefert. Ein änderungssensibler Test sollte nach Klauer (2011, 2014) nur mäßige Retest-Reliabilitäten aufweisen und die Zusammenhänge zwischen den Messungen sollten abnehmen, je mehr Zeit dazwischen liegt.

4.4 Validität

4.4.1 Konstruktvalidität

Zur Einschätzung der Konstruktvalidität wurden die Ergebnisse der Kinder im „Mathes 1“ zum zweiten Messzeitpunkt mit konstruktähnlichen und konstruktfernen Verfahren korreliert. Das Analysevorgehen ist analog dem im vorherigen Abschnitt 4.3 beschriebenen.
Die Zusammenhänge der Testergebnisse im „Mathes 1“ mit den mathematischen, den kognitiven sowie den Leseleistungen der Schülerinnen und Schüler werden in der Tabelle 3 dargestellt.

Tabelle 3: Zusammenhänge des „Mathes 1“ zu konstruktähnlichen und konstruktfernen Testverfahren

	Korrelation r mit		
	DEMAT 1+	KFT 1-2 R	WLLP
Mathes 1	.82** N = 1082	.56** N = 1155	.46** N = 1156

Anmerkungen: DEMAT 1+ – Deutscher Mathematiktest für erste Klassen (Krajewski et al., 2002; KFT 1-2 R – Kognitiver Fähigkeitstest für 1. und 2. Klassen (Kawthar & Perleth, 2005); WLLP – Würzburger Leise Leseprobe (Küspert & Schneider, 1998); *N* – Stichprobenumfang; *r* – Korrelationskoeffizient nach Pearson; ** – Korrelation ist auf dem Niveau von 0.01 (2-seitig) signifikant

Die Korrelationen fallen durchweg signifikant aus. Der Zusammenhang zwischen „Mathes 1“ und DEMAT 1+ (Krajewski et al., 2002) ist als hoch zu klassifizieren, ebenso der zum KFT 1-2 R (Kawthar & Perleth, 2005). Der Korrelationskoeffizient zur WLLP liegt nach Cohen (1988) im mittleren Bereich. Somit bleibt festzuhalten, dass die Zusammenhänge zu den jeweils konstruktfernen Verfahren geringer sind als zum konstruktähnlichen Verfahren, was für die Konstruktvalidität des „Mathes 1“ spricht.

4.4.2 Prognostische Validität

Vor dem Hintergrund der im Abschnitt 1.1 ausgewiesenen Zielstellungen des Verfahrens (Identifikation von Risikoschülerinnen und -schülern) ist die Vorhersagegüte des „Mathes 1“ von besonderer Bedeutung. Zur Prüfung der prognostischen Validität wurde in einem ersten Schritt der Zusammenhang der Ergebnisse im „Mathes 1“ zum Halbjahr der zweiten Klassenstufe mit denen im DEMAT 1+ (Krajewski et al., 2002) am Ende der des Schuljahres ermittelt. Das Analysevorgehen ist analog dem im Abschnitt 4.3 beschriebenen. Der Zusammenhang zwischen den Testleistungen fällt signifikant aus und liegt mit r = .73** (N = 872) im hohen Bereich, was für eine große prognostische Aussagekraft des „Mathes 1“ spricht.

Als Ergänzung zu den Ergebnissen der korrelativen Analysen wurde die klassifikatorische Güte des „Mathes 1“ ermittelt (Marx & Lenhard, 2010). Zur Berechnung der Kennwerte wurde Kindern, welche im „Mathes 1“ zur Mitte der ersten Klasse eine unterdurchschnittliche Leistung erzielten (Prozentrang < 25), ein Risikostatus in ihrer mathematischen Entwicklung zugeordnet. Dieser relativ milde Cut-Off-Wert wurde vor dem Hintergrund der Ergebnisse jüngerer Schulleistungsstudien gewählt, auf deren Grundlage davon auszugehen ist, dass etwa ein Fünftel bis ein Viertel aller Grundschülerinnen und -schüler Lernrückstände im Umfang von mehreren Schuljahren aufweist (z. B. TIMSS 2015: 23.3 %; Selter et al., 2016; s. Abschnit 1.2). Da ein wesentliches Ziel des Verfahrens ist, Kinder mit Schwierigkeiten beim Mathematiklernen für eine weiterführende Diagnostik und präventiv ausgerichtete Förderung zu erfassen, sollen möglichst keine falsch-negativen Klassifizierungen erfolgen. Als Kriterium diente ein Risikostatus im DEMAT 1+ (Krajewski et al., 2002), operationalisiert als ein Prozentrang < 16 am Ende der ersten Klassenstufe. Die Tabelle 4 stellt die Befunde der klassifikatorischen Analysen dar.

Tabelle 4: Ergebnisse zur klassifikatorischen Güte des "Mathes 1" bezüglich Risiken im Fach Mathematik

	Stichproben-umfang	*n* richtig positiv	*n* falsch positiv	*n* falsch negativ	*n* richtig negativ	Sensitivität	Spezifität	α - Fehlerquote	β - Fehlerquote	positiver prä-diktiver Wert	negativer prä-diktiver Wert	RATZ-Index
DEMAT 1+	870	96	103	28	643	0.77	0.86	0.14	0.23	0.48	0.96	0.71

Anmerkungen: n – Anzahl; Deutscher Mathematiktest für zweite Klassen (Krajewski et al., 2002)

Die Ergebnisse zur Sensitivität zeigen, dass 77 % der Kinder mit schwachen Mathematikleistungen am Ende der ersten Klassenstufe durch „Mathes 1“ bereits zum Halbjahr korrekt identifiziert werden. Die Spezifität gibt an, dass 86 % der Schülerinnen und Schüler mit zufriedenstellenden Mathematikleistungen im DEMAT 1+ (Krajewski et al., 2002) kein Risikostatus zugewiesen wird. Besonders erwähnenswert ist zudem der RATZ-Index, welcher mit 71 % im sehr hohen Bereich liegt und auf eine sehr gute Klassifikationsleistung des „Mathes 1“ hinweist (Marx & Lenhard, 2000).

5 Literaturverzeichnis

Blumenthal, Y., Hartke, B. & Koch, K. (2010). Mecklenburger Längsschnittstudie: Wie effektiv sind Diagnoseförderklassen*? Zeitschrift für Heilpädagogik, 61*, 331-341.

Bond, T. G. & Fox, C. M. (2015). *Applying the Rasch model: Fundamental measurement in the human sciences* (3rd ed.). New York & London: Routledge.

Bos, W., Bonsen, M., Baumert, J., Prenzel, M., Selter, C. & Walther, G. (Hrsg.) (2008). *TIMSS 2007. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich*. Münster: Waxmann.

Bos, W., Lankes, E.M., Prenzel, M., Schwippert, K., Walther, G. & Valtin, R. (Hrsg.) (2003). *Erste Ergebnisse aus IGLU. Schülerleistungen am Ende der vierten Jahrgangsstufe im internationalen Vergleich*. Münster: Waxmann.

Bos, W., Wendt, H., Köller, O. & Selter, C. (Hrsg.) (2012). *TIMSS 2011. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich*. Münster: Waxmann.

Bühner, M. (2011). *Einführung in die Test- und Fragebogenkonstruktion. 3., aktualisierte und erweiterte Auflage*. München: Pearson Studium.

Carstensen, C.H. & Taskinen, P. (2007). Feldtest. In M. Prenzel & W. Blum (Hrsg.), *Entwicklung eines Testverfahrens zur Überprüfung der Bildungsstandards in Mathematik für den Mittleren Schulabschluss. Technischer Bericht* (S. 16-23). Kiel: IPN.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale: Erlbaum.

Cronbach, L.J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*, 297-334.

Dilling H., Mombour, W. & Schmidt, M.H. (2011). *Internationale Klassifikation psychischer Störungen. ICD-10 Kapitel 5 (F). Klinisch-diagnostische Leitlinien*. Bern: Huber.

Flexer, R.J. (1986). The Power of Five: The Step Before the Power of Ten. *The Arithmetic Teacher, 34* (3), 5-9.

Fritz, A., Ricken, G. & Gerlach, M. (2007). *Kalkulie. Handreichung zur Durchführung der Diagnose*. Berlin: Cornelsen.

Fromme, M., Wartha, S. & Benz, C. (2011). Tragfähiges Operationsverständnis durch flexible Übersetzungen - Grundvorstellungen zur Subtraktion. *Grundschulmagazin, 4,* 35-40.

Gaidoschik, M. (2010). *Die Entwicklung von Lösungsstrategien zu den additiven Grundaufgaben im Laufe des ersten Schuljahres*. Dissertation. Universität Wien. Abruf am 28.4.20. Online verfügbar unter: http://othes.univie.ac.at/9155/1/2010-01-18_8302038.pdf

Gaidoschik, M. (2011). *Rechenschwäche – Dyskalkulie. Eine unterrichtspraktische Einführung für LehrerInnen und Eltern. 6. Auflage*. Hamburg: Persen.

Hartke, B. & Diehl, K. (2013). *Schulische Prävention im Bereich Lernen. Problemlösungen mit dem RTI-Ansatz*. Stuttgart: Kohlhammer.

Häsel-Weide, U., Nührenbörger, M., Moser Opitz, E. & Wittich, C. (2019). *Ablösung vom zählenden Rechnen. Fördereinheiten für heterogene Lerngruppen. 5. Auflage*. Seelze: Klett-Kallmeyer.

Hasemann, K. & Gasteiger, H. (2014). *Anfangsunterricht Mathematik. 3., überarbeitete und erweiterte Auflage*. Berlin: Springer.

Hasselhorn, M., Marx, H. & Schneider, W. (Hrsg.) (2005). *Diagnostik von Mathematikleistungen. Tests und Trends N.F. Bd. 4.* Göttingen: Hogrefe.

Heine, J.H. (2020). *pairwise: Rasch Model Parameters by Pairwise Algorithm. R package version 0.4.4-7*. Abruf am 28.4.20. Online verfügbar unter: http://CRAN.R-project.org/package=pairwise

Jacobs, C. & Petermann, F. (2012). *Diagnostik von Rechenstörungen. 2., überarbeitete und erweiterte Auflage*. Göttingen: Hogrefe.

Käpnick, F. (2014). *Mathematiklernen in der Grundschule*. Berlin: Springer.

Kawthar, K., A. & Perleth, C. (2005). *Kognitiver Fähigkeitstest für erste und zweite Klassen (KFT 1-2)* - Versuchsversion. Rostock: Universität Rostock.

Klauer, K.J. (2011). Lernverlaufsdiagnostik – Konzept, Schwierigkeiten und Möglichkeiten. *Empirische Sonderpädagogik, 3*, 207-224.

Klauer, K.J. (2014). Formative Leistungsdiagnostik: Historischer Hintergrund und Weiterentwicklung zur Lernverlaufsdiagnostik. In M. Hasselhorn, W. Schneider & U. Trautwein (Hrsg.), Lernverlaufsdiagnostik. *Tests und Trends, N.F. Bd. 12* (S. 1-18). Göttingen: Hogrefe.

Knopp, E. (2010). *Theoretische Grundlagen, Konzeption und Güte des Inventars „Rechenfische". Ein Verfahren zur Dokumentation von Fortschritten beim Erlernen arithmetischer Kenntnisse im Anfangsunterricht Mathematik*. München: Dr. Hut.

Koch, K. (2005). Probleme im Bereich des mathematischen Lernens. In S. Ellinger & M.C. Wittrock (Hrsg.), *Sonderpädagogik in der Regelschule. Konzepte, Forschung, Praxis* (S. 279-298). Stuttgart: Kohlhammer.

Koch, K., Hartke, B. & Blumenthal, Y. (2009). *Merkmale von Kindern mit besonderem Förderbedarf im ersten Schuljahr. Erste Ergebnisse der Mecklenburger Längsschnittstudie.* Hamburg: Dr. Kovač.

Koch, K. & Knopp, E. (2010). Mathematisches Lernen. In B. Hartke, K. Koch & K. Diehl (Hrsg.), *Förderung in der schulischen Eingangsstufe* (S. 91-118). Stuttgart: Kohlhammer.

Krajewski, K. (2003). *Vorhersage von Rechenschwäche in der Grundschule*. Hamburg: Dr. Kovač.

Krajewski, K. (2013). Wie bekommen die Zahlen einen Sinn: ein entwicklungspsychologisches Modell der zunehmenden Verknüpfung von Zahlen und Größen. In M. von Aster & J.H. Lorenz (Hrsg.), *Rechenstörungen bei Kindern: Neurowissenschaft, Psychologie, Pädagogik. 2., überarbeitete und erweiterte Auflage* (S. 156-179). Göttingen: Vandenhoeck & Ruprecht.

Krajewski, K., Küspert, P. & Schneider, W. (2002). *Deutscher Mathematiktest für erste Klassen (DEMAT 1+)*. Göttingen: Hogrefe.

Krauthausen, G. (2018): *Einführung in die Mathematikdidaktik – Grundschule. 4. Auflage*. Berlin: Springer.

Kultusministerkonferenz (2005). *Bildungsstandards im Fach Mathematik für den Primarbereich*. München: Luchterhand.

Küspert, P. & Schneider, W. (1998). *Würzburger Leise Leseprobe (WLLP)*. Göttingen: Hogrefe.

Lambert, K. (2015). *Rechenschwäche. Grundlagen, Diagnostik und Förderung*. Göttingen: Hogrefe.

Landerl, K. & Kaufmann, L. (2008). *Dyskalkulie. Modelle, Diagnostik, Intervention*. München: Reinhardt.

Linacre, J.M. (2002). What do Infit and Outfit, Mean-square and Standardized mean? *Rasch Measurement Transactions, 16* (2), 878. Abruf am 28.4.20. Verfügbar unter: http://www.rasch.org/rmt/rmt162f.htm

Lorenz, J.H. (2005). *Lernschwache Rechner fördern. 2. Auflage*. Berlin: Cornelsen.

Lorenz, J.H. (2014). Rechenschwäche. In G.W. Lauth, M. Grünke & J.C. Brunstein (Hrsg.), *Interventionen bei Lernstörungen. 2., überarbeitete und erweiterte Auflage* (S. 43-55). Göttingen: Hogrefe.

Marx, H., Jansen, H. & Skowronek, H. (2000). Prognostische, differentielle und konkurrente Validität des Bielefelder Screenings zur Früherkennung von Lese-Rechtschreibschwierigkeiten (BISC). In M. Hasselhorn, W. Schneider & H. Marx (Hrsg.), Diagnostik von Lese-Rechtschreibschwierigkeiten. *Tests und Trends, N.F. Bd. 1* (S. 9-34). Göttingen: Hogrefe.

Marx, H. & Lenhard, W. (2010). Diagnostische Merkmale von Screeningverfahren. In M. Hasselhorn & W. Schneider (Hrsg.), Frühprognose schulischer Kompetenzen. *Tests und Trends, N.F. Bd. 9* (S. 68-84). Göttingen: Hogrefe.

Moser Opitz, E. (2004). Dyskalkulie: Krankheit, Erfindung, Mythos, Etikett … ? Auseinandersetzung mit einem geläufigen, aber ungeklärten Begriff. *Vierteljahreszeitschrift für Heilpädagogik und ihre Nachbargebiete, 72*, 179-190.

Padberg, F. & Benz, C. (2011). *Didaktik der Arithmetik für Lehrerbildung und Lehrerfortbildung. 4. erweiterte, stark überarbeitete Auflage*. Berlin: Springer.

Prenzel, M. & Blum, W. (Hrsg.) (2007). *Entwicklung eines Testverfahrens zur Überprüfung der Bildungsstandards in Mathematik für den Mittleren Schulabschluss. Technischer Bericht.* Kiel: IPN.

R Core Team (2013). *R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing*. Abruf am 28.4.20 Online verfügbar unter: http://www.R-project.org/

Rasch, B., Friese, M., Hofmann, W. & Naumann, E. (2014). *Quantitative Methoden 1. Einführung in die Statistik für Psychologen und Sozialwissenschaftler. 4., überarbeitete Auflage*. Berlin: Springer.

Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Copenhagen: Nielsen & Lydiche.

Rost, J. (2004). *Lehrbuch Testtheorie – Testkonstruktion. 2. vollständig überarbeitete und erweiterte Auflage.* Bern: Huber.

Rost, D.H. (2013). *Interpretation und Bewertung pädagogisch-psychologischer Studien. 3., vollständig überarbeitete und erweiterte Auflage*. Bad Heilbrunn: Klinkhardt.

Schermelleh-Engel, K. & Werner, C.S. (2012). Methoden der Reliabilitätsbestimmung. In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion. 2., aktualisierte und überarbeitete Auflage* (S. 119-142). Berlin: Springer.

Selter, C., Walter, D., Walther, G. & Wendt, H. (2016). Mathematische Kompetenzen im internationalen Vergleich: Testkonzeption und Ergebnisse. In H. Wendt, W. Bos, C. Selter, O. Köller, K. Schwippert & D. Kasper (Hrsg.), *TIMSS 2015. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 79-136). Münster: Waxmann.

Sikora, S. (2017). *Lernverlaufsdiagnostik im Mathematikunterricht. Theoretische Grundlagen, Konzeption und Güte eines formativen Schulleistungstests für dritte Klassen*. Hamburg: Dr. Kovač.

Sikora, S. & Voß, S. (2018). *Mathematikunterricht in der inklusiven Grundschule*. Stuttgart: Kohlhammer.

Von Aster, M., Schweiter, M. & Weinhold Zulauf, M. (2007). Rechenstörungen bei Kindern. Vorläufer, Prävalenz und psychische Symptome. *Zeitschrift für Pädagogische Psychologie*, *39*, 85-96.

Wartha, S. & Schulz, A. (2013). *Rechenproblemen vorbeugen (2. Auflage)*. Berlin: Cornelsen

Winter, H. (2001). *Inhalte mathematischen Lernens. Expertise für das Ministerium für Bildung, Wissenschaft, Jugend und Kultur des Landes Rheinland-Pfalz*. Abruf am 4.5.20. Online verfügbar unter: http://grundschule.bildung-rp.de/fileadmin/user_upload/grundschule.bildung-rp.de/Downloads/Mathemathik/Winter_Inhalte_math_Lernens.pdf

Wittmann, E.Ch. & Müller, G.N. (2012). *Das Zahlenbuch 1 (Neubearbeitung). Begleitband*. Stuttgart: Klett.

6 Anhang

Anhang A: Schwierigkeiten und Trennschärfen der Items

Item	Schwierigkeit P_i zum Messzeitpunkt		Trennschärfe r_{pbis} zum Messzeitpunkt	
	1	2	1	2
1.1	0.83	0.93	0.46	0.32
1.2	0.76	0.91	0.46	0.40
2.1	0.80	0.91	0.36	0.33
3.1	0.79	0.91	0.41	0.44
3.2	0.69	0.89	0.52	0.55
3.3	0.36	0.69	0.54	0.57
3.4	0.16	0.46	0.46	0.53
4.1	0.44	0.75	0.49	0.56
4.2	0.40	0.72	0.50	0.55
5.1	0.88	0.95	0.38	0.32
5.2	0.86	0.95	0.39	0.37
5.3	0.71	0.89	0.39	0.41
6.1	0.71	0.88	0.55	0.50
6.2	0.67	0.87	0.60	0.61
6.3	0.38	0.80	0.65	0.67
6.4	0.36	0.78	0.62	0.68
7.1	0.43	0.72	0.48	0.49
7.2	0.59	0.83	0.54	0.58
7.3	0.38	0.65	0.45	0.49
8.1	0.87	0.94	0.48	0.42
8.2	0.84	0.95	0.48	0.38
8.3	0.51	0.82	0.60	0.57
8.4	0.57	0.85	0.60	0.60
8.5	0.78	0.90	0.53	0.49
8.6	0.77	0.90	0.53	0.45
8.7	0.64	0.88	0.57	0.53
9.1	0.78	0.87	0.51	0.49
9.2	0.72	0.82	0.52	0.52
9.3	0.73	0.85	0.47	0.51
9.4	0.81	0.90	0.48	0.42
9.5	0.36	0.77	0.54	0.55
9.6	0.29	0.69	0.53	0.57
9.7	0.29	0.72	0.54	0.62
10.1	0.43	0.72	0.54	0.67
10.2	0.26	0.67	0.55	0.67
10.3	0.67	0.85	0.57	0.56
10.4	0.24	0.65	0.46	0.54
11.1	0.37	0.65	0.51	0.59
11.2	0.65	0.83	0.52	0.55
11.3	0.64	0.84	0.62	0.64
11.4	0.20	0.60	0.52	0.58
12.1	0.24	0.75	0.38	0.58
12.2	0.16	0.61	0.37	0.54
12.3	0.06	0.34	0.28	0.40

Anhang B: Itemparameter der Rasch-Analyse

Item	Schwierigkeit	Infit	Item	Schwierigkeit	Infit
1.1	-1.69	1.07	8.4	-0.22	0.92
1.2	-1.27	1.10	8.5	-1.30	0.97
2.1	-1.32	1.17	8.6	-1.26	0.98
3.1	-1.41	1.17	8.7	-0.57	0.99
3.2	-0.88	1.02	9.1	-1.07	1.06
3.3	1.09	0.98	9.2	-0.60	1.07
3.4	2.70	0.96	9.3	-0.77	1.08
4.1	0.62	1.08	9.4	-1.40	1.02
4.2	0.89	1.07	9.5	0.87	0.97
5.1	-2.19	1.00	9.6	1.36	0.96
5.2	-2.02	1.01	9.7	1.25	0.89
5.3	-0.83	1.18	10.1	0.77	0.96
6.1	-0.85	1.00	10.2	1.59	0.82
6.2	-0.76	0.93	10.3	-0.66	0.98
6.3	0.69	0.78	10.4	1.60	0.98
6.4	0.82	0.79	11.1	1.21	1.02
7.1	0.75	1.15	11.2	-0.38	1.09
7.2	-0.22	1.01	11.3	-0.47	0.90
7.3	1.11	1.19	11.4	2.06	0.89
8.1	-2.09	0.92	12.1	1.28	1.01
8.2	-1.99	0.99	12.2	2.01	0.97
8.3	0.14	0.95	12.3	3.38	0.96

Anhang C: Normtabellen des „Mathes 1“

	Mitte Klasse 1 (20. / 21. Schulwoche) N = 1207		Anfang Klasse 2 (3. / 4. Schulwoche) N = 1218	
Rohwert	T-Wert	Prozentrang	T-Wert	Prozentrang
0	26	1	12	0
1	27	2	13	0
2	28	3	14	0
3	29	4	15	0
4	30	4	16	1
5	31	5	17	1
6	32	6	18	2
7	33	7	20	2
8	34	8	21	3
9	35	9	22	3
10	36	10	23	3
11	37	12	24	4
12	38	14	25	4
13	39	16	26	5
14	40	17	27	5
15	41	19	28	6
16	42	23	29	6
17	43	25	30	7
18	44	28	32	8
19	45	32	33	9
20	46	35	34	9
21	47	39	35	10
22	48	44	36	11
23	49	47	37	12
24	50	49	38	14
25	51	52	39	14
26	52	56	40	16
27	53	60	41	18
28	54	63	42	20
29	55	66	43	22
30	56	69	45	24
31	56	71	46	26
32	57	75	47	28
33	58	78	48	31
34	59	80	49	34
35	60	84	50	38
36	61	87	51	41
37	62	90	52	45
38	63	93	53	51
39	64	94	54	57
40	65	96	55	65
41	66	98	57	74
42	67	99	58	82
43	68	100	59	93
44	69	100	60	100