

Mathes 4

Formative Erfassung der mathematischen
Kompetenzen von Viertklässlern

Manual



Dieses Werk ist lizenziert unter einer [Creative Commons Namensnennung - Nicht-kommerziell - Weitergabe unter gleichen Bedingungen 4.0 International Lizenz](https://creativecommons.org/licenses/by-nc-sa/4.0/).



Inhalt

Abbildungsverzeichnis	3
Tabellenverzeichnis	3
Kurzinformation	4
1. Theoretische Grundlagen	5
1.1 Zielstellung der Mathes-Testreihe.....	5
1.2 Exkurs: Klassifikation und Epidemiologie von mathematischen Lernschwierigkeiten - Was soll verhindert werden?.....	5
2 Theoriebasierte Testentwicklung	6
2.1 Was wird von Schülerinnen und Schüler am Ende der Grundschulzeit erwartet?.....	6
2.2 Konstruktion inhaltsvalider Testanforderungen	7
2.3 Itempilottierung und -auswahl.....	9
2.4 Überlegungen zur Steigerung der Testökonomie.....	11
3 Testanwendung	11
3.1 Anwendungszeitraum und Zielgruppe	11
3.2 Testmaterial	12
3.3 Hinweise zur Testdurchführung	12
3.4 Auswertung und Interpretation.....	13
3.4.1 Manuelle Auswertung.....	13
3.4.2 Automatisierte Auswertung.....	13
3.4.3 Interpretation der Ergebnisse	13
4 Testgütekriterien	14
4.1 Itemkennwerte	15
4.2 Reliabilität.....	17
4.3 Gültigkeit des Messmodells.....	17
4.4 Validität.....	18
5 Literaturverzeichnis	19
6 Anhang	22

Abbildungsverzeichnis

Abbildung 1: Allgemeine und inhaltsbezogene mathematische Kompetenzen im Mathematikunterricht der Grundschule (nach KMK, 2005, S. 7)	7
Abbildung 2: Referenzniveaus als Interpretationshilfen für die erzielte Testleistung	14
Abbildung 3: Verteilungen von Aufgabenschwierigkeiten und Kompetenzwerten zu den zwei Messzeitpunkten..	16
Abbildung 4: Verteilung der Itemfit-Statistiken der vier Testskalen des „Mathes 4“	18

Tabellenverzeichnis

Tabelle 1: inhaltliche Anforderungen des "Mathes 4"	10
Tabelle 2: Durchführungszeiträume des "Mathes 4"	12
Tabelle 3: Ergebnisse der Analysen hinsichtlich der WLE-Reliabilität der Skalen des "Mathes 4"	17
Tabelle 4: Ergebnisse der Analysen hinsichtlich der Internen Konsistenz (Cronbachs α) der Subtests des "Mathes 4"	17

Kurzinformation

Titel, Autoren, Jahr	Mathes 4, Simon Sikora, 2020
Diagnostische Zielsetzung	Erfassung der Entwicklung der mathematischen Kompetenzen im Verlauf der vierten Klassenstufe bei allen Kindern
Anwendungsbereiche	Grundschulmathematikunterricht in der vierten Klasse
Aufbau	vier inhaltlich parallele Testformen A, B, C und D mit jeweils etwa 75 Aufgaben, verteilt auf vier Testskalen (Zahlen und Operationen, Größen und Messen, Daten, Häufigkeit und Wahrscheinlichkeit, Raum und Form)
Anwendungszeitraum	Mitte Klasse 4 (20. / 21. Schulwoche) Ende Klasse 4 (40. / 41. Schulwoche)
Durchführung	Gruppentestung im Klassenkontext durch die Lehrkraft bei freier Testbearbeitung innerhalb einer Unterrichtsstunde
Auswertung	skalenweise Auswertung mithilfe von Auswertungsvorlagen
Normen	Quotenstichprobe Rostock, $N = 945$ aus 39 Klassen von 13 Schulen
Reliabilität	Interne Konsistenz (WLE-Reliabilität) der testheftübergreifenden Skalen (N zum MZP 1 = 780, N zum MZP 2 = 691): Zahlen und Operationen: $r_{WLE} = .86$ und $r_{WLE} = .85$ Größen und Messen: jeweils $r_{WLE} = .81$ Daten, Häufigkeit und Wahrscheinlichkeit: $r_{WLE} = .83$ und $r_{WLE} = .82$ Raum und Form: $r_{WLE} = .74$ und $r_{WLE} = .72$
Validität	-- Angaben werden im Rahmen einer neuen Auflage ergänzt --

1. Theoretische Grundlagen

1.1 Zielstellung der Mathes-Testreihe

Taglich stellen sich Lehrkrafte Fragen wie die folgenden:

„Lernen alle Kinder in meinem Unterricht erfolgreich oder kommt jemand nicht mit?“
„Kann ich das aktuelle Themengebiet abschlieen und im Stoff weitermachen oder brauchen die Schulerinnen und Schuler noch mehr Lernzeit?“
„Was genau hat die Schulerin bzw. der Schuler noch nicht verstanden?“

Solche Fragen haben das Ziel, den Unterricht bestmoglich an die individuellen Lernausgangslagen und Forderbedurfnisse der Kinder anzupassen. Die Mathes-Testreihe soll Lehrkrafte bei vertretbarem Aufwand dabei unterstutzen, zu einer prazisen Einschatzung der aktuellen Lernstande sowie der Leistungsentwicklung der Schulerinnen und Schuler zu gelangen. Die Testverfahren konnen in dreierlei Hinsicht helfen, namlich

- beim genauen Einschatzen des Spektrums der Leistungen der Schulerinnen und Schuler, um den Unterricht daran bestmoglich anpassen zu konnen,
- beim rechtzeitigen Erkennen derjenigen Schulerinnen und Schuler mit Risiken bzw. bereits ausgepragten Schwierigkeiten im Kompetenzerwerb sowie
- bei der Planung effektiver Fordermanahmen, insbesondere fur diejenigen Schulerinnen und Schuler mit besonderen Unterstutzungsbedarfen.

Zu diesem Zweck wurden Mathes-Tests fur jede Klassenstufe entwickelt, welche jeweils zum Beginn und in der Mitte des Schuljahres eingesetzt werden konnen.

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

Schwierigkeiten beim Rechnen- bzw. mathematischen Lernen gehoren fur viele Schulerinnen und Schuler (und deren Lehrerinnen und Lehrer) zum schulischen Alltag. Allerdings erhalt nicht jedes Kind mit schwachen Leistungen in Mathematik die Diagnose „Dyskalkulie“ bzw. „Rechenstorung“. Nach den Kriterien der Weltgesundheitsorganisation leidet ein Kind nur dann unter einer Dyskalkulie, wenn seine Beeintrachtigung 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

2.1 Was wird von Schülerinnen und Schüler am Ende der Grundschulzeit erwartet?

Im Zusammenhang mit der Fragestellung, was Schülerinnen und Schüler bis zum Ende der Grundschulzeit lernen sollen, trat der Begriff der Kompetenz in den Fokus pädagogischer Überlegungen (Klieme et al., 2007; Blum, 2010). Die Kompetenz eines Individuums umfasst nach Weinert (2001) zusammenwirkende Aspekte wie Wissen, Fähigkeiten, Verständnis, Können, Handeln, Erfahrungen und Motivation, welche eine Person befähigen, konkrete

Anforderungen zu bewältigen. Dieses Verständnis spiegelt sich in den Rahmenplänen aller deutschen Bundesländer wider, welche sich an den länderübergreifenden Bildungsstandards (KMK, 2005) orientieren. Die Standards für den Primarbereich im Fach Mathematik beschreiben die mathematischen Kompetenzen, die Schülerinnen und Schüler am Ende der vierten Klassenstufe erreicht haben sollen, wobei zwischen inhaltsbezogenen und allgemeinen, eher prozessbezogenen, Kompetenzen unterschieden wird. Beide Kompetenzfelder sind in enger Verbindung zu sehen, wie die nachfolgende Abbildung 1 veranschaulicht.

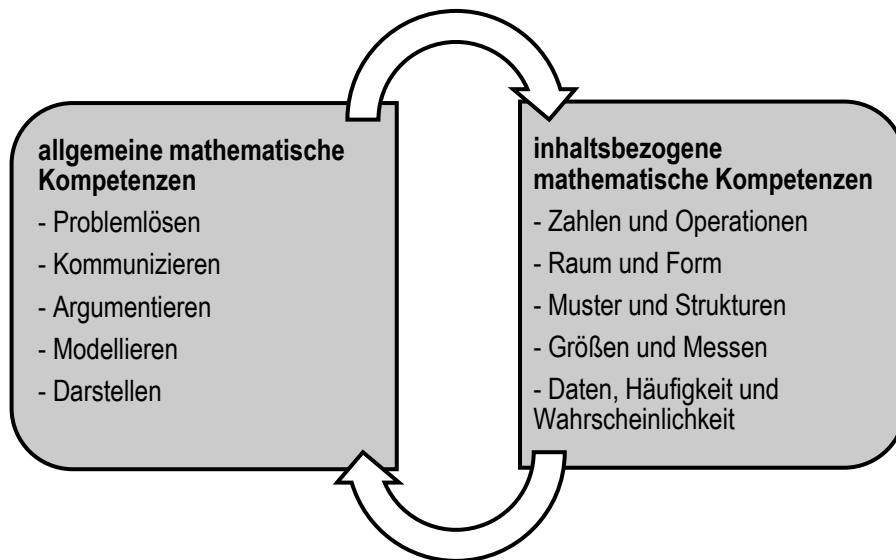


Abbildung 1: Allgemeine und inhaltsbezogene mathematische Kompetenzen im Mathematikunterricht der Grundschule (nach KMK, 2005, S. 7)

Aus den dem Kompetenzbegriff inhärenten Leistungen erwachsen neue Erwartungen und Anforderungen an den modernen Mathematikunterricht. Die Entwicklung von Kompetenzen „hängt nicht nur davon ab, welche Inhalte unterrichtet wurden, sondern in mindestens gleichem Maße davon, wie sie unterrichtet wurden“ (KMK, 2005, S. 6; Hervorhebungen im Original). Die Standards sollen einen Unterricht unterstützen, der „nicht auf die Aneignung von Kenntnissen und Fertigkeiten reduziert“ (KMK, 2005, S. 6) ist, sondern vielmehr auf „die Entwicklung eines gesicherten Verständnisses“ (ebd.; Hervorhebung im Original) abzielt.

Um die in Abschnitt 1.1 genannten Ziele für Lehrkräfte erfüllen zu können, müssen Mathematiktests am Ende der Grundschulzeit abbilden, welche inhalts- und prozessbezogenen Kompetenzen die Schülerinnen und Schüler bereits erworben haben und welche noch nicht.

2.2 Konstruktion inhaltsvalider Testanforderungen

Die Bildungsstandards der KMK (2005) beinhalten eine Zielvorgabe des Mathematikunterrichts in der Grundschule, welche ebenfalls den aktuellen Stand der mathematikdidaktischen Fachdiskussion widerspiegelt (u.a. Walther, van den Heuvel-Panhuizen et al., 2011; Grassmann et al., 2014). Kritik wurde allerdings an der Leitidee Muster und Strukturen geübt. Wittmann und Müller stellen heraus, dass es sich bei diesem Bereich „nicht einfach um einen Aspekt des Mathematikunterrichts unter anderen handelt, sondern dass dieser Aspekt grundlegend ist“ (2011,

S. 42). In Anlehnung an Devlin (1998) können Muster und Strukturen gar als das Wesen der Mathematik gesehen werden. Folglich ist die Berücksichtigung mathematischer Gesetzmäßigkeiten in den Bildungsstandards keineswegs falsch, vielmehr wird ihr Stellenwert höher eingeschätzt.

Die Bildungsstandards der KMK (2005) sind damit auch aus wissenschaftlicher Perspektive ein geeigneter Referenzpunkt für die Erarbeitung eines formativen Schulleistungstests für das Fach Mathematik in vierten Klassen. Die in den Standards formulierten Leitideen liefern einen sinnvollen inhaltlichen Rahmen für ein solches Verfahren: Indem regelmäßig Einblick in die Entwicklung in den inhaltlichen Schwerpunkten der dritten Klassenstufe genommen wird, kann der „Weg“ des Kompetenzerwerbs bezogen auf Zahlen und Operationen, Raum und Form, Größen und Messen sowie Daten, Häufigkeit und Wahrscheinlichkeit verfolgt werden. Zu jedem Messzeitpunkt kann so ein Leistungsprofil der Schülerinnen und Schüler erstellt werden, aus dem leicht ersichtlich ist, in welchem Inhaltsbereich eine gezielte Förderung am dringlichsten wäre, aber auch, sofern das Verfahren mehrfach eingesetzt wird, wie die Entwicklung im jeweiligen Bereich verläuft.

Diese Überlegungen werden durch empirische Erkenntnisse zur Struktur mathematischer Kompetenzen im späten Grundschulalter gestützt. Winkelmann, Robitzsch, Stanat und Köller (2012) untersuchten die fachspezifischen Kompetenzen anhand von zwei Datensätzen mit $N = 10\,328$ und $N = 6\,638$ Dritt- und Viertklässlern, die bei der Entwicklung von Messinstrumenten zur Überprüfung der Bildungsstandards im Fach Mathematik in der Grundschule generiert wurden. Ihre Dimensionsanalysen belegen, dass die Daten am besten durch ein Modell mit fünf inhaltsbezogenen Faktoren entsprechend den Leitideen der Bildungsstandards (KMK, 2005) abgebildet werden.

Die Bildungsstandards der KMK (2005) bilden nicht nur einen sinnvollen inhaltlichen Rahmen für einen Test zur Lernverlaufdiagnostik, sondern legen darüber hinaus implizit Anforderungen an die Aufgaben eines solchen Verfahrens fest. Die Aufgaben dürfen somit nicht nur die Wiedergabe von Kenntnissen oder ein rezeptartiges Nachmachen und Ausführen von Algorithmen erfordern, sondern sollten komplexere Probleme an die Schülerinnen und Schüler stellen, deren Lösungen der Aktivierung allgemeiner mathematischer Kompetenzen bedürfen. Dennoch darf nicht unerwähnt bleiben, dass auch technische Grundfertigkeiten wie Rechenkompetenzen oder Zahlvorstellungen in dem Verfahren enthalten sein müssen. Schließlich kann eine komplexere Aufgabenlösung auch an der Reproduktion von Wissen, Fertigkeiten und Grundvorstellungen scheitern. Ein förderrelevantes Verfahren muss demnach auch solche Basiskompetenzen umfassen, aber eben nicht ausschließlich.

Die Testinhalte wurden zudem, wenn möglich, an fundamentalen Ideen der Mathematik (zusammenfassend Vohns, 2007) ausgerichtet. Diese mathematischen Grundideen sind die zentralen, fachtypischen „Fragestellungen, Bezeichnungsweisen, Strukturierungs- und Begründungsformen“ (Führer, 1997, S. 83), welche sich wie ein roter Faden durch das mathematische Denken ziehen. Für die Grundschulmathematik existieren zum aktuellen Zeitpunkt zwei etablierte Aufstellungen fundamentaler Ideen, und zwar von Wittmann (1995, 1999) bzw. Wittmann und Müller (2012) sowie Winter (2001). Beide wurden bei der Entwicklung der Testaufgaben berücksichtigt.

Zusammenfassend ergeben sich aus den dargestellten mathematikdidaktischen Überlegungen folgende Konsequenzen für einen formativ ausgerichteten Schulleistungstest bei Schülerinnen und Schülern vierter Klassen:

- Zuordnung der Aufgaben zu Testskalen nach Leitideen der Bildungsstandards der KMK (2005)
- Integration des Kompetenzbereichs Muster und Strukturen in die anderen Testskalen
- Konstruktion von Aufgaben, deren Lösung für vorhandene inhaltsbezogene und allgemeine mathematische Kompetenzen gemäß den Bildungsstandards der KMK (2005) sprechen
- Konstruktion von Aufgaben, welche mathematische Basiskompetenzen erfassen
- Konstruktion von Aufgaben, welche die fundamentalen Ideen der Mathematik berücksichtigen und so das Verständnis der zugrundeliegenden mathematischen Konzepte und Strukturen abprüfen

2.3 Itempilottierung und -auswahl

Insgesamt wurden 309 Items konstruiert und auf vier Itempools verteilt. 105 wurden dem „Mathes 3“ (Sikora & Hartke, 2012) entnommen, da zukünftig angedacht ist, eine schuljahresübergreifende Messung der Lernentwicklung zu ermöglichen. Dafür sind allerdings weitere Studien nötig (s. Kapitel 4).

Alle Items wurden so formuliert, dass nur zwei Ausprägungen (richtig oder falsch) möglich sind. Die entwickelten Aufgaben wurden Viertklässlerinnen und Viertklässlern Grundschulen zu zwei Messzeitpunkten im Januar 2018 ($N = 780$) und Juni 2019 ($N = 691$) pilotiert. Dabei bearbeitete jedes Kind zu jedem Messzeitpunkt einen der Itempools innerhalb von einer Unterrichtsstunde. Der Datenausfall aufgrund von Nichtteilnahme an einem Messzeitpunkt, z. B. durch Krankheit, war relativ gering und unsystematisch.

Mit den so gewonnenen Daten wurden die Items einer kritischen Prüfung unterzogen. Unpassende Aufgaben wurden in Anlehnung an Prenzel und Blum (2007) in einem stufenweisen Prozess nach verschiedenen Kriterien (fehlerhafte Items, zu hoher [$< 5\%$ Lösungsrate] oder niedriger Schwierigkeitsgrad [$> 95\%$ Lösungsrate], ungenügender Modellfit [weighted Mean Square $wMNSQ < 0.7$ bzw. $wMNSQ > 1.3$], geringe Trennschärfe [$r_{pbis} < .20$]) ausgeschlossen.

Insgesamt wurden 191 der pilotierten Items in das endgültige Verfahren mit dem Titel „Mathes 4“ aufgenommen und auf vier Testformen so verteilt, dass die wesentlichen Inhalte aus jedem Kompetenzbereich abgedeckt werden. Dabei wurde ein Multi-Matrix-Design (Mislevy, Beaton, Kaplan & Sheehan, 1992) verwendet, sodass die Ergebnisse in den vier Testformen vergleichbar sind. In Übereinstimmung mit den Empfehlungen von Kolen und Brennan (2004) wurden in jeden Subtest anteilmäßig mindestens 20 % Anker-Items mit verschiedenen Schwierigkeitsgraden aufgenommen. Insgesamt kommen 97 Aufgaben in mindestens zwei der Testformen vor und fungieren somit als Linking-Items. 66 Aufgaben sind zudem in mindestens einer Testform des „Mathes 3“ (Sikora & Hartke, 2012) enthalten, um zukünftig die Kompetenzentwicklung auch schuljahresübergreifend abbilden zu können. Jede Form des „Mathes 4“ erhielt insgesamt etwa 75 Items, sodass der Test innerhalb einer Schulstunde durchführbar ist.

Durch das gewählte Testdesign mit Anker-Items resultieren folgende Vorteile für die Durchführung des „Mathes 4“ im Klassenverband:

- Bei Klassentestungen können nebeneinandersitzenden Schülerinnen und Schülern verschiedene Testformen des „Mathes 4“ dargeboten werden, sodass Abschreibeffekte ausgeschlossen werden können.
- Das auf Klassenebene gemessene Konstrukt ist somit umfangreicher (höhere inhaltliche Validität), da mehr Facetten durch die Mehrzahl von Aufgaben erfasst werden können.
- Bei wiederholten Messungen können einem Kind verschiedene Testformen des „Mathes 4“ dargeboten werden, sodass Erinnerungseffekte ausgeschlossen werden können.
- Bei wiederholten Messungen ist dementsprechend auch das auf Schülerebene gemessene Konstrukt umfangreicher (höhere inhaltliche Validität der Lernverlaufsmessungen).

Die Tabelle 1 ordnet die inhaltlichen Anforderungen des „Mathes 4“ den vier Testformen/-versionen A bis D zu.

Tabelle 1: inhaltliche Anforderungen des "Mathes 4"

	Aufbengruppe bzw. Anforderung	Testform des „Mathes 4“			
Zahlen und Operationen	Stellentafel	A			
	Zahlenstrahlen		B	C	D
	Reihen ergänzen	A		C	
	große Zahlen bilden		B		
	Zahlen ordnen				D
	Zahlvergleiche	A	B		
	Zahlen runden			C	D
	arithmetische Muster	A	B	C	D
	Zahlenrätsel (Zahlenwissen)	A	B	C	D
	abzählende Kombinatorik	A	B	C	D
	Verdoppeln / Halbieren	A	B	C	D
	Kopfrechnen	A	B	C	D
	Rechenstrategien	A	B	C	D
	halbschriftliches Rechnen	A	B	C	D
	schriftliches Rechnen	A	B	C	D
	Überschlagsrechnen	A	B	C	D
	Rechengesetze	A	B	C	D
	Größen und Messen	Schätzen	A	B	C
Längen schätzen		A	B		
Umrechnungen		A	B	C	D
Größenvergleiche		A	B	C	D
Zeitspannen ablesen		A	B	C	D
Argumentieren mit Massen		A	B	C	D
Argumentieren mit Geld		A	B	C	D
Messverständnis		A			
Modellieren		A	B	C	D
Daten, Häufigkeit und Wahrscheinlichkeit		Fahrplan	A	B	C
	Säulendiagramm	A	B	C	D
	Kreisdiagramm		B	C	
	Strichliste	A	B	C	D
	Wahrscheinlichkeitsbegriff	A	B	C	D
	Siegchancen bewerten	A	B	C	D

Raum und Form	Figur-Grund-Diskrimination	A	B	C	D
	Spiegelachsen ebener Figuren	A	B	C	D
	Begriffswissen ebener Figuren (Figuren ergänzen)	A	B	C	D
	visuelle Diskrimination im Raum	A	B	C	D
	räumliches Vorstellungsvermögen	A	B	C	D
	Begriffswissen geometrischer Körper	A	B	C	D
	geometrische Muster	A	B		D
	Abbildungsgeometrie (maßstabsgetreues Verkleinern)				D

2.4 Überlegungen zur Steigerung der Testökonomie

Zur Erreichung einer hohen Testökonomie und Zumutbarkeit der Durchführung des „Mathes 4“ wurden folgende Maßnahmen getroffen:

- freie Testbearbeitung ohne Zeitbegrenzung
- 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 (Leistungsprofil entsprechend den Inhaltsbereichen) auf Kind- und Klassenebene

3 Testanwendung

Hinweis: Alle für die Durchführung und Auswertung des „Mathes 4“ 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 (inklusive) Grundschulklasse konzipiert.

„Mathes 4“ kann zu zwei Messzeitpunkten in der vierten Klassenstufe durchgeführt werden. Die nachfolgende Tabelle 2 stellt die Durchführungszeiträume dar.

Tabelle 2: Durchführungszeiträume des "Mathes 4"

	Schulwoche	Kommentar
MZP 1	20./21.	vor dem Halbjahreszeugnis
MZP 2	40./41.	vor den Sommerferien

3.2 Testmaterial

Für die Durchführung des „Mathes 4“ werden folgende Materialien benötigt:

- 1 Testheft pro Kind,
- 1 Füller, 1 Bleistift,
- 1 Testheft für die Lehrkraft zur Demonstration,
- 1 [Durchführungsanleitung](#) für die Lehrkraft.

3.3 Hinweise zur Testdurchführung

Die Durchführung des „Mathes 4“ 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.
- Die Schülerinnen und Schüler haben insgesamt eine Unterrichtsstunde Zeit für die Aufgabenbearbeitung. Falls es Kinder gibt, die es in dieser Zeit nicht schaffen, alle Aufgaben zu bearbeiten, sollten diese in einer anderen Stunde die Möglichkeit erhalten, den Test zu beenden.
- „Mathes 4“ wird ohne Pause durchgeführt.
- 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.
- Nebeneinandersitzende Schülerinnen und Schüler sollten unterschiedliche Testformen erhalten, um Abschreiben zu verhindern.
- 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 4“ sollte möglichst nicht in der letzten Unterrichtsstunde und nicht im Nachmittagsunterricht durchgeführt werden.
- Bei mehrfachem Einsatz des „Mathes 4“ sollte jedes Kind eine andere Testform erhalten, um Erinnerungseffekte auszuschließen.

3.4 Auswertung und Interpretation

3.4.1 Manuelle Auswertung

Die [Auswertungsvorlagen](#) unterstützen eine objektive und ökonomische Auswertung des „Mathes 4“. Alle richtig gelösten Aufgaben werden mit einem Punkt, falsch gelöste mit null Punkten bewertet. Die erreichten Punkte werden für jede Testskala aufsummiert. Mithilfe der Normtabellen 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 D, S. 33ff.).

Differenzierte Informationen zur Auswertung des „Mathes 4“ 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 4“ angeboten. Bei dieser Variante müssen lediglich die erreichten Rohwerte in den vier Testskalen (*Zahlen und Operationen; Größen und Messen; Daten, Häufigkeit und Wahrscheinlichkeit; Raum und Form*) mithilfe der [Auswertungsvorlagen](#) wie in Abschnitt 3.4.1 beschrieben ermittelt und auf der Internetplattform eingetragen werden. Anschließend werden automatisch 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 4“ stellt die Internetplattform ebenfalls den Lernverlauf 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.

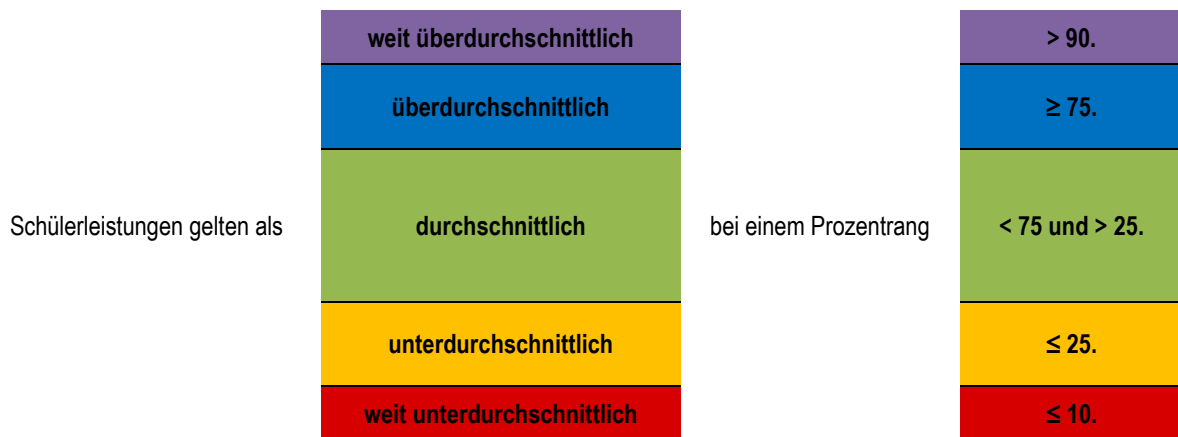


Abbildung 2: Referenzniveaus als Interpretationshilfen für die erzielte Testleistung

4 Testgütekriterien

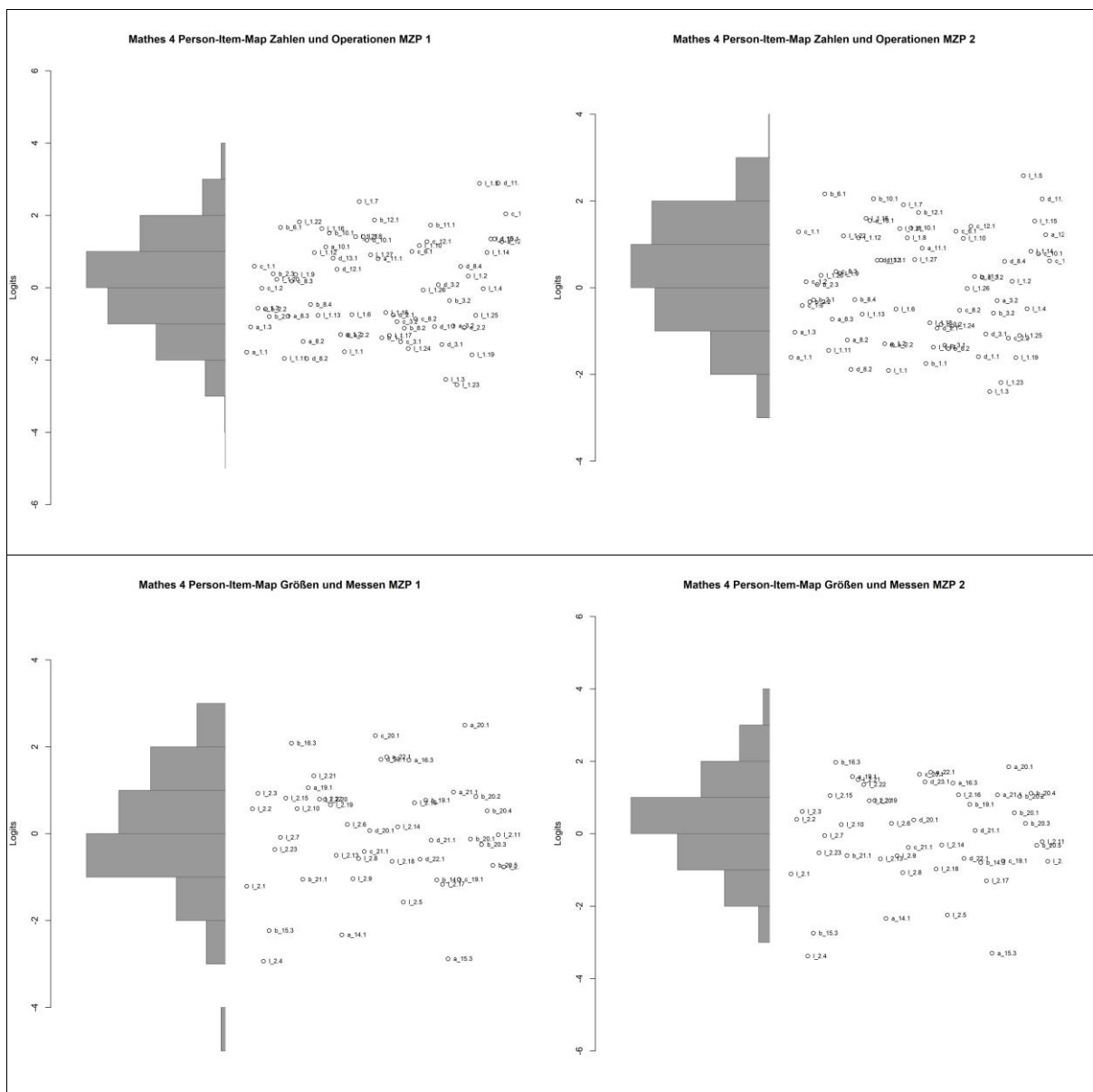
Zur Einschätzung der psychometrischen Güte des „Mathes 4“ wurde eine Studie in einer Rostocker Stichprobe im Schuljahr 2018 / 2019 durchgeführt. Die Schulleitungen aller Grundschulen Rostocks wurden um Teilnahme gebeten. 13 Schulen mit 39 Klassen konnten dafür gewonnen werden. Die Untersuchungsgruppe verteilt sich prozentual gemäß ihrem wahren Anteil auf das Rostocker Stadtgebiet und entspricht somit einer Quotenstichprobe für die Hansestadt. Eine Analyse der VERA-Leistungen (ausführlich in Sikora, 2017) erbrachte keine bedeutsamen Unterschiede zwischen Rostocker Schülerinnen und Schülern und allen anderen Drittklässlerinnen und Drittklässlern in Mecklenburg-Vorpommern. Es ist somit sehr wahrscheinlich, sodass die im Rahmen der vorliegenden Studie ermittelten Vergleichswerte auch über die Stadtgrenzen Rostocks hinaus bedeutsam sind.

Die Stichprobe umfasste 945 Schülerinnen und Schüler. Informationen zum Alter und zum Geschlecht der Kinder wurden nicht erhoben. Jedes Kind bearbeitete zu beiden Messzeitpunkten eine andere Version des Verfahrens innerhalb von einer Unterrichtsstunde, wobei die Reihenfolge der Testformen zufällig festgelegt wurde. Eine Schule nahm zum zweiten Messzeitpunkt nicht mehr an der Untersuchung teil. Ansonsten war der Datenausfall relativ gering und unsystematisch und ist vermutlich auf Krankheit einzelner Kinder zu einem der Messzeitpunkte zurückzuführen. Fehlende Werte wurden daher nicht ersetzt und die betreffenden Kinder verblieben zum übrigen Messzeitpunkt im Datensatz.

In der vorliegenden Untersuchung konnten keine weiteren Messverfahren zu Validierungszwecken durchgeführt werden. Zukünftig ist eine weitere Studie geplant, in der geprüft werden soll, ob die mathematische Kompetenzentwicklung von Schülerinnen und Schülern im Verlauf der dritten und vierten Klassenstufe mithilfe des „Mathes 3“ (Sikora & Hartke, 2012) sowie des „Mathes 4“ abgebildet werden kann (s. Abschnitt 2.3). In diesem Rahmen sollen auch Instrumente zur Validierung des „Mathes 4“ eingesetzt werden. Angaben zur verlaufdiagnostischen Eignung sowie zur Validität werden daher in einer neuen Auflage des vorliegenden Testmanuals ergänzt.

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 sowohl zur Mitte als auch am Ende des Schuljahres alle Leistungsbereiche abdecken und möglichst zu keinem Messzeitpunkt Boden- bzw. Deckeneffekte aufweisen. Ob die Schwierigkeit der Aufgaben des „Mathes 4“ für Schülerinnen und Schüler vierter Klassen angemessen ist, wurde mittels der Übereinstimmung der im Rasch-Modell (s. Abschnitt 4.3) geschätzten Item- und Personenparameter geprüft (Carstensen & Taskinen, 2007). Die nachfolgende Abbildung 3 stellt die Verteilungen der Item- und Personenparameter der vier testheftübergreifenden Skalen des „Mathes 4“ zu beiden Messzeitpunkten dar.



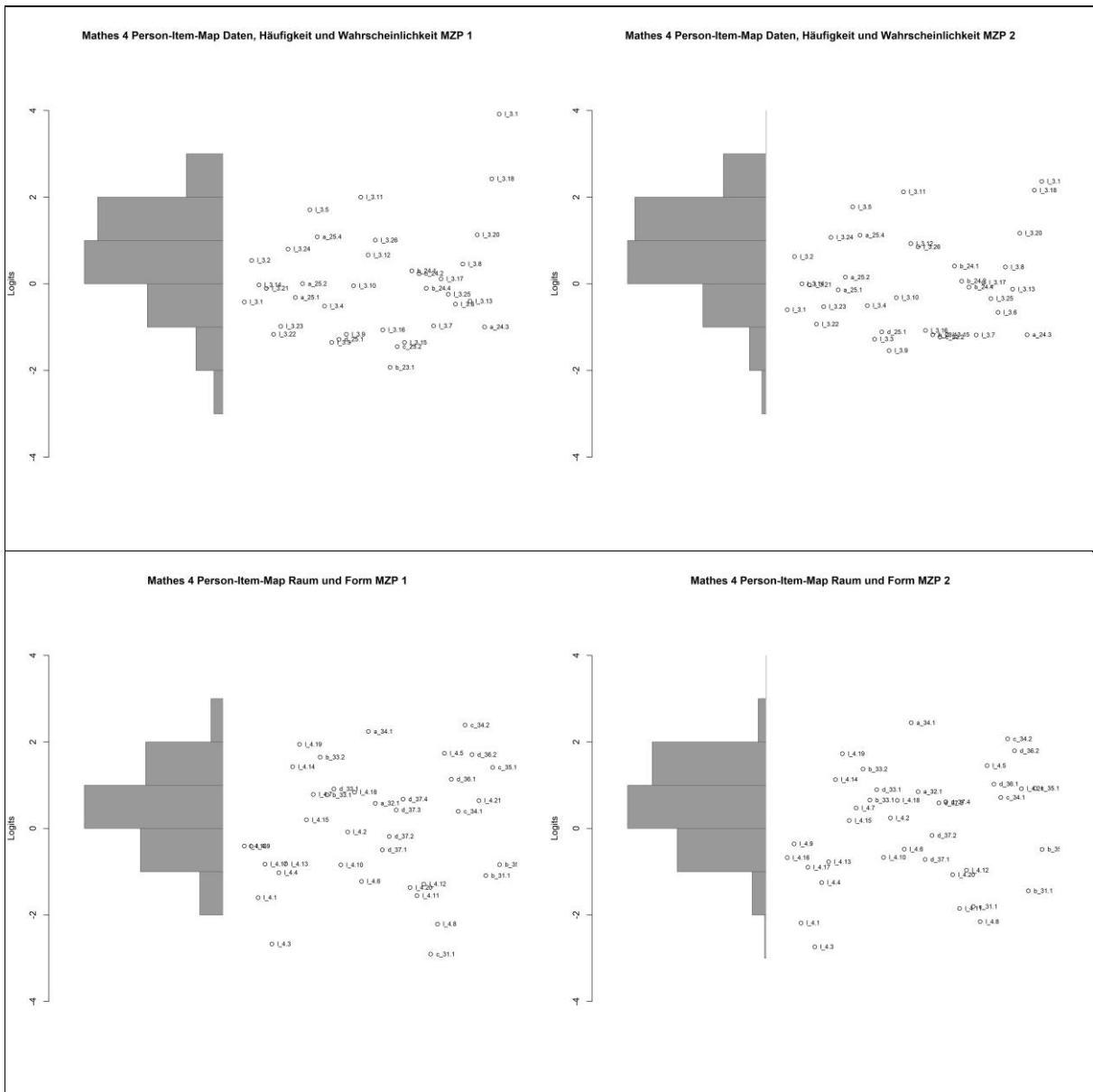


Abbildung 3: Verteilungen von Aufgabenschwierigkeiten und Kompetenzwerten zu den zwei Messzeitpunkten

Aus den Darstellungen geht hervor, dass „Mathes 4“ in allen vier Inhaltsbereichen den mittleren und oberen Leistungsbereich gut bis sehr gut abdeckt, mit Ausnahme der Skala Größen und Messen, welche zum zweiten Messzeitpunkt bei sehr leistungsstarken Schülerinnen und Schülern nicht mehr ausreichend differenziert. In den Skalen Zahlen und Operationen sowie Raum und Form sind auch im unteren Leistungsbereich zumindest ausreichend Items vertreten. Auffällig ist, dass es in der Skala Größen und Messen zum ersten Messzeitpunkt einige wenige Schülerinnen und Schüler gibt, die keine Aufgabe lösen können.

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 4“ zu den beiden Messzeitpunkten befindet sich im Anhang A auf Seite 22ff.

4.2 Reliabilität

Zur Schätzung der Zuverlässigkeit des Verfahrens im Sinne der Internen Konsistenz wurde die sogenannte WLE-Reliabilität (Rost, 2004) berechnet. Die resultierenden Koeffizienten sind von der Höhe vergleichbar mit Cronbachs α (Rost, 2004), welcher zusätzlich für jeden Subtest der vier Testformen des „Mathes 4“ ermittelt wurde. Die WLE-Reliabilitätskoeffizienten liegen zwischen $r_{WLE} = .72$ und $r_{WLE} = .86$. Die Tabelle 3 stellt die WLE-Reliabilitäten der vier Testskalen zu den beiden Messzeitpunkten dar.

Tabelle 3: Ergebnisse der Analysen hinsichtlich der WLE-Reliabilität der Skalen des "Mathes 4"

WLE-Reliabilität von				
	Zahlen und Operationen	Größen und Messen	Daten, Häufigkeit und Wahrscheinlichkeit	Raum und Form
MZP 1	.86 <i>N</i> = 780	.81 <i>N</i> = 780	.83 <i>N</i> = 780	.74 <i>N</i> = 780
MZP 2	.85 <i>N</i> = 691	.81 <i>N</i> = 691	.82 <i>N</i> = 691	.72 <i>N</i> = 691

Anmerkungen: MZP – Messzeitpunkt; *N* – Stichprobenumfang

Die Interne Konsistenz der Subtests des „Mathes 4“ liegt überwiegend im hohen Bereich mit Koeffizienten zwischen $\alpha = .66$ und $\alpha = .88$ (*MW*: $\alpha = .82$, *SD* = $.07$). Die nachfolgende Tabelle 4 stellt die Alphakoeffizienten der Subtests des „Mathes 4“ getrennt nach Testform zu den beiden Messzeitpunkten dar.

Tabelle 4: Ergebnisse der Analysen hinsichtlich der Internen Konsistenz (Cronbachs α) der Subtests des "Mathes 4"

Reliabilitätskoeffizient Cronbachs α von					
Form	MZP	Zahlen und Operationen	Größen und Messen	Daten, Häufigkeit und Wahrscheinlichkeit	Raum und Form
A	1	.85	.85	.86	.68
	2	.84	.84	.85	.66
B	1	.87	.86	.86	.69
	2	.86	.85	.83	.73
C	1	.87	.88	.84	.70
	2	.86	.88	.81	.68
D	1	.86	.87	.85	.76
	2	.85	.88	.83	.73

Anmerkungen: MZP – Messzeitpunkt, α – Koeffizient nach Cronbach

4.3 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 sensibler im Bereich mittlerer Fähigkeitsausprägungen ausfallen (Linacre, 2002), wurden für den „Mathes 4“ 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.

Nachfolgend werden die Ergebnisse der Raschskalierung präsentiert. Die Infit-Werte wurden aus Gründen der Übersichtlichkeit in der Abbildung 4 mittels Boxplots graphisch aufbereitet, sodass die Streuung der Werte des „Mathes 4“ schnell ersichtlich ist.

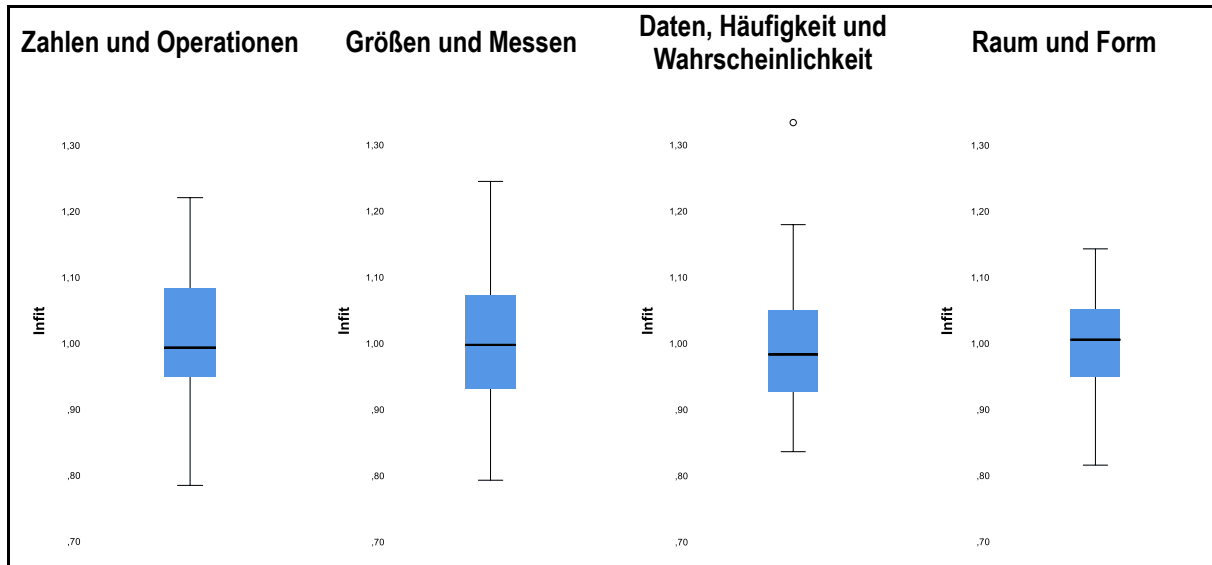


Abbildung 4: Verteilung der Itemfit-Statistiken der vier Testskalen des „Mathes 4“

Über alle Skalen des „Mathes 4“ hinweg liegen die Infit-Werte der Aufgaben innerhalb des zulässigen Wertebereiches, mit Ausnahme eines Items zum Wahrscheinlichkeitsbegriff (Form C: 28.1, Form D: 28.1) in der Skala Daten, Häufigkeit und Wahrscheinlichkeit, welches einen leichten Underfit mit einem Infit-Wert von 1.33 aufweist. Dennoch ist der Großteil der Items in allen Testskalen in der Nähe des Erwartungswertes von 1. Eine Auflistung der im Rasch-Modell geschätzten Fitstatistiken zu jedem Item kann im Anhang B auf S. 27ff. eingesehen werden.

Eine zentrale Voraussetzung für Veränderungsmessungen ist die Stichprobeninvarianz der Itemparameter. Diese wurde analog zu Vorarbeiten mit ähnlicher Zielstellung (z. B. Gebhardt, Heine, Zeuch & Förster, 2015) auf verschiedene Messzeitpunkte bezogen und mittels graphischer Modelltests geprüft. Die Itemparameter erweisen sich – insgesamt betrachtet – zwischen den Messzeitpunkten als annähernd konstant, d. h. messinvariant über die Zeit (s. Anhang C, S. 32f. für weitere Informationen). Nur bei wenigen Aufgaben sind diesbezüglich kleinere Abweichungen festzustellen.

4.4 Validität

-- Angaben werden im Rahmen einer neuen Auflage ergänzt --

5 Literaturverzeichnis

- Blum, W. (2010). Die Bildungsstandards Mathematik. In W. Blum, C. Drücke-Noe, R. Hartung & O. Köller (Hrsg.), *Bildungsstandards Mathematik: konkret. Sekundarstufe I. Aufgabenbeispiele, Unterrichts Anregungen, Fortbildungsideen*. 4. Auflage (S. 14-32). Berlin: Cornelsen.
- 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., Bensen, 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.
- 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.
- Devlin, K. (1998). *Muster der Mathematik*. Heidelberg: Spektrum.
- Dilling H., Mombour, W. & Schmidt, M.H. (2011). Internationale Klassifikation psychischer Störungen. ICD-10 Kapitel 5 (F). Klinisch-diagnostische Leitlinien. Bern: Huber.
- Führer, L. (1997). *Pädagogik des Mathematikunterrichts. Eine Einführung in die Fachdidaktik für Sekundarstufen*. Braunschweig & Wiesbaden: Vieweg.
- Gaidoschik, M. (2011). *Rechenschwäche – Dyskalkulie. Eine unterrichtspraktische Einführung für LehrerInnen und Eltern*. 6. Auflage. Hamburg: Persen.
- Grassmann, M., Eichler, K.-P., Mirwald, E. & Nitsch, B. (2014). *Mathematikunterricht. Kompetent im Unterricht der Grundschule*. 3. korrigierte und veränderte Auflage. Baltmannsweiler: Schneider Verlag Hohengehren.
- Hartke, B. & Diehl, K. (2013). *Schulische Prävention im Bereich Lernen. Problemlösungen mit dem RTI-Ansatz*. Stuttgart: Kohlhammer.
- 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.

- Klieme, E., Avenarius, H., Blum, W., Gruber, H., Reiss, K., Riquarts, K., Rost, J., Tenorth, H.-E., Vollmer, H.J. (2007). *Zur Entwicklung nationaler Bildungsstandards. Expertise*. Berlin: BMBF.
- 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. & Knopp, E. (2010). Mathematisches Lernen. In B. Hartke, K. Koch & K. Diehl (Hrsg.), *Förderung in der schulischen Eingangsstufe* (S. 91-118). Stuttgart: Kohlhammer.
- Kolen, M.J. & Brennan, R.J. (2004). *Test equating, scaling, and linking: methods and practice*. 2nd edition. New York: Springer.
- Krajewski, K. (2003): Vorhersage von Rechenschwäche in der Grundschule. Hamburg: Dr. Kovač.
- Kultusministerkonferenz (2005). *Bildungsstandards im Fach Mathematik für den Primarbereich*. München: Luchterhand.
- 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.
- Mislevy, R.J., Beaton, A.E., Kaplan, B. & Sheehan, K.M. (1992). Estimating population characteristics from sparse matrix samples of item responses. *Journal of Educational Measurement*, 29, 133-161.
- Moser Opitz, E. (2004). Dyskalkulie: Krankheit, Erfindung, Mythos, Etikett ... ? Auseinandersetzung mit einem geläufigen, aber ungeklärten Begriff. *Vierteljahrszeitschrift für Heilpädagogik und ihre Nachbargebiete*, 72, 179-190.
- 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, 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.
- 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. & Hartke, B. (2012). *Mathes 3. Formative Erfassung der mathematischen Kompetenzen von Drittklässlern*. Abruf am 13.5.14. Online verfügbar unter <http://www.lernlinie.de/to/mathes3>

- Sikora, S. (2017). *Lernverlaufsdagnostik 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.
- Vohns, A. (2007). *Grundlegende Ideen und Mathematikunterricht. Entwicklung und Perspektiven eines fachdidaktischen Prinzips*. Norderstedt: Books on Demand.
- Von Aster, M., Schweiter, M. & Weinhold Zulauf, M. (2007). Rechenstörungen bei Kindern. Vorläufer, Prävalenz und psychische Symptome. *Zeitschrift für Pädagogische Psychologie*, 39, 85-96.
- Walther, G., van den Heuvel-Panhuizen, M., Granzer, D. & Köller, O. (Hrsg.) (2011). *Bildungsstandards für die Grundschule. Mathematik konkret. 5. Auflage*. Berlin: Cornelsen.
- Weinert, F.E. (2001). Vergleichende Leistungsmessung in Schulen – eine umstrittene Selbstverständlichkeit. In F.E. Weinert (Hrsg.), *Leistungsmessungen in Schulen* (S. 17-31). Weinheim: Beltz.
- Winkelmann, H., Robitzsch, A., Stanat, P. & Köller, O. (2012). Mathematische Kompetenzen in der Grundschule. Struktur, Validierung und Zusammenspiel mit allgemeinen kognitiven Fähigkeiten. *Diagnostica*, 58, 15-30.
- Winter, H. (2001). *Inhalte mathematischen Lernens. Expertise für das Ministerium für Bildung, Wissenschaft, Jugend und Kultur des Landes Rheinland-Pfalz*. Abruf am 15.4.20. Online verfügbar unter: http://grundschule.bildung-rp.de/fileadmin/user_upload/grundschule.bildung-rp.de/Downloads/Mathematik/Winter_Inhalte_math_Lernens.pdf
- Wittmann, E.Ch. (1995). Aktiv-entdeckendes und soziales Lernen im Arithmetikunterricht. In G.N. Müller & E.Ch. Wittmann (Hrsg.), *Mit Kindern rechnen* (S. 10-41). Frankfurt am Main: Arbeitskreis Grundschule – Der Grundschulverband – e.V.
- Wittmann, E.Ch. (1999). Konstruktion eines Geometrieunterrichts ausgehend von Grundideen der Elementargeometrie. In H. Henning (Hrsg.), *Mathematik lernen durch Handeln und Erfahrung. Festschrift zum 75. Geburtstag von Heinrich Besuden* (S. 205-223). Oldenburg: Bültmann und Gerriets.
- Wittmann, E.Ch. & Müller, G.N. (2011). Muster und Strukturen als fachliches Grundkonzept. In G. Walther, M. van den Heuvel-Panhuizen, D. Granzer & O. Köller (Hrsg.), *Bildungsstandards für die Grundschule: Mathematik konkret. 5. Auflage* (S. 42-65). Berlin: Cornelsen.
- Wittmann, E.Ch. & Müller, G.N. (2012). *Das Zahlenbuch 1 (Neubearbeitung). Begleitband*. Stuttgart: Klett.

6 Anhang

Anhang A: Schwierigkeiten und Trennschärfen der Items

Zahlen und Operationen

Item				Schwierigkeit P_i zum Messzeitpunkt		Trennschärfe r_{pbis} zum Messzeitpunkt	
Form A	Form B	Form C	Form D	1	2	1	2
1.1				0.84	0.86	0.28	0.31
1.2				0.77	0.83	0.46	0.35
1.3				0.71	0.76	0.42	0.43
2.1			7.1	0.65	0.66	0.62	0.59
2.2		6.2		0.17	0.23	0.51	0.48
3.1		3.1		0.81	0.86	0.48	0.43
3.2				0.74	0.78	0.58	0.59
3.3		3.3		0.48	0.53	0.56	0.48
4.1	2.1			0.88	0.90	0.42	0.30
4.2				0.70	0.60	0.52	0.44
5.1			4.1	0.25	0.35	0.43	0.42
6.1	3.2			0.55	0.64	0.67	0.54
6.2			5.2	0.15	0.16	0.49	0.48
7.1	6.1			0.45	0.53	0.60	0.59
7.2		7.2		0.31	0.34	0.53	0.52
8.1			9.1	0.81	0.80	0.49	0.49
8.2				0.75	0.76	0.57	0.54
8.3				0.64	0.72	0.58	0.49
8.4		8.4		0.35	0.41	0.52	0.59
9.1	8.1	9.1	10.1	0.65	0.68	0.56	0.52
9.2	8.2			0.31	0.36	0.35	0.30
10.1	9.2	10.1	11.2	0.31	0.29	0.52	0.49
10.2	9.1	10.2	11.1	0.25	0.34	0.49	0.58
11.1				0.28	0.30	0.41	0.43
12.1				0.35	0.42	0.43	0.40
13.1				0.31	0.29	0.50	0.33
	1.1			0.74	0.82	0.51	0.51
	1.2		1.2	0.73	0.78	0.56	0.51
	1.3		1.3	0.65	0.70	0.45	0.50
	2.2			0.58	0.65	0.45	0.37
	3.1	4.1		0.79	0.81	0.63	0.51
	4.1			0.64	0.63	0.61	0.65
	4.2			0.61	0.64	0.57	0.66
	4.3			0.44	0.58	0.58	0.64
	5.1	6.1		0.46	0.51	0.55	0.57
	5.2		7.2	0.27	0.33	0.50	0.52
	6.2		6.2	0.24	0.38	0.56	0.55
	7.1	8.1		0.88	0.87	0.52	0.35

7.2		0.70	0.78	0.57	0.54	
7.3	9.3	0.77	0.71	0.56	0.38	
7.4		0.61	0.63	0.60	0.57	
10.1		0.23	0.25	0.44	0.51	
11.1		0.19	0.49	0.40	0.41	
12.1		0.20	0.31	0.42	0.51	
13.1		0.22	0.22	0.41	0.45	
	1.1	0.39	0.34	0.39	0.43	
	1.2	0.51	0.57	0.52	0.48	
	1.3	0.62	0.66	0.52	0.51	
	2.1	0.30	0.28	0.41	0.32	
	3.2	0.71	0.75	0.59	0.45	
	4.2	5.1	0.65	0.74	0.63	0.58
	5.1	0.76	0.79	0.44	0.28	
	5.2	0.67	0.70	0.44	0.46	
	7.1	6.1	0.51	0.55	0.55	0.48
	8.2	0.65	0.64	0.62	0.66	
	8.3	0.48	0.54	0.57	0.55	
	9.2	10.2	0.36	0.47	0.47	0.38
	11.1	0.31	0.41	0.50	0.51	
	12.1	0.14	0.42	0.19	0.42	
	13.1	0.29	0.29	0.55	0.40	
		1.1	0.71	0.81	0.51	0.44
		2.1	0.65	0.72	0.45	0.41
		3.1	0.77	0.74	0.51	0.45
		3.2	0.50	0.50	0.42	0.42
		8.1	0.28	0.34	0.46	0.47
		9.2	0.79	0.82	0.57	0.59
		9.4	0.41	0.43	0.49	0.42
		12.1	0.12	0.24	0.44	0.59
		13.1	0.40	0.48	0.42	0.56
		14.1	0.35	0.48	0.48	0.50

Größen und Messen

Item				Schwierigkeit P_i zum Messzeitpunkt		Trennschärfe r_{pbis} zum Messzeitpunkt	
Form A	Form B	Form C	Form D	1	2	1	2
14.1				0.88	0.91	0.53	0.39
14.2	15.1			0.68	0.76	0.54	0.43
14.3		14.4		0.37	0.50	0.46	0.50
14.4			15.4	0.30	0.44	0.48	0.49
15.1	16.1			0.90	0.94	0.49	0.44
15.2	16.2			0.80	0.87	0.38	0.48
15.3				0.88	0.93	0.57	0.44
16.1		16.3		0.56	0.58	0.67	0.66

16.2	17.1	16.2	17.1	0.52	0.57	0.62	0.64
16.3				0.23	0.32	0.51	0.58
17.1		17.2	18.1	0.68	0.76	0.65	0.60
17.2	18.1	17.1		0.72	0.67	0.52	0.55
17.3	18.2		18.3	0.40	0.52	0.51	0.52
18.1	19.2	18.1	19.2	0.49	0.56	0.41	0.42
18.2	19.1	18.2	19.1	0.63	0.66	0.55	0.44
19.1				0.28	0.29	0.46	0.44
20.1				0.09	0.21	0.24	0.48
21.1				0.29	0.28	0.32	0.30
22.1				0.25	0.26	0.39	0.32
	14.1			0.67	0.68	0.52	0.47
	15.2		15.1	0.68	0.74	0.53	0.46
	15.3			0.71	0.70	0.34	0.21
	15.4	14.3		0.46	0.60	0.54	0.47
	16.3			0.83	0.91	0.55	0.51
	17.2		17.3	0.39	0.40	0.52	0.53
	17.3			0.18	0.25	0.51	0.54
	18.3	17.3	18.2	0.36	0.36	0.56	0.50
	20.1			0.34	0.38	0.41	0.44
	21.1			0.50	0.38	0.53	0.52
	21.2			0.31	0.30	0.51	0.53
	21.3			0.53	0.44	0.54	0.52
	21.4			0.36	0.27	0.49	0.51
	21.5			0.62	0.56	0.59	0.56
		14.1	15.2	0.72	0.74	0.43	0.53
		14.2	15.3	0.61	0.72	0.48	0.46
		15.1	16.2	0.33	0.41	0.58	0.69
		15.2	16.1	0.31	0.41	0.52	0.68
		15.3	16.4	0.24	0.30	0.50	0.56
		15.4	16.3	0.31	0.34	0.50	0.61
		16.1	17.2	0.57	0.65	0.56	0.60
		19.1		0.69	0.66	0.47	0.42
		20.1		0.18	0.27	0.39	0.45
		21.1		0.67	0.67	0.52	0.54
			20.1	0.58	0.54	0.52	0.55
			21.1	0.51	0.52	0.59	0.50
			22.1	0.59	0.67	0.44	0.48
			23.1	0.21	0.31	0.18	0.37

Daten, Häufigkeit und Wahrscheinlichkeit

Item				Schwierigkeit P_i zum Messzeitpunkt		Trennschärfe r_{pbis} zum Messzeitpunkt	
Form A	Form B	Form C	Form D	1	2	1	2
23.1	22.2			0.69	0.75	0.56	0.55
23.2			24.1	0.52	0.54	0.46	0.47
24.1			25.2	0.83	0.84	0.53	0.54
24.2			25.3	0.71	0.74	0.59	0.57
24.3				0.77	0.80	0.58	0.64
25.1				0.64	0.66	0.58	0.60
25.2				0.58	0.61	0.58	0.58
25.3	24.3			0.28	0.30	0.43	0.40
25.4				0.38	0.43	0.54	0.55
26.1	25.2			0.69	0.74	0.57	0.54
26.2			26.1	0.75	0.82	0.69	0.56
26.3	25.3	23.3	26.3	0.51	0.56	0.45	0.46
27.1			29.2	0.80	0.87	0.49	0.44
27.2			29.1	0.63	0.71	0.45	0.45
27.3			29.3	0.21	0.22	0.25	0.28
28.1	27.1			0.50	0.48	0.55	0.47
29.1			30.1	0.68	0.66	0.51	0.56
	22.1	25.2		0.63	0.66	0.54	0.53
	23.1			0.91	0.87	0.49	0.48
	23.2	24.3		0.86	0.86	0.59	0.57
	23.3	24.1		0.85	0.87	0.52	0.42
	24.1			0.60	0.60	0.53	0.47
	24.2			0.61	0.66	0.58	0.57
	24.4			0.66	0.69	0.52	0.53
	25.1	23.2		0.59	0.64	0.49	0.35
	26.1	26.1		0.16	0.20	0.29	0.35
	26.2	26.2		0.04	0.18	0.14	0.37
	28.1	27.1		0.38	0.39	0.44	0.42
		22.1	27.2	0.76	0.79	0.58	0.58
		22.2	27.1	0.74	0.74	0.58	0.56
		22.3	27.3	0.41	0.44	0.52	0.57
		23.1	26.2	0.65	0.70	0.63	0.62
		24.2		0.88	0.88	0.51	0.48
		25.1	24.2	0.64	0.66	0.56	0.55
		28.1	28.1	0.41	0.47	0.33	0.24
			25.1	0.81	0.83	0.57	0.50

Raum und Form

Item				Schwierigkeit P_i zum Messzeitpunkt		Trennschärfe r_{pbis} zum Messzeitpunkt	
Form A	Form B	Form C	Form D	1	2	1	2
30.1	29.2			0.87	0.93	0.39	0.35
30.2			31.1	0.58	0.58	0.40	0.38
31.1			32.2	0.92	0.95	0.53	0.31
31.2		29.1		0.77	0.83	0.51	0.46
31.3	30.4			0.22	0.30	0.28	0.43
32.1				0.43	0.43	0.41	0.42
33.1			34.2	0.90	0.91	0.39	0.40
33.2	32.1			0.66	0.70	0.45	0.43
34.1			36.2	0.76	0.72	0.54	0.30
34.2	34.1			0.43	0.53	0.40	0.46
35.1				0.14	0.17	0.26	0.44
36.1			38.1	0.78	0.76	0.34	0.39
	29.1	30.2		0.83	0.86	0.46	0.54
	30.1		32.1	0.76	0.78	0.64	0.57
	30.2	29.2		0.76	0.78	0.43	0.41
	30.3	29.3		0.25	0.33	0.30	0.25
	31.1			0.78	0.83	0.49	0.47
	32.2	32.1		0.70	0.77	0.51	0.46
	33.1			0.45	0.50	0.55	0.56
	33.2			0.27	0.35	0.48	0.53
	34.2	33.1		0.53	0.57	0.44	0.45
	35.1			0.75	0.71	0.40	0.50
		29.4	32.3	0.38	0.49	0.44	0.42
		30.1	31.2	0.73	0.78	0.51	0.51
		31.1		0.93	0.89	0.34	0.39
		32.2	34.1	0.81	0.79	0.42	0.39
		33.2	36.1	0.19	0.24	0.33	0.37
		34.1		0.29	0.39	0.39	0.33
		35.1	35.3	0.45	0.41	0.49	0.46
		35.2		0.50	0.46	0.47	0.50
		35.3		0.12	0.18	0.19	0.28
			33.1	0.41	0.44	0.35	0.37
			35.1	0.35	0.39	0.45	0.43
			35.2	0.23	0.25	0.19	0.39
			37.1	0.66	0.74	0.51	0.48
			37.2	0.59	0.65	0.58	0.54
			37.3	0.46	0.50	0.52	0.50
			37.4	0.41	0.50	0.47	0.57

Anhang B: Itemkennwerte der Rasch-Analyse

Zahlen und Operationen

Form A	Form B	Form C	Form D	N	Schwierigkeit	Infit
1.1				750	-1.69	1.11
1.2				372	-1.29	1.01
1.3				378	-1.05	1.14
2.1			7.1	728	-0.62	0.85
2.2		6.2		737	2.13	0.92
3.1		3.1		728	-1.83	0.96
3.2				728	-1.31	0.85
3.3		3.3		356	0.23	1.00
4.1	2.1			356	-2.46	1.01
4.2				356	-0.66	1.10
5.1			4.1	372	1.29	1.00
6.1	3.2			356	-0.27	0.85
6.2			5.2	356	2.71	1.00
7.1	6.1			378	0.35	0.90
7.2		7.2		365	1.15	0.95
8.1			9.1	378	-1.71	0.99
8.2				1471	-1.35	0.91
8.3				378	-0.75	0.97
8.4		8.4		378	1.06	0.98
9.1	8.1	9.1	10.1	1471	-0.69	0.97
9.2	8.2			356	0.90	1.22
10.1	9.2	10.1	11.2	356	1.44	0.99
10.2	9.1	10.2	11.1	378	1.62	0.97
11.1				378	1.33	1.12
12.1				372	0.86	1.08
13.1				356	1.24	1.05
	1.1			721	-1.55	0.96
	1.2		1.2	721	-1.35	0.93
	1.3		1.3	721	-0.75	1.05
	2.2			356	-0.47	1.21
	3.1	4.1		356	-1.72	0.85
	4.1			378	-0.56	0.87
	4.2			378	-0.47	0.89
	4.3			378	0.25	0.87
	5.1	6.1		743	0.26	0.95
	5.2		7.2	737	1.39	0.94
	6.2		6.2	378	1.49	0.89
	7.1	8.1		356	-2.42	0.97
	7.2			365	-1.25	0.88
	7.3		9.3	365	-1.24	0.99
	7.4			1471	-0.37	0.89
	10.1			378	1.76	0.97

11.1			365	0.92	1.11
12.1			372	1.81	0.97
13.1			378	1.90	1.03
	1.1		750	0.91	1.18
	1.2		378	0.06	1.08
	1.3		750	-0.49	1.01
	2.1		365	1.15	1.10
	3.2		356	-1.13	0.96
	4.2	5.1	356	-0.93	0.85
	5.1		365	-1.41	1.12
	5.2		365	-0.90	1.11
	7.1	6.1	365	-0.04	1.00
	8.2		365	-0.70	0.79
	8.3		378	0.28	0.98
	9.2	10.2	372	0.78	1.10
	11.1		356	1.05	1.01
	12.1		356	1.26	1.14
	13.1		365	1.34	1.01
		1.1	356	-1.32	1.00
		2.1	365	-0.84	1.09
		3.1	356	-1.31	1.03
		3.2	356	0.16	1.19
		8.1	372	1.35	0.97
		9.2	378	-1.91	0.95
		9.4	356	0.60	1.10
		12.1	356	2.43	0.95
		13.1	378	0.58	1.10
		14.1	378	0.74	1.06

Größen und Messen

Form A	Form B	Form C	Form D	N	Schwierigkeit	Infit
14.1				372	-2.34	0.95
14.2	15.1			378	-1.17	1.10
14.3		14.4		378	0.48	1.13
14.4			15.4	378	0.77	1.06
15.1	16.1			1471	-3.13	0.89
15.2	16.2			365	-1.83	0.99
15.3				356	-3.05	0.94
16.1		16.3		372	0.24	0.80
16.2	17.1	16.2	17.1	1471	-0.08	0.86
16.3				365	1.56	0.96
17.1		17.2	18.1	372	-0.80	0.85
17.2	18.1	17.1		728	-0.84	0.97
17.3	18.2		18.3	1471	0.42	1.04

18.1	19.2	18.1	19.2	356	-0.12	1.11
18.2	19.1	18.2	19.1	356	-0.77	1.00
19.1				378	1.28	1.13
20.1				356	2.15	1.04
21.1				356	1.01	1.20
22.1				372	1.73	1.09
	14.1			378	-0.85	1.11
	15.2		15.1	372	-0.59	1.04
	15.3			356	-0.93	1.25
	15.4	14.3		365	-0.06	1.05
	16.3			378	-2.44	0.80
	17.2		17.3	750	0.92	1.04
	17.3			378	2.03	0.95
	18.3	17.3	18.2	365	0.88	1.00
	20.1			365	0.79	1.17
	21.1			356	0.21	0.96
	21.2			356	0.93	0.90
	21.3			356	0.00	0.96
	21.4			356	0.79	0.92
	21.5			356	-0.52	0.90
		14.1	15.2	356	-1.22	1.00
		14.2	15.3	365	-0.79	1.08
		15.1	16.2	378	0.76	0.79
		15.2	16.1	378	0.83	0.85
		15.3	16.4	378	1.40	0.90
		15.4	16.3	378	1.01	0.93
		16.1	17.2	1471	-0.44	0.94
		19.1		356	-0.90	1.07
		20.1		372	1.94	0.94
		21.1		372	-0.40	1.02
			20.1	372	0.20	1.03
			21.1	365	-0.03	1.01
			22.1	365	-0.63	1.13
			23.1	372	1.58	1.18

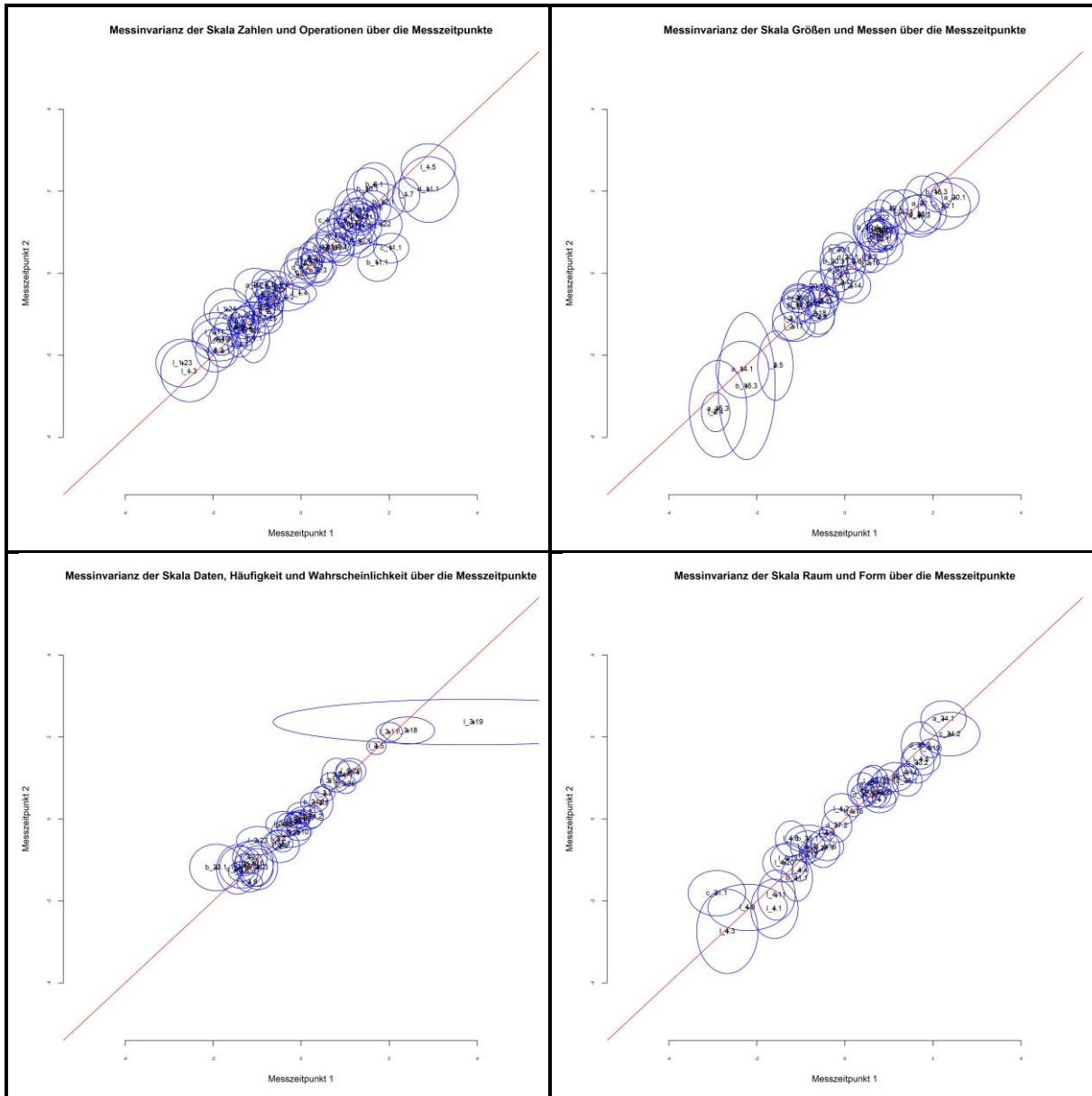
Daten, Häufigkeit und Wahrscheinlichkeit

Form A	Form B	Form C	Form D	N	Schwierigkeit	Infit
23.1	22.2			1471	-0.49	0.99
23.2			24.1	1471	0.58	1.12
24.1			25.2	1471	-1.32	0.93
24.2			25.3	1471	-0.51	0.92
24.3				356	-1.08	0.90
25.1				743	-0.23	0.94
25.2				743	0.08	0.96
25.3	24.3			1471	1.74	1.07
25.4				743	1.11	0.97
26.1	25.2			365	-0.55	0.96
26.2			26.1	365	-1.07	0.84
26.3	25.3	23.3	26.3	365	0.43	1.14
27.1			29.2	1471	-1.32	1.02
27.2			29.1	1471	-0.16	1.16
27.3			29.3	1471	2.06	1.10
28.1	27.1			1471	0.79	1.03
29.1			30.1	365	-0.27	1.03
	22.1	25.2		1471	-0.01	1.02
	23.1			372	-1.54	0.95
	23.2	24.3		372	-1.27	0.86
	23.3	24.1		372	-1.07	0.91
	24.1			728	0.35	1.06
	24.2			728	0.16	0.95
	24.4			728	-0.09	1.04
	25.1	23.2		365	0.08	1.18
	26.1	26.1		356	2.29	1.01
	26.2	26.2		356	2.96	1.00
	28.1	27.1		365	1.15	1.10
		22.1	27.2	378	-1.05	0.91
		22.2	27.1	378	-0.76	0.93
		22.3	27.3	378	0.93	0.98
		23.1	26.2	365	-0.28	0.88
		24.2		372	-1.34	0.90
		25.1	24.2	1471	-0.06	0.98
		28.1	28.1	1471	0.94	1.33
			25.1	1471	-1.20	0.92

Raum und Form

Form A	Form B	Form C	Form D	N	Schwierigkeit	Infit
30.1	29.2			378	-1.85	0.93
30.2			31.1	372	0.06	1.11
31.1			32.2	378	-2.71	0.96
31.2		29.1		1471	-1.13	0.91
31.3	30.4			356	1.59	1.04
32.1				372	0.70	1.08
33.1			34.2	356	-2.19	0.92
33.2	32.1			1471	-0.39	1.04
34.1			36.2	372	-0.83	1.01
34.2	34.1			378	0.64	1.07
35.1				372	2.33	1.01
36.1			38.1	378	-0.77	1.06
	29.1	30.2		365	-1.69	0.90
	30.1		32.1	365	-1.13	0.82
	30.2	29.2		1471	-0.80	0.97
	30.3	29.3		1471	1.27	1.14
	31.1			356	-1.26	0.96
	32.2	32.1		378	-0.54	0.95
	33.1			378	0.71	0.92
	33.2			378	1.52	0.93
	34.2	33.1		1471	0.18	1.06
	35.1			356	-0.67	1.00
		29.4	32.3	372	0.74	1.05
		30.1	31.2	1471	-0.86	0.94
		31.1		365	-2.29	0.99
		32.2	34.1	365	-1.23	1.04
		33.2	36.1	1471	1.83	1.01
		34.1		356	1.16	1.07
		35.1	35.3	356	0.77	1.01
		35.2		356	0.53	0.99
		35.3		356	2.22	1.02
			33.1	378	0.90	1.14
			35.1	356	1.08	1.02
			35.2	356	1.74	1.13
			37.1	372	-0.60	0.96
			37.2	372	-0.18	0.91
			37.3	372	0.49	0.97
			37.4	372	0.64	0.95

Anhang C: Messinvarianz der Testskalen des „Mathes 4“ über die beiden Messzeitpunkte



Anmerkungen: Zur Prüfung der Messinvarianz der Aufgaben über die Zeit wurden die Itemparameter im Rahmen von graphischen Modelltests jeweils an der X- und Y-Achse abgetragen. Die Konfidenzintervalle (95 %) werden durch Ellipsen dargestellt.

Anhang D: Normtabellen des „Mathes 4“

Mathes 4 Form A (20./21. Schulwoche)

	Rohwert	Prozentrang	T-Wert		Rohwert	Prozentrang	T-Wert
Zahlen und Operationen	0	0	<22	Größen und Messen	0	0	23
	1	0	22		1	5	26
	2	1	24		2	6	28
	3	2	26		3	8	31
	4	3	28		4	9	34
	5	5	30		5	12	36
	6	8	33		6	19	39
	7	12	35		7	28	42
	8	17	37		8	37	44
	9	22	39		9	43	47
	10	26	42		10	50	49
	11	31	44		11	61	52
	12	38	46		12	71	55
	13	45	48		13	80	57
	14	52	50		14	92	60
	15	62	53		15	97	62
	16	71	55		16	99	65
	17	79	57		17	100	68
	18	86	59		18	100	70
	19	92	61		19	100	>70
	20	96	64				
	21	99	66				
	22	100	68				
	23	100	70				
	24	100	>70				
	25	100	>70				
26	100	>70					

	Rohwert	Prozentrang	T-Wert		Rohwert	Prozentrang	T-Wert
Daten, Häufigkeit und Wahrscheinlichkeit	0	1	23	Raum und Form	0	1	11
	1	4	25		1	2	16
	2	6	27		2	3	21
	3	9	30		3	5	26
	4	11	32		4	7	31
	5	13	35		5	11	36
	6	15	37		6	21	41
	7	19	40		7	39	46
	8	23	42		8	61	51
	9	30	45		9	84	56
	10	37	47		10	97	61
	11	44	50		11	100	66
	12	56	52		12	100	70
	13	67	54				
	14	79	57				
	15	91	59				
	16	97	62				
17	100	64					

Mathes 4 Form A (40./41. Schulwoche)

	Rohwert	Prozentrang	T-Wert		Rohwert	Prozentrang	T-Wert
Zahlen und Operationen	0	0	<24	Größen und Messen	0	0	23
	1	0	<24		1	3	26
	2	0	24		2	4	28
	3	1	26		3	5	31
	4	2	28		4	6	34
	5	3	30		5	10	36
	6	7	33		6	16	39
	7	10	35		7	21	42
	8	13	37		8	26	44
	9	18	39		9	33	47
	10	21	42		10	42	49
	11	27	44		11	52	52
	12	31	46		12	62	55
	13	39	48		13	73	57
	14	47	50		14	86	60
	15	57	53		15	95	62
	16	68	55		16	99	65
	17	76	57		17	100	68
	18	83	59		18	100	70
	19	90	61		19	100	>70
	20	95	64				
	21	98	66				
	22	99	68				
	23	100	70				
	24	100	73				
	25	100	>73				
26	100	>73					

	Rohwert	Prozentrang	T-Wert		Rohwert	Prozentrang	T-Wert
Daten, Häufigkeit und Wahrscheinlichkeit	0	1	23	Raum und Form	0	1	11
	1	3	25		1	1	16
	2	5	27		2	2	21
	3	7	30		3	3	26
	4	8	32		4	5	31
	5	10	35		5	9	36
	6	12	37		6	18	41
	7	16	40		7	33	46
	8	20	42		8	55	51
	9	25	45		9	79	56
	10	31	47		10	93	61
	11	41	50		11	99	66
	12	51	52		12	100	70
	13	64	54				
	14	76	57				
	15	90	59				
	16	97	62				
	17	100	64				

Mathes 4 Form B (20./21. Schulwoche)

	Rohwert	Prozentrang	T-Wert		Rohwert	Prozentrang	T-Wert
Zahlen und Operationen	0	0	20	Größen und Messen	0	2	22
	1	0	20		1	4	24
	2	1	25		2	5	26
	3	2	27		3	6	28
	4	5	29		4	7	30
	5	7	31		5	7	33
	6	9	33		6	10	35
	7	12	35		7	13	37
	8	15	37		8	16	39
	9	19	39		9	23	41
	10	24	41		10	32	44
	11	29	43		11	38	46
	12	35	45		12	46	48
	13	42	47		13	53	50
	14	49	49		14	60	52
	15	56	51		15	70	55
	16	66	53		16	78	57
	17	75	55		17	87	59
	18	81	58		18	93	61
	19	87	60		19	97	63
	20	93	62		20	99	65
	21	96	64		21	100	68
	22	99	66		22	100	70
	23	100	68				
	24	100	70				
	25	100	72				
26	100	>72					

	Rohwert	Prozentrang	T-Wert		Rohwert	Prozentrang	T-Wert
Daten, Häufigkeit und Wahrscheinlichkeit	0	3	24	Raum und Form	0	1	18
	1	6	27		1	2	21
	2	8	30		2	3	25
	3	12	33		3	5	29
	4	15	35		4	7	33
	5	18	38		5	13	36
	6	21	41		6	21	40
	7	26	43		7	33	44
	8	34	46		8	50	48
	9	43	49		9	64	51
	10	54	52		10	78	55
	11	67	54		11	90	59
	12	81	57		12	96	63
	13	91	60		13	99	66
	14	99	63		14	100	70
	15	100	65				
	16	100	>65				

Mathes 4 Form B (40./41. Schulwoche)

	Rohwert	Prozentrang	T-Wert		Rohwert	Prozentrang	T-Wert
Zahlen und Operationen	0	0	20	Größen und Messen	0	1	22
	1	0	22		1	2	24
	2	1	25		2	2	26
	3	1	27		3	4	28
	4	3	29		4	4	30
	5	4	31		5	5	33
	6	7	33		6	8	35
	7	9	35		7	10	37
	8	12	37		8	13	39
	9	16	39		9	20	41
	10	19	41		10	25	44
	11	25	43		11	31	46
	12	29	45		12	37	48
	13	36	47		13	46	50
	14	44	49		14	53	52
	15	52	51		15	64	55
	16	60	53		16	71	57
	17	68	55		17	82	59
	18	77	58		18	91	61
	19	83	60		19	96	63
	20	89	62		20	99	65
	21	95	64		21	100	68
	22	98	66		22	100	>68
	23	100	68				
	24	100	70				
	25	100	>70				
26	100	>70					

	Rohwert	Prozentrang	T-Wert		Rohwert	Prozentrang	T-Wert
Daten, Häufigkeit und Wahrscheinlichkeit	0	2	24	Raum und Form	0	1	18
	1	4	27		1	2	21
	2	6	30		2	2	25
	3	9	33		3	3	29
	4	11	35		4	5	33
	5	14	38		5	10	36
	6	16	41		6	17	40
	7	22	43		7	26	44
	8	28	46		8	40	48
	9	38	49		9	55	51
	10	50	52		10	69	55
	11	69	54		11	84	59
	12	82	57		12	93	63
	13	91	60		13	98	66
	14	98	63		14	100	70
	15	100	65				
	16	100	68				

Mathes 4 Form C (20./21. Schulwoche)

	Rohwert	Prozentrang	T-Wert		Rohwert	Prozentrang	T-Wert	
Zahlen und Operationen	0	0	<27	Größen und Messen	0	2	28	
	1	0	<27		1	5	30	
	2	2	27		2	7	32	
	3	4	29		3	11	34	
	4	6	31		4	16	37	
	5	10	33		5	21	39	
	6	14	35		6	28	41	
	7	17	37		7	32	43	
	8	21	39		8	38	45	
	9	25	41		9	46	48	
	10	30	43		10	51	50	
	11	35	45		11	58	52	
	12	40	47		12	66	54	
	13	48	49		13	77	56	
	14	55	51		14	86	58	
	15	62	53		15	91	61	
	16	69	55		16	95	63	
	17	76	57		17	98	65	
	18	84	59		18	100	67	
	19	90	61		19	100	69	
	20	95	63					
	21	97	65					
	22	99	67					
	23	99	68					
	24	100	70					
	25	100	72					
26	100	>72						

	Rohwert	Prozentrang	T-Wert		Rohwert	Prozentrang	T-Wert
Daten, Häufigkeit und Wahrscheinlichkeit	0	3	22	Raum und Form	0	0	18
	1	6	25		1	2	22
	2	8	28		2	3	26
	3	10	31		3	5	29
	4	12	34		4	7	33
	5	16	37		5	15	37
	6	20	41		6	24	41
	7	25	44		7	39	45
	8	33	47		8	55	48
	9	45	50		9	69	52
	10	59	53		10	82	56
	11	77	56		11	91	60
	12	91	59		12	96	64
	13	99	62		13	98	68
	14	100	65		14	100	71
	15	100	>65		15	100	75

Mathes 4 Form C (40./41. Schulwoche)

	Rohwert	Prozentrang	T-Wert		Rohwert	Prozentrang	T-Wert	
Zahlen und Operationen	0	0	<25	Größen und Messen	0	2	28	
	1	0	25		1	4	30	
	2	1	27		2	6	32	
	3	1	29		3	9	34	
	4	4	31		4	13	37	
	5	7	33		5	18	39	
	6	9	35		6	22	41	
	7	14	37		7	27	43	
	8	17	39		8	32	45	
	9	22	41		9	36	48	
	10	27	43		10	41	50	
	11	31	45		11	49	52	
	12	37	47		12	58	54	
	13	43	49		13	69	56	
	14	49	51		14	80	58	
	15	57	53		15	87	61	
	16	64	55		16	93	63	
	17	72	57		17	96	65	
	18	81	59		18	100	67	
	19	87	61		19	100	69	
	20	92	63					
	21	96	65					
	22	98	67					
	23	100	68					
	24	100	70					
	25	100	70					
26	100	74						

	Rohwert	Prozentrang	T-Wert		Rohwert	Prozentrang	T-Wert
Daten, Häufigkeit und Wahrscheinlichkeit	0	2	22	Raum und Form	0	0	18
	1	3	25		1	1	22
	2	5	28		2	2	26
	3	7	31		3	3	29
	4	10	34		4	5	33
	5	14	37		5	10	37
	6	17	41		6	17	41
	7	21	44		7	30	45
	8	29	47		8	46	48
	9	41	50		9	62	52
	10	56	53		10	77	56
	11	74	56		11	86	60
	12	89	59		12	93	64
	13	99	62		13	97	68
	14	100	65		14	100	71
	15	100	68		15	100	75

Mathes 4 Form D (20./21. Schulwoche)

	Rohwert	Prozentrang	T-Wert		Rohwert	Prozentrang	T-Wert
Zahlen und Operationen	0	0	22	Größen und Messen	0	1	29
	1	1	25		1	4	31
	2	3	27		2	7	33
	3	5	29		3	12	35
	4	8	31		4	17	37
	5	10	33		5	24	39
	6	13	35		6	29	41
	7	17	37		7	35	44
	8	22	40		8	42	46
	9	25	42		9	49	48
	10	29	44		10	55	50
	11	35	46		11	61	52
	12	42	48		12	69	54
	13	50	50		13	76	56
	14	61	52		14	85	58
	15	69	55		15	90	61
	16	79	57		16	93	63
	17	86	59		17	97	65
	18	92	61		18	99	67
	19	96	63		19	100	69
	20	98	65		20	100	>69
	21	99	67				
	22	100	69				
	23	100	72				
	24	100	>72				
25	100	>72					

	Rohwert	Prozentrang	T-Wert		Rohwert	Prozentrang	T-Wert
Daten, Häufigkeit und Wahrscheinlichkeit	0	1	20	Raum und Form	0	0	14
	1	4	23		1	1	17
	2	6	26		2	2	21
	3	8	28		3	3	24
	4	9	31		4	4	27
	5	12	34		5	6	30
	6	15	36		6	9	34
	7	19	39		7	12	37
	8	23	42		8	18	40
	9	28	45		9	28	43
	10	36	47		10	42	46
	11	46	50		11	56	50
	12	58	53		12	67	53
	13	74	55		13	79	56
	14	88	58		14	88	59
	15	98	61		15	96	63
	16	100	63		16	99	66
			17	100	69		
			18	100	72		

Mathes 4 Form D (40./41. Schulwoche)

	Rohwert	Prozentrang	T-Wert		Rohwert	Prozentrang	T-Wert
Zahlen und Operationen	0	0	<25	Größen und Messen	0	1	29
	1	0	25		1	2	31
	2	1	27		2	5	33
	3	3	29		3	9	35
	4	5	31		4	14	37
	5	8	33		5	19	39
	6	11	35		6	24	41
	7	13	37		7	28	44
	8	16	40		8	33	46
	9	19	42		9	38	48
	10	26	44		10	46	50
	11	32	46		11	52	52
	12	41	48		12	58	54
	13	48	50		13	67	56
	14	55	52		14	78	58
	15	62	55		15	84	61
	16	73	57		16	91	63
	17	81	59		17	95	65
	18	90	61		18	97	67
	19	95	63		19	100	69
	20	98	65		20	100	>69
	21	100	67				
	22	100	69				
	23	100	>69				
	24	100	>69				
25	100	>69					

	Rohwert	Prozentrang	T-Wert		Rohwert	Prozentrang	T-Wert
Daten, Häufigkeit und Wahrscheinlichkeit	0	1	20	Raum und Form	0	0	<17
	1	3	23		1	0	17
	2	4	26		2	1	21
	3	5	28		3	2	24
	4	7	31		4	2	27
	5	9	34		5	4	30
	6	12	36		6	5	34
	7	14	39		7	9	37
	8	19	42		8	15	40
	9	24	45		9	26	43
	10	29	47		10	38	46
	11	41	50		11	51	50
	12	54	53		12	62	53
	13	70	55		13	73	56
	14	87	58		14	85	59
	15	97	61		15	93	63
	16	100	63		16	97	66
			17	99	69		
			18	100	72		