

Aus der Klinik und Poliklinik für Kinder- und Jugendpsychiatrie, Psychosomatik
und Psychotherapie

Klinik der Ludwig-Maximilians-Universität München

Direktor: Prof. Dr. med. Gerd Schulte-Körne

Das kognitive Profil der Rechenstörung

Dissertation

zum Erwerb des Doktorgrades der Humanbiologie

an der Medizinischen Fakultät der

Ludwig-Maximilians-Universität zu München

vorgelegt von

Stefan Haberstroh

aus

Pfaffenhofen an der Ilm

2020

Mit Genehmigung der Medizinischen Fakultät
der Universität München

Berichterstatter: Prof. Dr. med. Gerd Schulte-Körne

Mitberichterstatter: Prof. Dr. Beate Sodian

Prof. Dr. Andrea Szelényi

Dekan: Prof. Dr. med. dent. Reinhard Hickel

Tag der mündlichen Prüfung: 17.03.2020

Inhaltsverzeichnis

1	Einleitung.....	5
1.1	Rechnen als neuronales Netzwerk domänenspezifischer und domänenübergreifender Fähigkeiten.....	7
1.2	Kognitive Befunde zur Rechenstörung	10
1.2.1	Domänenspezifische Fähigkeiten: Bisherige Forschungsergebnisse.....	10
1.2.2	Domänenübergreifende Fähigkeiten: Bisherige Forschungsergebnisse ..	14
1.3	Datenkodierung nach der Bottom-up-Methode	23
2	Zielsetzung	27
3	Methode.....	29
3.1	Einschlusskriterien	29
3.2	Suchstrategie	33
3.3	Kodierung.....	34
3.3.1	Deskriptive Studienangaben.....	34
3.3.2	Studienergebnisse.....	36
3.3.3	Ablauf der Kodierung.....	48
3.4	Auswertung	49
3.4.1	Ermittlung der Effektstärken.....	50
3.4.2	Multivariate Mehrebenen-Metaanalyse	51
3.4.3	Biasprüfung	53
3.4.4	Heterogenität.....	56
4	Ergebnisse.....	59
4.1	Suchverlauf und Studienbeschreibung	59
4.2	Ergebnisse der multivariaten Mehrebenen-Metaanalysen	68
4.2.1	Domänenspezifische Fähigkeiten: Ermittelte Effektstärken.....	73
4.2.2	Domänenübergreifende Fähigkeiten: Ermittelte Effektstärken	74
4.3	Heterogenität und Publikationsbias	78
4.4	Kognitives Profil der Rechenstörung	83
5	Diskussion	89
5.1	Domänenspezifische Fähigkeiten: Stabile Defizite	89
5.1.1	Rechnen.....	89

5.1.2	Mathematisches Schlussfolgern	90
5.1.3	Basiskompetenzen	91
5.2	Domänenübergreifende Fähigkeiten: Stabile Defizite	93
5.2.1	Arbeitsgedächtnis und Kurzzeitgedächtnis	93
5.2.2	Aufmerksamkeit.....	94
5.2.3	Verarbeitungsgeschwindigkeit	96
5.2.4	Raumvorstellung.....	98
5.2.5	Phonologische Verarbeitung.....	100
5.3	Einschränkungen	101
5.4	Implikationen für Theorie und Praxis	104
6	Zusammenfassung	109
	Literaturverzeichnis	111
	Abbildungsverzeichnis.....	131
	Tabellenverzeichnis.....	133
	Danksagung	135
	Eidesstattliche Versicherung	137

1 Einleitung

Ungefähr 3 bis 7 % aller Personen leiden unter einer Rechenstörung. Trotz normaler Intelligenz und regulärer Beschulung haben sie ausgeprägte Schwierigkeiten in Mathematik, die sich über das Alter hinweg als persistent erweisen (Fortes et al., 2016; Moll, Kunze, Neuhoff, Bruder & Schulte-Körne, 2014; Shalev, 2007). Bereits im Vorschulalter zeigen Kinder, bei denen im Grundschulalter eine Rechenstörung diagnostiziert wurde, deutliche Probleme beim einfachen Umgang mit Zahlen und Mengen (Butterworth, Varm & Laurillard, 2011; Desoete, Ceulemans, de Weerd & Pieters, 2012; Morgan, Farkas & Wu, 2009). Mit Eintritt in die Schule kommen betroffene Kinder mit den Grundrechenarten und größeren Zahlenräumen in Kontakt. Der Aufbau eines Faktenwissens, das es ihnen ermöglicht, die Ergebnisse einfacher Berechnungen direkt aus dem Gedächtnis abzurufen anstatt zu berechnen, gelingt ihnen nur mangelhaft (Geary, Hoard, Nugent & Bailey, 2012). Die zunehmende Komplexität des Lernstoffs führt letztlich zu Schwierigkeiten in allen weiterführenden mathematischen Anforderungen (z. B. schriftliches Dividieren, Textaufgaben) (Kaufmann & von Aster, 2012; Shalev, Manor & Gross-Tsur, 2005) und kann schließlich einen früheren Schulaustritt und geringeren Schulabschluss bedingen (Sälzer & Heine, 2016; Schulte-Körne, 2016).

Das Scheitern in Mathematik und die damit verbundenen Misserfolgserfahrungen in Schule und Alltag können sich im Weiteren negativ auf die psychische Gesundheit auswirken (Endlich, Dummert, Schneider & Schwenck, 2014). Internalisierende Symptome wie Ängste oder Traurigkeit sowie externalisierende Auffälligkeiten wie Aggressivität und Unruhe sind bei ca. 10 bis 40 % der Betroffenen zu beobachten (Fischbach, Schuchardt, Mähler & Hasselhorn, 2010; Kohn, Wyschkon & Esser, 2013; Willcutt et al., 2013). Langfristig besteht bei einer Rechenstörung schließlich die Gefahr eines geringeren sozioökonomischen Status. Betroffene weisen beispielsweise ein geringeres Jahreseinkommen sowie eine höhere Arbeitslosigkeits- und Kriminalitätsrate auf (Ritchie & Bates, 2013). Durch die mannigfaltigen Folgen einer Rechenstörung auf die Bildungs- und Berufskarriere sowie die weitere Lebensführung ist ihre Feststellung und korrekte Diagnose für Betroffene daher essenziell. Neben der subjektiven Gewissheit hinsichtlich der Ursachen der eigenen Mathematikprobleme

ermöglicht die Diagnose den Zugang zu Förder- und Unterstützungsmaßnahmen in Schule, Ausbildung und Studium sowie gegebenenfalls auch die Übernahme von Therapiekosten.

Um jedoch zwischen Personen mit und ohne Rechenstörung zu unterscheiden und damit eine Rechenstörung diagnostizieren zu können, muss über das zugrundeliegende kognitive Profil der Rechenstörung Klarheit herrschen (Lewis & Fisher, 2016). Ein solches Profil umfasst dabei diejenigen kognitiven Fähigkeiten, die in ihrer Ausprägung charakteristisch für eine Rechenstörung sind.

Dadurch ist es möglich, die Rechenstörung eindeutig mit psychologischen Tests zu identifizieren und sie von anderen Lernstörungen, insbesondere der Lese- und / oder Rechtschreibstörung, abzugrenzen (Cirino, Fuchs, Elias, Powell & Schumacher, 2015; Nolan, Hammeke & Barkley, 1983).

In mehreren Übersichtsarbeiten wurden bisher verschiedenste kognitive Defizite bei einer Rechenstörung beschrieben (z. B. Cowan & Powell, 2014; Karagiannakis, Baccaglini-Frank & Papadatos, 2014; Peng, Cuicui & Jessica, 2018; Shin & Bryant, 2015). Hierbei zeigte sich, dass eine Rechenstörung nicht nur mit Defiziten in spezifischen mathematischen Fähigkeiten (z. B. Faktenwissen) einherging, sondern auch mit Defiziten in allgemeinen nicht-mathematischen Fähigkeiten (z. B. Arbeitsgedächtnis) assoziiert war (Henik, Rubinsten & Ashkenazi, 2011). Mathematische Fähigkeiten werden in diesem Zusammenhang auch als *domänenspezifische Fähigkeiten* bezeichnet, wohingegen nicht-mathematische Fähigkeiten als *domänenübergreifende Fähigkeiten* gelten. Ein eindeutiges kognitives Profil der Rechenstörung konnte jedoch nicht ermittelt werden. Zum einen wurde festgestellt, dass die Schwere der Defizite in jeder Fähigkeit davon abhängt, mit welchen Testaufgaben sie erfasst werden (Hirsch, Lambert, Coppens & Moeller, 2018). Zum anderen konzentrieren sich die meisten Studien ausschließlich auf Defizite in der Richtigkeit (z. B. Anzahl gelöster Aufgaben), während weiterhin unklar ist, wie schwer und vielfältig die Defizite in der benötigten Zeit (z. B. mittlere benötigte Zeit, um eine Aufgabe zu lösen) sind und ob die benötigte Zeit somit zur Diagnostik verwendet werden kann. Insgesamt bleibt daher festzuhalten, dass weiterhin kein konsistentes kognitives Profil der Rechenstörung vorliegt (Geary, 2010; Träff, Olsson, Östergren & Skagerlund, 2017).

Ziel dieser Arbeit war es demzufolge, das kognitive Profil der Rechenstörung zu ermitteln. Hierzu wurde ein systematisches Review mit metaanalytischer Auswertung angewandt, um (1) diejenigen domänenspezifischen und domänenübergreifenden Fähigkeiten sowie ihre zugrundeliegenden Testaufgaben zu ermitteln, die am besten zwischen Personen mit und ohne Rechenstörung unterschieden und damit als stabile Defizite galten sowie (2) zu vergleichen, ob die Schwere der Defizite in den Fähigkeiten und Testaufgaben sich änderte, wenn die Richtigkeit oder die benötigte Zeit erfasst wurde.

Aus praktischer Sicht hilft diese Arbeit somit sämtlichen Fachkräften (z. B. Psychologen/innen, Kinder- und Jugendpsychiater/innen, Lehrer/innen) und Forscher/innen, die mit Personen mit Rechenstörung zu tun haben, sich in ihrer Arbeit auf die Fähigkeiten und Testaufgaben zu fokussieren, die nachweislich am besten zur Diagnose einer Rechenstörung geeignet sind. Aus theoretischer Sicht trägt diese Arbeit zu einem besseren Verständnis der Heterogenität und Komplexität der Rechenstörung bei. So können beispielsweise neue Tests entwickelt werden, die domänenspezifische als auch domänenübergreifende Fähigkeiten bei der Diagnosestellung berücksichtigen. Von methodischer Seite wurde in dieser Arbeit überdies ein neuer Ansatz bei der Datenkodierung und -auswertung sowie beim Umgang mit Heterogenität verwendet, der als Blaupause für zukünftige metaanalytische Profilstudien dient.

1.1 Rechnen als neuronales Netzwerk domänenspezifischer und domänenübergreifender Fähigkeiten

Während das Wissen um spezifische Defizite in Mathematik bereits länger existierte und die „Rechenstörung“ erstmals im ICD-9 als eigenständige Störungskategorie „Umschriebene Rechenschwäche“ geführt wurde (World Health Organization, 1978), so fand erst in den letzten 20 Jahren eine deutliche Intensivierung der Forschungsaktivität in diesem Gebiet statt (Lewis & Fisher, 2016). Zum einen vervierfachte sich nahezu die Anzahl an Publikationen, die sich spezifisch mit der Rechenstörung befassten (Datenbanken: *MEDLINE*, *PsycInfo* und *PSYINDEX*; Suchbegriffe aus Tabelle 1). Zum anderen fand eine genauere Betrachtung des Rechnens im Allgemeinen und der normalen Entwicklung der Rechenfähigkeiten statt. Insbesondere die neurobiologische Forschung unter Verwendung bildgebender Verfahren konnte hierbei aufzeigen, dass beim

Rechnen ein komplexes neuronales Netzwerk domänenspezifischer sowie domänenübergreifender Fähigkeiten beteiligt ist, welches sämtliche Gehirnlappen sowie vielfältige subkortikale Regionen miteinbezieht (für eine Übersicht: Peters & de Smedt, 2018). Abbildung 1 skizziert die beim Rechnen beteiligten Gehirnareale.

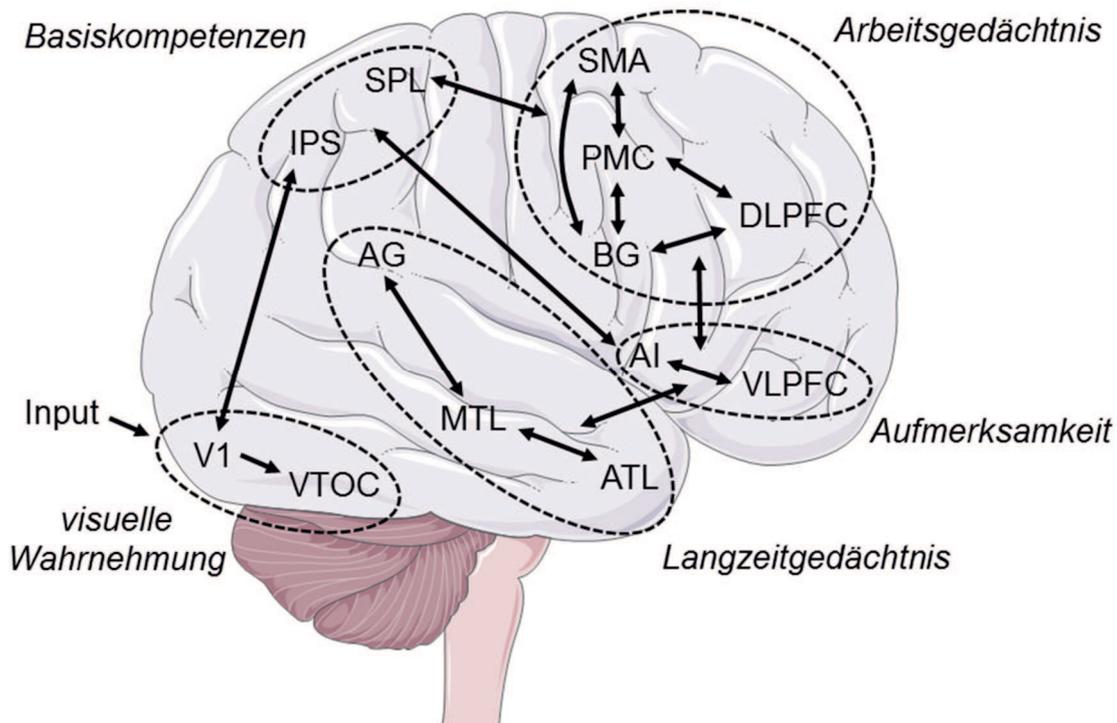


Abbildung 1. Neuronales Netzwerk beim Rechnen.

V1 = primäre Sehrinde / primary visual cortex; VTOC = ventraler temporo-okzipitaler Cortex / ventral temporal-occipital cortex; IPS = Sulcus intraparietalis / intraparietal sulcus; SPL = Lobulus parietalis superior / superior parietal lobe; SMA = supplementär-motorische Rinde / supplementary motor area; PMC = prämotorische Rinde / premotor Cortex; BG = Basalganglien / basal ganglia; DLPFC = dorsolateraler Präfrontalcortex / dorsolateral prefrontal cortex; AG = Gyrus angularis / angular gyrus; MTL = medialer Temporallappen / medial temporal lobe; ATL = anteriorer Temporallappen / anterior temporal lobe; AI = Inselrinde / anterior insula; VLPFC = ventrolateraler Präfrontalcortex / ventrolateral prefrontal cortex.

Abbildung „Cerveau Tronc“ von Servier Medical Art

(https://smart.servier.com/smart_image/brain-12/). Lizenz: CC BY 3.0.

Bezeichnungen und Zuordnung der Gehirnareale nach Kucian (2016).

Der Input erfolgt hierbei in den meisten Fällen visuell, indem arabische Zahlen (z. B. 1, 2, 3) vorgegeben werden, deren Dekodierung im visuellen Cortex (V1, VTOC) stattfindet. Die parietalen Areale IPS und SPL weisen den so dekodierten Zahlen schließlich ihre Bedeutung zu, indem sie eine visuelle Repräsentation der hinter den Zahlen stehenden Mengen generieren (z. B. $2 = \bullet\bullet$). Diese parietale Region, auch bezeichnet als *Zahlensinn* oder *Number Sense* (Arsalidou & Taylor, 2011), ist zentral für die Zahlen- und Mengenverarbeitung. Die gewonnenen Informationen werden für mehrere Sekunden im Arbeitsgedächtnis gehalten (SMA, PMC, DLPFC, subkortikal: BG) und je nach Aufgabenstellung entsprechend bearbeitet. Ebenso, falls notwendig, können weitere zur Aufgabenbearbeitung notwendige Informationen aus dem Langzeitgedächtnis abgerufen werden. Insbesondere der Gyrus angularis (GA) im parietalen Cortex spielt dabei für den Zugriff auf das Faktenwissen eine entscheidende Rolle (Grabner et al., 2009). Generell bedeutsam für den Zugriff auf das Langzeitgedächtnis sind überdies der mediale und anteriore Temporallappen (MTL, ATL), die subkortikal den Hippocampus sowie die Amygdala umfassen. Während dieses gesamten Prozesses findet präfrontal (AI, VLPFC) die Kontrolle der Aufmerksamkeit statt, die eine zielgerichtete Aufgabenbearbeitung ermöglicht (Kucian, 2016; Menon, 2014).

In Bezug auf die Rechenstörung existiert nur eine geringe Anzahl an fMRT-Studien, die das Aktivierungsmuster der Areale bei Personen mit und ohne Rechenstörung verglichen (de Smedt, Noël, Gilmore & Ansari, 2013). Diese Studien unterscheiden sich zudem im Stichprobenalter sowie den bei der Untersuchung verwendeten Testaufgaben (z. B. einfache Aufgaben zur Mengenverarbeitung, Grundrechenarten), sodass keine klaren Aussagen zum neuronalen Profil einer Rechenstörung gemacht werden konnten (Peters & de Smedt, 2018). Bezogen auf die domänenspezifischen Fähigkeiten wurde bei der Rechenstörung zumeist eine geringere (Ashkenazi, Rosenberg-Lee, Tenison & Menon, 2012; Kucian, Loenneker, Martin & von Aster, 2011), jedoch bisweilen auch eine höhere Aktivierung in den korrespondierenden parietalen Regionen beobachtet (Kaufmann, Wood, Rubinsten & Henik, 2011; Rosenberg-Lee et al., 2014), wenn betroffene Personen Testaufgaben zur Zahlen- und Mengenverarbeitung bearbeiteten. Hinsichtlich der domänenübergreifenden Fähigkeiten zeigten sich im Vergleich zu Personen ohne Rechenstörung abweichende Akti-

vitätsmuster, insbesondere für das Arbeitsgedächtnis (Rotzer et al., 2009) und der Aufmerksamkeit (Davis et al., 2009; Rosenberg-Lee et al., 2014) in den jeweils korrespondierenden Gehirnregionen. Dabei war bei Kindern mit Rechenstörung eine deutlich höhere Aktivierung festzustellen, die als Kompensationsmechanismus gesehen wurde. Bei der Bearbeitung von Aufgaben zur Zahlen- und Mengenverarbeitung griffen diese Kinder stärker auf das Arbeitsgedächtnis und die Aufmerksamkeit zurück, da die Herausbildung der Regionen rund um den IPS, die zentral für die Zahlen- und Mengenverarbeitung sind, verlangsamt stattfand (McCaskey et al., 2018). Ihr Bearbeitungsstil verbraucht damit mehr kognitive Ressourcen als bei Kindern ohne Rechenstörung. Diese hingegen können die Bedeutung von Zahlen und Mengen leichter erfassen und vom Gedächtnis abrufen, wodurch mehr Ressourcen für die Lösung der Aufgaben zur Verfügung stehen.

Für das kognitive Profil der Rechenstörung und im weiteren Sinne ihrer Diagnostik ist daher festzuhalten, dass Rechnen auf einem neuronalen Netzwerk fußt, welches domänenspezifische wie domänenübergreifende Fähigkeiten umfasst und bei Personen mit Rechenstörung ein anderes Aktivierungsmuster aufweist (Jolles et al., 2016). Diese Beobachtung bestätigt damit auch die Ergebnisse von Studien, die sich mit den kognitiven Unterschieden von Personen mit und ohne Rechenstörung befassten. Diese wiesen anhand von Verhaltensdaten und psychometrischen Tests frühzeitig darauf hin, dass eine Rechenstörung nicht nur mit kognitiven Defiziten in domänenspezifischen Fähigkeiten einhergeht, sondern auch Minderleistungen in domänenübergreifenden Fähigkeiten umfasst (Fias, Menon, & Szucs, 2013).

1.2 Kognitive Befunde zur Rechenstörung

1.2.1 Domänenspezifische Fähigkeiten: Bisherige Forschungsergebnisse

Das Leitsymptom einer Rechenstörung ist, entsprechend der Störungsbezeichnung, eine anhaltende Schwierigkeit beim Lösen von Rechenaufgaben (Kuhn, 2015). Das ICD-10, erstellt in einer Zeit, in der die Forschung zur Rechenstörung noch in ihren Anfängen war, listet daher auch als einziges Symptom einer Rechenstörung Defizite in den Grundrechenarten auf (World Health Organization, 1992). Mit der Zunahme der Forschungsaktivität zur Rechenstörung in den letzten zwei Jahrzehnten wurde jedoch deutlich, dass Schwierigkei-

ten im Rechnen auch durch Defizite in der zugrundeliegenden Zahlen- und Mengenverarbeitung bedingt sein können (Geary, 2010). Selbst das Rechnen an sich kann, je nach Komplexität der Testaufgabe, anders stattfinden. So werden die Ergebnisse einfacher Rechenaufgaben beispielsweise direkt aus dem Langzeitgedächtnis abgerufen (d. h. Faktenwissen), ohne sie überhaupt zu berechnen (Grabner et al., 2009). Diese zunehmende Ausdifferenzierung der spezifisch mathematischen Defizite bei einer Rechenstörung floss folglich in die jeweils aktuelle Fassung der beiden gängigen Klassifikationsschemata, DSM-5 (American Psychiatric Association, 2013) und ICD-11 (World Health Organization, 2018), ein. So beschreiben beide eine Rechenstörung als domänenspezifische Defizite im Rechnen und Faktenwissen (d. h. Einmaleins) als auch in den Basiskompetenzen (d. h. Zahlen- und Mengenverarbeitung) und dem mathematischen Schlussfolgern. Während das Faktenwissen und Rechnen klar definierte Fähigkeiten sind und daher in Studien mit ähnlichen Testaufgaben erfasst wurden, so existieren für die Basiskompetenzen und das mathematische Schlussfolgern verschiedene Annahmen darüber, welche Anforderungen sie umfassen und mit welchen Testaufgaben sie gemessen werden (Berch, 2005; Gersten, Jordan & Flojo, 2005).

1.2.1.1 Rechnen

Zum *Rechnen* zählen die Fähigkeiten *Grundrechenarten* und *Faktenwissen*, die mit ähnlichen Testaufgaben erfasst werden, sich jedoch zentral in ihrer Methode zur Aufgabenbearbeitung unterscheiden. Auf das Faktenwissen wird zurückgegriffen, wenn einfache arithmetische Aufgaben (z. B. Einmaleins) dadurch gelöst werden, indem das Ergebnis direkt aus dem Gedächtnis abgerufen wird, anstatt es eigens zu berechnen. Rechnen hingegen findet bei arithmetischen Aufgaben Verwendung, die komplexer sind und mehrere Schritte zur Lösung benötigen und somit auf die korrekte Anwendung und das richtige Verständnis der Grundrechenarten fußen (Busch, Schmidt & Grube, 2015). Beide Fähigkeiten sind von der Operationalisierung her leicht durch Rechenaufgaben zu erfassen (z. B. $3 * 4$ für Faktenwissen; $818 + 104$ für Rechnen) und Defizite in diesen Fähigkeiten gelten als Leitsymptom einer Rechenstörung (Jordan & Hanich, 2003). Auffallend ist jedoch, dass Testaufgaben für gewöhnlich nur die Richtigkeit messen, wohingegen die benötigte Zeit zwar oft berichtet, allerdings nicht hinsichtlich ihrer Validität, zwischen Personen mit und ohne Rechenstörung zu

unterscheiden, analysiert und diskutiert wurde (z. B. Hanich, Jordan, Kaplan & Dick, 2001; Mammarella, Caviola, Cornoldi & Lucangeli, 2013). Obwohl anzunehmen ist, dass Personen mit Rechenstörung aufgrund ihres mangelnden Verständnisses von Rechenoperationen sowie dem defizitären Abruf von Fakten aus dem Gedächtnis länger zu Bearbeitung dieser Aufgaben benötigen, bleibt ungeklärt, inwiefern die benötigte Zeit im Rahmen der Diagnosestellung berücksichtigt werden kann.

1.2.1.2 Mathematisches Schlussfolgern

Mathematisches Schlussfolgern ist grundsätzlich zu verstehen als die Anwendung mathematischen Wissens zur Lösung nicht-vertrauter mathematischer Problemstellungen (Lindquist, Philpot, Mullis & Cotter, 2017). Die Testaufgaben hierzu können sich jedoch stark unterscheiden. Kaufmann und von Aster (2012) und Casey et al. (2015) betrachteten zum Beispiel mathematisches Schlussfolgern als die korrekte Anwendung des gesamten erlernten mathematischen Wissens, um komplexe Probleme zu lösen. Bei dieser Herangehensweise sind vor allem Text- und Geometrieaufgaben sowie Testaufgaben, bei denen Tabellen oder Grafiken interpretiert werden müssen, typisch (Proctor, 2012). Zhang et al. (2017) hingegen sah mathematisches Schlussfolgern als die Fähigkeit, einfache, jedoch nicht-vertraute Probleme durch Schlussfolgern zu lösen. In diesem Fall sind Zahlenfolgen als Testaufgaben üblich, bei denen bei einer vorgegebenen Sequenz von Zahlen ein Muster durch korrekte Anwendung der Grundrechenarten erkannt werden muss (z. B. 2 4 6 _ 10). Obwohl beide diese Arten von Testaufgaben hoch korrelieren und damit höchstwahrscheinlich die gleiche Fähigkeit erfassen (Lee, Ng, Bull, Pe & Ho, 2011; Taub, Keith, Floyd & McGrew, 2008; Zhang et al., 2017), so ist weiterhin unklar, ob sie auch gleich gut zwischen Personen mit und ohne Rechenstörung unterscheiden. Studien, die beide Testaufgaben bei Personen mit Rechenstörung durchführten, existieren nicht. Zusätzlich ist zu berücksichtigen, dass Zahlenfolgen als Testaufgabe ebenso hoch mit der allgemeinen Fähigkeit zum logischen Schlussfolgern korreliert (Englisch: Fluid Reasoning), was wiederum als gängiges Maß zur Messung der nonverbalen Intelligenzleistung gilt (W. J. Schneider & McGrew, 2018). Demnach ist nicht auszuschließen, dass Defizite bei Zahlenfolgen auch von der allgemeinen Intelligenz der Testpersonen moderiert werden und nicht ausschließlich auf das Vorhandensein einer Rechenstörung zurückzuführen

sind. Klärungsbedarf hinsichtlich der korrekten Operationalisierung des mathematischen Schlussfolgerns besteht insofern, da gängige Tests im deutschsprachigen Raum zur Diagnostik einer Rechenstörung sowohl Zahlenfolgen (z. B. BADYS 1-4+ (Meridian, Meridian & Schardt, 2015), CODY-M 2-4 (Kuhn, Schwenk, Raddatz, Dobel & Holling, 2017), HRT 1-4 (Haffner, Baro, Parzer & Resch, 2005)) als auch Textaufgaben (z. B. DEMAT 3+ (Roick, Göllitz & Hasselhorn, 2018), ERT 4+ (Schaupp, Lenart & Holzer, 2010), ZAREKI-R (von Aster, Weinhold-Zulauf & Horn, 2006)) beinhalten.

1.2.1.3 Basiskompetenzen

Die *Basiskompetenzen* sind im Allgemeinen als die grundlegende Fähigkeit zur Zahlen- und Mengenverarbeitung zu verstehen (Butterworth et al., 2011), die für die spätere Leistung in Mathematik als äußerst prädiktiv gelten (Desoete et al., 2012; Geary et al., 2012; Jordan, Kaplan, Ramineni & Locuniak, 2009). Sie können mit verschiedenen Testaufgaben erfasst werden (für einen Überblick siehe: Hirsch et al., 2018; Kaufmann & von Aster, 2012), die jedoch unterschiedlich gut zwischen Personen mit und ohne Rechenstörung unterscheiden. Beispielsweise waren Kinder mit Rechenstörung signifikant langsamer bei Testaufgaben, in denen sie die Größere zweier Zahlen bestimmen mussten (z. B. $3 > 1$), als bei vergleichbaren Testaufgaben, bei denen sie statt zwei Zahlen zwei Mengen nach ihrer Größe verglichen (z. B. $\bullet\bullet\bullet > \bullet$) (Schwenk et al., 2017). Ebenso korrelierte die Leistung in Mathematik höher mit der Leistung in Testaufgaben zum Zahlenvergleich als zum Mengenvergleich (M. Schneider et al., 2017). Hinsichtlich der Simultanerfassung von Mengen, also dem schnellen Erfassen von Punktmengen (z. B. $\bullet\bullet\bullet = 3$), ergaben sich bisher widersprüchliche Ergebnisse und nicht alle Personen mit Rechenstörung benötigten mehr Zeit oder zeigten eine geringere Richtigkeit (Landerl, 2013; Moll, Göbel & Snowling, 2015; Szűcs, Devine, Soltesz, Nobes & Gabriel, 2013). Im Gegensatz hierzu waren Defizite im Zahlenstrahl (z. B. Position einer Zahl auf dem Zahlenstrahl schätzen) (Cowan & Powell, 2014; Geary, Hoard, Nugent & Byrd-Craven, 2008) und beim Transkodieren (z. B. gehörte Zahl ausschreiben) (Iglesias-Sarmiento & Deaño, 2016; R. Moura et al., 2015; Raddatz, Kuhn, Holling, Moll & Dobel, 2017) bei Personen mit Rechenstörung mehrheitlich stabil. Für die widersprüchlichen Ergebnisse zur Rolle der Mengenverarbeitung im Vergleich zu den stabilen Defiziten in der Zahlenverarbeitung bei der Rechenstörung gibt

es zwei Annahmen: Studien, die von Defiziten sowohl in Testaufgaben mit Mengen (z. B. Mengenvergleich, Simultanerfassung) als auch in Testaufgaben mit Zahlen (z. B. Zahlenvergleich, Zahlenstrahl) berichten, gehen davon aus, dass eine defizitäre Repräsentation von Mengen konsequenterweise zu Schwierigkeiten bei der Zahlenverarbeitung führt. Demgegenüber stehen Studien, die ausschließlich Schwierigkeiten bei der Verarbeitung von Zahlen ermitteln und annehmen, dass die Mengenverarbeitung zwar intakt ist, jedoch der Zugriff darauf bei Vorgabe von Zahlen eingeschränkt bzw. lediglich die Verknüpfung zwischen Zahl und Menge gestört ist (Bartelet, Ansari, Vaessen & Blomert, 2014; de Smedt & Gilmore, 2011; Olsson, Östergren & Träff, 2016; Skagerlund & Träff, 2016; Träff et al., 2017). Inwiefern eine Rechenstörung somit tatsächlich mit einer defizitären Mengenverarbeitung einhergeht, die im Rahmen der Diagnostik erfasst werden soll, blieb ungeklärt. Übergreifend lassen die Resultate zu den einzelnen Testaufgaben jedoch mutmaßen, dass bei einer Rechenstörung Defizite in der Zahlenverarbeitung größer und stabiler sind als in der Mengenverarbeitung.

1.2.2 Domänenübergreifende Fähigkeiten: Bisherige Forschungsergebnisse

Eine Rechenstörung betrifft nicht nur isoliert die mathematischen Fähigkeiten, sondern sie geht, wie eingangs erwähnt, mit einer Vielzahl von Auffälligkeiten in den allgemeinen nicht-mathematischen Fähigkeiten, den sogenannten domänenübergreifenden Fähigkeiten, einher. Im besonderen Fokus der Forschung standen hierbei in den letzten Jahren die exekutiven Funktionen (d. h. Arbeitsgedächtnis, Aufmerksamkeit). Zum einen ging es um die Frage, welche Rolle das Arbeitsgedächtnis und ihre einzelnen Komponenten beim Rechnen und bei der Zahlen- und Mengenverarbeitung spielen (McDonald & Berg, 2018; Watson & Gable, 2013). Dies war auch dahingehend von Interesse, da lange Zeit versucht wurde, Schwierigkeiten in der Mathematik mittels eines Arbeitsgedächtnistrainings zu beheben, was letztlich jedoch nicht gelang (Kassai, Futo, Demetrovics & Takacs, 2019; Sala & Gobet, 2017). Zum anderen wurde überprüft, ob sich bei einer Rechenstörung charakteristische Defizite in der Aufmerksamkeitskontrolle zeigen, da bis zu 40 % der Personen mit Rechenstörung Aufmerksamkeitsdefizite aufweisen, wenngleich nicht in allen Fällen eine Aufmerksamkeitsdefizit-/Hyperaktivitätsstörung (ADHS) vorliegt (Fischbach et al.,

2010; Fortes et al., 2016; Gross-Tsur, Manor & Shalev, 1996). Andere Studien konzentrierten sich hingegen auf basale kognitive Fähigkeiten wie die Verarbeitungsgeschwindigkeit, die für das schnelle Dekodieren und Weiterverarbeiten eingehender Informationen zuständig ist. Hierbei war die Annahme, dass eine Rechenstörung bereits auf eine defizitäre Verarbeitung einfacher (numerischer) Stimuli zurückzuführen ist, die in weiterer Folge zu einem mangelhaften Aufbau der Rechenkompetenz führt (Niileksela & Reynolds, 2014; Proctor, 2012). Zwei weitere Forschungsstränge entstanden zur Raumvorstellung (d. h. visuelle Wahrnehmung und mentale Manipulation von komplexen Objekten im Raum) sowie zur phonologischen Verarbeitung. Bei der Raumvorstellung ging es um die Frage, inwiefern Defizite in der Mengenvorstellung auf Schwierigkeiten in der Raumvorstellung zurückzuführen sind (Mussolin, Martin & Schiltz, 2011). Die phonologische Verarbeitung wurde vor allem in Zusammenhang mit dem Aufbau des Faktenwissens untersucht. Hierbei werden gemäß Dehaene (1992) die Fakten (z. B. $2 * 3 = 6$) bzw. die zugrundeliegenden Zahlwörter (d. h. zwei * drei = sechs) in einen auditiv-verbale Code transformiert und im semantischen Gedächtnis gespeichert. Zusätzlich weist die Rechenstörung eine hohe Komorbidität mit der Lesestörung auf (Moll et al., 2014), sodass die Annahme bestand, dass bei Personen mit Rechenstörung ebenso charakteristische Defizite in der phonologischen Verarbeitung vorliegen (Hecht, Torgesen, Wagner & Rashotte, 2001).

1.2.2.1 Arbeitsgedächtnis und Kurzzeitgedächtnis

Hinsichtlich des *Arbeitsgedächtnisses* untersuchten die meisten Studien Baddeleys Arbeitsgedächtnismodell (Baddeley, 2012; Baddeley & Hitch, 1974). Das Modell geht von einer übergeordneten Komponente, der sogenannten *zentralen Exekutive* aus, welche für die Überwachung und Regulierung komplexer kognitiver Prozesse zuständig ist. Der zentralen Exekutive waren ursprünglich zwei Komponenten, der visuell-räumliche Notizblock und die phonologische Schleife, untergeordnet, welche visuell-räumliche bzw. sprachliche Informationen kurzfristig speichern. Jüngere Forschung bewertet hingegen ausschließlich die zentrale Exekutive als Arbeitsgedächtnis, wohingegen der visuell-räumliche Notizblock und die phonologische Schleife als *visuell-räumliches bzw. auditives Kurzzeitgedächtnis* gelten. Die Trennung beruht darauf, dass das Arbeitsgedächtnis nicht nur für die ausschließliche Speicherung von Informationen, son-

dern auch für die Überwachung und Regulierung ihrer Weiterverarbeitung zuständig ist (Swanson, 2012).

Bei Personen mit Rechenstörung wurden in mehreren Übersichtsarbeiten konsistente Defizite im Arbeitsgedächtnis beschrieben, während die Ergebnisse zum Kurzzeitgedächtnis widersprüchlich waren (David, 2012; Johnson, Humphrey, Mellard, Woods & Swanson, 2010; Peng et al., 2018; Peng & Fuchs, 2014; Swanson & Jerman, 2006). Beispielsweise berichteten Szűcs et al. (2013) ausschließlich über Defizite im visuell-räumlichen Kurzzeitgedächtnis, wohingegen Cowan und Powell (2014) zusätzlich auch Defizite im auditiven Kurzzeitgedächtnis bei Personen mit Rechenstörung feststellten. Beide Studien unterscheiden sich vor allem in ihrem Umgang mit komorbiden Fällen einer Rechen- und Lesestörung, die bei Szűcs et al. (2013) ausgeschlossen und bei Cowan und Powell (2014) eingeschlossen waren. Eine Studie von Schuchardt, Mähler und Hasselhorn (2008) wies in diesem Zusammenhang bereits auf mögliche störungsspezifische Defizite im Kurzzeitgedächtnis hin. Darin zeigten lediglich Personen mit Rechen- und Lesestörung sowohl im auditiven als auch visuell-räumlichen Kurzzeitgedächtnis Defizite. Personen mit einer isolierten Lesestörung hatten hingegen ausschließlich Schwierigkeiten im auditiven Kurzzeitgedächtnis und demgegenüber Personen mit einer isolierten Rechenstörung nur im visuell-räumlichen Kurzzeitgedächtnis.

Hinsichtlich der Testaufgaben liegen für das Kurzzeitgedächtnis und insbesondere für das Arbeitsgedächtnis unterschiedliche Befunde vor (Aben, Stapert & Blokland, 2012). In einer Studie von Cirino et al. (2015) zeigten Personen mit Rechenstörung beispielsweise größere Defizite beim Lösen sogenannter komplexer Spannen als bei rückwärtigen Spannen. Eine Spanne meint dabei eine bestimmte Anzahl an Symbolen oder Inhalten, die erinnert werden muss und die sukzessive ansteigt (z. B. eine Reihenfolge von Wörtern). Bei einer komplexen Spanne bearbeiten Testpersonen simple Aufgaben (z. B. einfache Fragen mit Ja/Nein beantworten: „Ist Gras grün?“), während sie sich gleichzeitig die korrekte Darbietungsreihenfolge spezifischer Aufgabeninhalte merken (z. B. das letzte Wort jeder gestellten Frage: „grün“). Das Prinzip rückwärtiger Spannen basiert hingegen darauf, eine vorwärts vorgegebene Reihenfolge an Inhalten in umgekehrter Reihenfolge wiederzugeben. Beide Aufgabentypen haben gemein,

dass das Erinnern von Informationen gepaart ist mit einer zusätzlichen kognitiven Leistung, die an diese Informationen geknüpft ist (z. B. Fragen beantworten, Reihenfolge umkehren). Die gefundenen Unterschiede zwischen beiden Aufgabentypen bei Personen mit Rechenstörungen wurden von Mähler und Schuchardt (2016) und Desoete et al. (2012) jedoch nicht bestätigt, sodass unklar bleibt, ob die unterschiedlichen Operationalisierungen dieser Aufgabentypen in einem Zusammenhang mit der Rechenstörung stehen.

1.2.2.2 Aufmerksamkeit

Aufmerksamkeit ist auch zu verstehen als die Fähigkeit, die kognitive Leistung in Bezug zu sich wechselnden Aufgabenstellung zu überwachen, anzupassen und zu regulieren (van der Sluis, de Jong & van der Leij, 2004). Gemäß Miyake, Friedman, Emerson, Witzki und Howerter (2000) setzt sich die Kontrolle der Aufmerksamkeit aus den folgenden 3 Komponenten zusammen: *Inhibition* (d. h. dominante, jedoch falsche Antworttendenzen unterdrücken), *Shifting* (d. h. flexibel zwischen Aufgabenstellungen wechseln) und *Updating* (d. h. neue aufgabenbezogenen Informationen mitberücksichtigen).

Für Personen mit Rechenstörung ergaben sich bisher widersprüchliche Ergebnisse hinsichtlich dieser 3 Komponenten. Übergreifend für alle Komponenten berichtete zum Beispiel Peng et al. (2018) kleine und Johnson et al. (2010) große Defizite. Bezüglich der Inhibition konnten in mehreren Studien Unterschiede in der benötigten Zeit zwischen Personen mit und ohne Rechenstörung festgestellt werden, wenn die Testaufgabe auf dem Stroop-Effekt basierte (d. h. Merkmal eines Objekts benennen und dabei ein dominantes Merkmal des oder eines anderen Objekts bewusst unterdrücken) (Peng, Congying, Beilei & Sha, 2012; Skagerlund & Träff, 2016; Szűcs et al., 2013; Wang, Tasi & Yang, 2012). Eine mögliche Erklärung für den Zusammenhang zwischen Inhibition und Rechenstörung lieferten Geary, Hamson und Hoard (2000). Diese nahmen an, dass Personen mit Rechenstörung Schwierigkeiten im Faktenwissen haben, da sie nicht in der Lage sind, ähnliche aber dennoch falsche Lösungen zu unterdrücken. Beispielsweise triggert die Aufgabe $3 * 4$ neben der korrekten Lösung 12 auch ähnliche Lösungen wie 8 (d. h. $2 * 4$), 9 (d. h. $3 * 3$) oder 16 (d. h. $4 * 4$), die nicht ausgeblendet werden können und von der korrekten Lösung ablenken. Die berichteten Defizite in der Inhibition und dabei insbesondere bei der

Testaufgabe Stroop-Effekt wurden jedoch in Studien von Censabella und Noël (2008) sowie van der Sluis et al. (2004) nicht festgestellt. Auch bei Verwendung anderer Testaufgaben zur Erfassung der Inhibition (z. B. Go/Nogo) oder wenn statt der benötigten Zeit die Richtigkeit erfasst wurde, zeigten sich entweder nur kleine oder keine Effekte (Censabella & Noël, 2008; Swanson, 2012; van der Sluis et al., 2004; Willburger, Fussenegger, Moll, Wood & Landerl, 2008). Eine endgültige Bewertung der Rolle der Inhibition für die Rechenstörung ist somit noch ausstehend. Für die anderen beiden Komponenten, Shifting und Updating, ist die Studienlage zur Rechenstörung deutlich geringer und die Ergebnisse nicht eindeutig genug, um Schlussfolgerungen daraus ziehen zu können (Peng et al., 2012; Szűcs et al., 2013; van der Sluis et al., 2004; Willcutt et al., 2013).

1.2.2.3 Verarbeitungsgeschwindigkeit

Der *Verarbeitungsgeschwindigkeit* ist beim Lösen einer Vielzahl unterschiedlicher Aufgaben relevant. Grundsätzlich meint die Verarbeitungsgeschwindigkeit, schnell Informationen zu dekodieren und auf Basis dieser Informationen einfache und sich wiederholende Aufgaben schnellstmöglich zu lösen (Conway, Cowan, Bunting, Theriault & Minkoff, 2002). Im Fokus steht dabei insbesondere Perceptual Speed bzw. die *visuelle Verarbeitungsgeschwindigkeit*, welche die Geschwindigkeit darstellt, mit der vorgegebene visuelle Stimuli (z. B. Symbole) in einem visuellen Feld gesucht und/oder visuelle Stimuli hinsichtlich ihrer Ähnlichkeit verglichen werden (W. J. Schneider & McGrew, 2018). Die Operationalisierung von Testaufgaben beruht folglich darauf, ausschließlich visuelle Stimuli zueinander in Bezug zu setzen, um eine Aufgabe zu lösen. Weitere Komponenten der Verarbeitungsgeschwindigkeit sind gemäß Salthouse (2000) beispielsweise die *Reaktions- und Entscheidungsgeschwindigkeit* sowie die *psychomotorische Geschwindigkeit*. Bei der Reaktions- und Entscheidungsgeschwindigkeit steht im Vordergrund, auf kognitiv wenig komplexe Stimuli jedweder Art (z. B. Hören eines Tones, Erscheinen eines Punktes auf einem Bildschirm) schnellstmöglich zu reagieren (z. B. Knopf drücken) oder schnelle Entscheidungen hinsichtlich der korrekten Reaktion zu treffen (z. B. linken bzw. rechten Knopf drücken, wenn Punkt links bzw. rechts erscheint) (W. J. Schneider & McGrew, 2018). Im Unterschied zur visuellen Verarbeitungsgeschwindigkeit ist die Operationalisierung somit nicht nur auf visuelle Stimuli beschränkt. Auch ist für die Auswertung mehrheitlich die benötigte Zeit von Inte-

resse, da durch den geringen Komplexitätsgrad der Testaufgaben davon auszugehen ist, dass alle Testpersonen sie bei ausreichender Zeit lösen. Bei der psychomotorischen Geschwindigkeit geht es hingegen um die Geschwindigkeit, mit der vertraute Körperbewegungen auf Basis von zuvor verarbeiteten Informationen ausgeführt werden (Chaiken, Kyllonen & Tirre, 2000). Wichtig bei der Operationalisierung ist daher, das Ausmaß an komplexen Informationen und unvertrauten Körperbewegungen gering zu halten, andernfalls erfassen die Testaufgaben mehrheitlich die Befähigung zur Informationsverarbeitung oder die Motorik. Typische Testaufgaben sind demzufolge zum Beispiel die Sprech- oder Schreibgeschwindigkeit, bei denen bekannte Sequenzen (z. B. das Alphabet) möglichst schnell gesprochen oder geschrieben werden müssen (Keys & White, 2000).

Bezogen auf die Hauptkategorie Verarbeitungsgeschwindigkeit berichteten die Übersichtsarbeiten von Peng et al. (2018) und Johnson et al. (2010) von Defiziten bei Personen mit Rechenstörung. Eine mögliche Begründung hierfür war, dass eine langsamere Verarbeitungsgeschwindigkeit ebenso die Geschwindigkeit beeinträchtigt, mit der Fakten aus dem Langzeitgedächtnis abgerufen werden. Im Falle der Rechenstörung bedeutet dies, dass betroffene Personen länger benötigen, um den Wert einer vorgegebenen Zahl zu erfassen bzw. der Zahl ihre dahinterstehende Menge zuzuweisen. Dies führt in weiterer Folge zu Defiziten, insbesondere beim Zählen und dem Faktenwissen (z. B. Einmaleins) (Bull & Johnston, 1997; Fuchs et al., 2008). In einer Studie von Cirino et al. (2015) wiesen dabei ausschließlich Personen mit isolierter Rechenstörung oder kombinierter Rechen- und Lesestörung, nicht jedoch Personen mit isolierter Lesestörung, Defizite in der Verarbeitungsgeschwindigkeit auf. Hieraus wurde geschlussfolgert, dass Defizite in der Verarbeitungsgeschwindigkeit ein möglicher Indikator der Rechenstörung darstellen.

Bei der Differenzierung in die einzelnen Komponenten der Verarbeitungsgeschwindigkeit zeigten sich in mehreren Studien Defizite bei der visuellen Verarbeitungsgeschwindigkeit sowohl in der Richtigkeit als auch in der benötigten Zeit (Compton, Fuchs, Fuchs, Lambert & Hamlett, 2012; Cowan & Powell, 2014; Fuchs et al., 2008; Tolar, Fuchs, Fletcher, Fuchs & Hamlett, 2016). Eine Differenzierung der Defizite hinsichtlich der verwendeten Testaufgaben ist je-

doch nicht zu erkennen. So lagen bei Personen mit Rechenstörung ähnlich große Defizite bei Buchstaben-oder-Zahlen-Folgen (z. B. Buchstaben in alphabetischer Reihenfolge mit Linien verbinden), beim Kodieren (z. B. Zahlen auf Basis eines Kodierschlüssels schnell ihren korrespondierenden Symbolen zuordnen) sowie beim visuellen Vergleich (z. B. ein Objekt unter mehreren Objekten finden) vor (D'Amico & Passolunghi, 2009; Poletti, 2016; Szűcs et al., 2013). Die Studienlage zu den anderen Komponenten ist hingegen deutlich geringer und weniger eindeutig. Für die Reaktions- und Entscheidungsgeschwindigkeit liegen vor allem Studien vor, die die Reaktionszeit als Kontrollvariable für anderweitige Fragestellungen erfassten und nicht näher untersuchten bzw. systematisch variierten. Keine Unterschiede zwischen Personen mit und ohne Rechenstörung fanden jedoch A. J. Wilson et al. (2015) und Moll et al. (2015) bei Testaufgaben, in denen auf das Erscheinen eines Stimulus schnellstmöglich reagiert werden musste. Demgegenüber berichteten Kuhn, Ise, Raddatz, Schwenk und Dobel (2016) von Defiziten bei Personen mit Rechenstörung, wenn die Testaufgaben eine zusätzliche Entscheidungskomponente beinhalten und Testpersonen je nach Stimulus anders reagieren mussten, um die Aufgabe zu lösen. Für die psychomotorische Geschwindigkeit liegen nur Daten einer publizierten Studie vor, die die Leistung zwischen Personen mit und ohne Rechenstörung verglich. Hierbei waren Personen mit Rechenstörung signifikant langsamer beim schnellen Schreiben von Zahlen- und Buchstabensequenzen (Swanson, 2012). Für die größeren Defizite in der visuellen Verarbeitungsgeschwindigkeit im Vergleich zur Reaktions- und Entscheidungsgeschwindigkeit sowie der psychometrischen Geschwindigkeit liegen keine expliziten Theorien vor. Hierbei ist jedoch zu berücksichtigen, dass die Verarbeitungsgeschwindigkeit in Intelligenztestbatterien standardmäßig durch die visuelle Verarbeitungsgeschwindigkeit erfasst wird, sodass mehr Daten zu dieser Komponente vorliegen und vergleichende Analysen mit anderen Komponenten der Verarbeitungsgeschwindigkeit im Rahmen der Studien nicht erfolgen konnten (Cirino et al., 2015).

1.2.2.4 Raumvorstellung

Die *Raumvorstellung* weist eine lange Forschungstradition auf und etliche Fähigkeiten und Testaufgaben wurden ihr in der Vergangenheit zugeordnet (für einen Überblick siehe: Hegarty & Waller, 2005), die übergreifend die korrekte

Wahrnehmung und mentale Manipulation von Objekten im Raum gemein haben (W. J. Schneider & McGrew, 2018). Die Raumvorstellung unterscheidet sich dabei von der visuellen Verarbeitungsgeschwindigkeit dahingehend, dass im Vordergrund nicht das schnelle Dekodieren einfacher visueller Stimuli (z. B. zweidimensionale Symbole), sondern das korrekte Verarbeiten komplexer visuell-räumlicher Stimuli (z. B. dreidimensionale Objekte) steht.

Hinsichtlich des Rechnens beschäftigten sich die meisten Studien mit der Rolle der Raumvorstellung für die allgemeine Rechenfähigkeit (für einen Überblick siehe: Cheng & Mix, 2014). Dabei zeigte sich, dass eine bessere Leistung in der Raumvorstellung mit einer besseren Mathematikleistung einherging (McCaskey, von Aster, O’Gorman Tuura & Kucian, 2017). Zum einen ist die Raumvorstellung bei der Mengenverarbeitung involviert, die einen stark visuell-räumlich geprägten Verarbeitungsprozess darstellt (z. B. schnelles Erfassen von Mengen im Raum) (Mussolin et al., 2011). Zum anderen wird sie bei der Bearbeitung komplexer Aufgaben des mathematischen Schlussfolgerns benötigt (z. B. Diagramme interpretieren) (Hegarty & Kozhevnikov, 1999). Für die wenigen Studien mit Bezug zur Rechenstörung ergibt sich hinsichtlich der Raumvorstellung kein unmittelbar eindeutiges Bild, wenngleich die Mehrzahl der Studien die Bedeutung der Raumvorstellung als möglichen Indikator einer Rechenstörung hervorhob. So berichtete die Übersichtarbeit von Swanson (2012) von mittleren bis großen Defiziten in der Raumvorstellung bei Personen mit Rechenstörung. Demgegenüber steht jedoch ein Review von Peng et al. (2018), bei dem lediglich kleine Defizite in der Raumvorstellung festgestellt wurden. Auffallend dabei ist, dass Peng et al. (2018) komorbide Fälle von Personen mit Rechen- und Lesestörung für die Analyse nicht ausschloss. Diese wurden im Gegensatz dazu bei Swanson (2012) separat betrachtet, woraufhin sich zeigte, dass ausschließlich bei Gruppen mit Rechenstörung (d. h. isolierte sowie komorbide Fälle) Defizite in der Raumvorstellung vorlagen. Die Gruppe der Personen mit isolierter Lesestörung wies hingegen eine normale Raumvorstellungsleistung auf. Ein geringerer Zusammenhang der Raumvorstellung mit der Lese- als mit der Rechenleistung zeigte sich ebenso bei Niileksela und Reynolds (2014).

Bezogen auf den der Raumvorstellung zugeordneten Fähigkeiten liegt der Schwerpunkt vor allem auf der *Visualisierung*, bei der es darum geht, mehrere gedankliche Operationen bei Bildern und Objekten durchzuführen (W. J. Schneider & McGrew, 2018). Testaufgaben bestehen zum Beispiel darin, Objekte mental zu rotieren oder Einzelteile eines Bildes entsprechend zusammenzusetzen. In mehreren Studien wurden hierzu Defizite bei Personen mit Rechenstörung festgestellt (McCaskey et al., 2017; Skagerlund & Träff, 2014; Swanson, 2012), wenngleich vereinzelt Gegenbelege vorlagen (Szűcs et al., 2013). Für andere Testaufgaben, bei denen zum Beispiel Längen geschätzt oder geometrische Figuren kopiert werden müssen, gibt es nur eine geringe Anzahl an Studien, weswegen keine eindeutigen Rückschlüsse hinsichtlich der Defizite bei Personen mit Rechenstörung gezogen werden können (Landerl, Bevan & Butterworth, 2004; Landerl, Fussenegger, Moll & Willburger, 2009; Murphy, Mazzocco, Hanich & Early, 2007; Pieters, Desoete, Roeyers, Vanderswalmen & van Waelvelde, 2012).

1.2.2.5 Phonologische Verarbeitung

Phonologische Verarbeitung definiert sich gemäß Wagner und Torgesen (1987, S. 192) als „the use of phonological information (...) in processing written and oral language“. Ihre Kernkomponenten sind dabei insbesondere die *phonologische Bewusstheit* (d. h. Phoneme, Silben und Anlaute erkennen und verwenden) sowie das *schnelle Benennen* (RAN; d. h. vertraute Bilder und Symbole schnell benennen). Auch besteht eine Verbindung zum auditiven Kurzzeitgedächtnis (Swanson, 2004).

Für Personen mit Rechenstörung wurden in einer Übersichtsarbeit von Peng et al. (2018) bei der phonologischen Verarbeitung insgesamt kleine Defizite festgestellt. Dieses Bild ergibt sich auch in den dazugehörigen Komponenten. So zeigten Personen mit Rechenstörung keine bis kleine Defizite in der phonologischen Bewusstheit, die üblicherweise mit Testaufgaben erfasst wird, bei denen Wörter laut ohne bestimmte Laute gesprochen werden müssen (z. B. „Katze“ ohne [k]). Die Unterschiede zu Personen ohne Rechenstörung waren zwar teilweise signifikant, jedoch deutlich geringer als zu Personen mit Lesestörung (Cirino et al., 2015; Landerl et al., 2009; Moll et al., 2015; Rubinsten & Henik, 2006; Willcutt et al., 2013). Ähnlich verhält es sich beim schnellen Benennen,

wo Personen mit Rechenstörung zwar kleine und zum Teil signifikante Unterschiede zeigten, ihre Defizite jedoch deutlich geringer waren als die der Personen mit Lesestörung (Donker, Kroesbergen, Slot, van Viersen & de Bree, 2016; Murphy et al., 2007; van der Sluis et al., 2004). Dies war sowohl bei der Richtigkeit als auch der benötigten Zeit der Fall und ebenso, wenn der Inhalt der Items variiert wurde (z. B. Zahlen, Buchstaben, Objekte) (Willburger et al., 2008).

Die phonologische Verarbeitung gilt als Basiskompetenz der Lesefähigkeit (Moll et al., 2015), weswegen Defizite bei Personen mit isolierter Rechenstörung über viele Studien hinweg entsprechend klein ausfallen (Vanbinst & de Smedt, 2016). Die dennoch gefundenen signifikanten Unterschiede werden auch darauf zurückgeführt, dass Schwierigkeiten in der phonologischen Verarbeitung durchaus als ein frühes Indiz einer Rechenstörung gelten (Peng et al., 2018). Die Entwicklung des Zahlenwissens (z. B. Zählen) und der Aufbau eines Faktenwissens sind auch sprachliche Prozesse, bei denen die Bedeutung von Symbolen (d. h. $3 = \bullet\bullet\bullet$) sowie deren Schreibung und Aussprache (d. h. drei) dekodiert und im semantischen Gedächtnis abgespeichert werden (de Smedt & Boets, 2010; Dehaene, 1992; Pollack & Ashby, 2018). Später werden diese abgespeicherten Fakten lediglich aus dem Gedächtnis abgerufen, wodurch auch Personen mit Rechenstörung weniger auf die phonologische Verarbeitung angewiesen sind.

1.3 Datenkodierung nach der Bottom-up-Methode

Metaanalysen fassen die Outcomes verschiedener Studien, die vorab kodiert wurden, statistisch zusammen. Für gewöhnlich erfolgt die Kodierung der Outcomes nach einer Top-down-Methode. Basierend auf der Fragestellung der jeweiligen Metaanalyse wird ein Kodierschema entwickelt, welches alle relevanten Outcomes sowie die Testaufgaben, die diese Outcomes erfassen, enthält (D. B. Wilson, 2009). Dies hat zur Folge, dass Outcomes, die nicht als relevant betrachtet werden, nicht im Kodierschema gelistet sind und für die Metaanalyse nicht weiter berücksichtigt werden (Lipsey, 2009). Während sich diese Top-down-Strategie für Metaanalysen mit eng gefassten Fragestellungen gut eignet (z. B. Verbessert Intervention A die Leistung in Outcome B?), so kommt sie bei explorativen und breit formulierten Fragestellungen, wie bei Profilstudien, an

ihre Grenzen. Ziel solcher Studien ist es, anhand der Leistung einer definierten Gruppe in verschiedenen Fähigkeiten ein bestimmtes Leistungsprofil zu ermitteln (Karagiannakis et al., 2014). Hier gestaltet es sich schwierig, vorab relevante Outcomes zu definieren, da die Qualität eines Profils zu großen Teilen von der Vielfalt und Vielzahl an Outcomes abhängt, die es umfasst, und von der Detailliertheit, mit denen sie die Testaufgaben beschreibt, die sich zur Erfassung dieser Outcomes eignen. Ein Kodierschema, das demzufolge vorab bereits zu viele Outcomes ausschließt, bedingt einen Informationsverlust und erhöht die Gefahr eines Bias, wenn bestimmte Fähigkeiten relevant wären, jedoch durch das Schema ausgeschlossen sind.

Brown, Upchurch und Acton (2003) schlugen aus diesem Grund eine Bottom-up-Methode beim Kodieren vor, bei der die relevanten Outcomes und Testaufgaben direkt aus den in der Metaanalyse eingeschlossenen Studien extrahiert werden. Zuerst wird ein Teil der eingeschlossenen Studien begutachtet und sämtliche dort berichteten Outcomes werden in das Kodierschema aufgenommen. Im nächsten Schritt wird dieses Kodierschema an einem anderen Teil eingeschlossener Studien pilotiert und weiter spezifiziert. Einerseits werden so die Outcomes in das Kodierschema eingeschlossen, die bisher nicht berichtet wurden. Andererseits werden Outcomes aus dem Kodierschema entfernt, die nicht gut genug zwischen den zu untersuchenden Gruppen trennen (d. h. wenn ein Outcome nur in einer Studie erfasst wurde). Das Ergebnis ist ein Kodierschema, welches alle Outcomes umfasst, die von einer ausreichend großen Anzahl an Studien, die in die Metaanalyse eingeschlossen sind, berichtet werden.

Hinsichtlich der Rechenstörung wurde dieser Ansatz bisher von keiner Metaanalyse angewandt. Das Kodierschema der Metaanalysen von Johnson et al. (2010), Peng et al. (2018) und Swanson und Jerman (2006) berücksichtigte zwar mehrere Outcomes, die jedoch durch eine Top-down-Strategie gewonnen wurden. Auch unterschieden diese Studien zwischen Personen mit und ohne Rechenstörung nur bei den Fähigkeiten. Für die Testaufgaben, die den jeweiligen Fähigkeiten zugeordnet waren, wurden nur Beispiele gegeben, wodurch unklar ist, welche Vielfalt an unterschiedlichen Testaufgaben je Fähigkeit berücksichtigt und in die Metaanalyse eingeschlossen war. Andere Metaanalysen zur Rechenstörung konzentrierten sich lediglich auf eine einzige Fähigkeit, bei-

spielsweise das Arbeitsgedächtnis (David, 2012). Bis heute existiert somit kein kognitives Profil der Rechenstörung, das allumfassend sämtliche in Studien empirisch erhobene Outcomes und verwendete Testaufgaben umfasst.

2 Zielsetzung

Ziel der vorliegenden Studie war es, das kognitive Profil der Rechenstörung zu ermitteln. Dies geschah vor dem Hintergrund, dass Personen mit Rechenstörung Defizite sowohl in domänenspezifischen als auch domänenübergreifenden Fähigkeiten zeigen. Da über die genaue Vielfalt, Schwere und Stabilität dieser Defizite jedoch weiterhin Unklarheit herrscht und keine zusammenfassenden Studien diesbezüglich vorliegen, erfolgt die Diagnostik einer Rechenstörung nicht einheitlich und Fehldiagnosen werden begünstigt. Langfristig soll diese Studie somit die Diagnosesicherheit bei der Rechenstörung erhöhen und zu einem besseren Verständnis der Heterogenität und Komplexität der Rechenstörung beitragen.

Um dies zu erreichen, wurde ein systematisches Review mit metaanalytischer Auswertung durchgeführt, bei der die Leistungen von Personen mit und ohne Rechenstörung in domänenspezifischen und domänenübergreifenden Fähigkeiten verglichen wurden. Die Ziele dieser Analyse waren:

- (1) Identifizierung derjenigen domänenspezifischen und domänenübergreifenden Fähigkeiten und ihren zugrundeliegenden Testaufgaben, die am besten zwischen Personen mit und ohne Rechenstörung unterschieden und deren Defizit als stabil zu bewerten war.
- (2) Überprüfung, ob sich die Schwere der Defizite in den Fähigkeiten und Testaufgaben unterschied, wenn die Richtigkeit oder die benötigte Zeit erfasst wurde.

3 Methode

Zum kognitiven Profil der Rechenstörung wurden vom Autor dieser Arbeit Literaturrecherchen und Auswertungen im Rahmen der Erstellung der S3-Leitlinie „Diagnostik und Behandlung der Rechenstörung“ (AWMF-Registernummer: 028-046) durchgeführt (Haberstroh & Schulte-Körne, 2019). Für die vorliegende Arbeit wurde der Umfang dieser bestehenden Analysen zusätzlich auf die Testaufgaben ausgeweitet, um differenziertere Befunde zur Heterogenität der Rechenstörung zu ermitteln. Dies bedurfte einer erneuten Literatursuche sowie einem anderen methodischen Vorgehen hinsichtlich Kodierung und Datenauswertung im Vergleich zur Leitlinie.

Zur Erreichung der Ziele der vorliegenden Dissertation wurde eine systematische Literaturrecherche nach Studien durchgeführt, die über Gruppenunterschiede zwischen Personen mit und ohne Rechenstörung berichteten. Anschließend erfolgten auf Basis dieser Studien mehrere multivariate Mehrebenen-Metaanalysen. Diese ermittelten die durchschnittlichen Effektstärken von domänenspezifischen und domänenübergreifenden Fähigkeiten und Testaufgaben für die Richtigkeit und die benötigte Zeit.

3.1 Einschlusskriterien

Eingeschlossene Studien erfüllten die folgenden Kriterien:

- (a) Studien waren in Deutsch oder Englisch veröffentlicht. Studien in Deutsch wurden zumeist auch in den deutschsprachigen Ländern durchgeführt und wiesen folglich die höchste Relevanz für die Nutzergruppe dieser Arbeit auf. Das Gros der Studien zur Rechenstörung war in Englisch publiziert, weswegen Studien in englischer Sprache ebenso eingeschlossen waren.
- (b) Studien wurden 1992 oder später veröffentlicht. Die Festlegung auf das Jahr 1992 entsprach dem Erscheinungsdatum des bis heute gültigen ICD-10 (World Health Organization, 1992).
- (c) Studien berichteten über Unterschiede zwischen einer Gruppe mit und einer Gruppe ohne Rechenstörung. Die Gruppenunterschiede durften nicht in Zusammenhang mit einer Intervention stehen, die im Rahmen

- der Studie durchgeführt wurde. Es gab keine Mindestanzahl an Personen pro Gruppe (d. h. mindestens zwei Personen pro Gruppe).
- (d) Bei den Testpersonen der Gruppe mit Rechenstörung wurde eine Rechenstörung diagnostiziert mit (1) einem Prozentrang kleiner oder gleich 25 in einem standardisierten Test zur Erfassung der Mathematikleistung oder (2) einem Rückstand der Mathematikleistung von mindestens 12 Lebens- bzw. 10 Unterrichtsmonaten (d. h. ein Jahr bzw. eine Jahrgangsstufe) oder (3) einer bereits vorhandenen oder im Rahmen der Studie festgestellten Diagnose einer Rechenstörung (DSM-IV/5: 315.1; ICD-10: F81.2) ausgestellt von einer entsprechenden Fachkraft. Zur Diagnose einer Rechenstörung existieren verschiedene Vorgehensweisen und Cut-Offs (Kaufmann et al., 2013; Möller, Fischer, Cress & Nuerk, 2012). Ein Prozentrang von kleiner oder gleich 25 stellte ein gängiges Kriterium zur Identifikation von Personen mit schwach sowie stark ausgeprägter Rechenstörung dar (Jolles et al., 2016; Swanson, Jerman & Zheng, 2009; A. J. Wilson et al., 2015).
- (e) Bei den Testpersonen der Gruppe ohne Rechenstörung wurde eine Rechenstörung ausgeschlossen mit (1) einem Prozentrang größer 25 in einem standardisierten Test zur Erfassung der Mathematikleistung oder (2) einem Rückstand der Mathematikleistung von weniger als 12 Lebens- bzw. 10 Unterrichtsmonaten (d. h. ein Jahr bzw. eine Jahrgangsstufe) oder (3) einer nicht vorhandenen Diagnose einer Rechenstörung (DSM-IV/5: 315.1; ICD-10: F81.2) ausgestellt von einer entsprechenden Fachkraft.
- (f) Testpersonen der Gruppen mit und ohne Rechenstörung besuchten mindestens die 1. Jahrgangsstufe einer Regelschule (d. h. kein sonderpädagogischer Förderbedarf) oder hatten ihre Schullaufbahn auf einer Regelschule bereits absolviert. Ein Risiko für eine Rechenstörung ist bereits im Kindergarten und der Schuleingangsphase festzustellen und entsprechende Defizite in der Mathematikleistung zeigen sich über die gesamte Schullaufbahn hinweg (Desoete et al., 2012; Geary et al., 2012; Jordan, Glutting & Ramineni, 2010; Stock, Desoete & Roeyers, 2010). Aus diesem Grund wurden Studien mit Testpersonen ab der 1. Jahr-

gangsstufe eingeschlossen. Es gab keine obere Altersgrenze für die Testpersonen.

- (g) Defizite in der Mathematikleistung der Gruppe mit Rechenstörung waren nicht primär bedingt durch neurogenetische oder hirnorganische Erkrankungen. Hierzu zählen in Zusammenhang mit der Mathematikleistung insbesondere das Fragile-X-Syndrom (Murphy, 2009), Turner-Syndrom (Baker & Reiss, 2016), Deletionssyndrom 22q11 (Brankaer, Ghesquière, de Wel, Swillen & de Smedt, 2016; de Smedt, Swillen, Verschaffel & Ghesquiere, 2009) sowie eine Neurofibromatose Typ 1 (Orraca-Castillo, Estevez-Perez & Reigosa-Crespo, 2014), infantile Zerebralparese (van Rooijen, Verhoeven & Steenbergen, 2015) oder Epilepsie (van Iterson, de Jong & Zijlstra, 2015). Ebenso waren Defizite in der Mathematikleistung nicht hauptsächlich Folge einer Frühgeburt (Jaekel & Wolke, 2014) oder einem geringen Geburtsgewicht (Taylor, Espy & Anderson, 2009).
- (h) Testpersonen der Gruppen mit und ohne Rechenstörung wiesen keine Intelligenzminderung gemäß ICD-10 (F70-F79) auf. Der IQ sämtlicher Studienteilnehmer war mindestens 70. Da ein Mindest-IQ von 70 relativ niedrig war, galt das Intelligenzkriterium auch als erfüllt, wenn innerhalb der Studien ein „durchschnittlicher IQ“ als Einschlussbedingung erwähnt, jedoch die genaue Höhe nicht spezifiziert wurde.
- (i) Testpersonen der Gruppen mit und ohne Rechenstörung wiesen keine Lesestörung auf. Eine Lesestörung lag nicht vor bei (1) einem Prozentrang größer 16 in einem standardisierten Test zur Erfassung der Leseleistung oder (2) einem Rückstand der Leseleistung von weniger als 18 Lebens- bzw. 15 Unterrichtsmonaten (d. h. 1,5 Jahre bzw. 1,5 Jahrgangsstufen) oder (3) einer nicht vorhandenen Diagnose einer Lesestörung (DSM-IV/5: 315.0; ICD-10: F81.0, F81.3) ausgestellt von einer entsprechenden Fachkraft. Bei ca. 25 bis 45 % der Personen mit Rechenstörung tritt zusätzlich eine Lesestörung auf (Dirks, Spyer, van Lieshout & de Sonnevile, 2008; Fischbach et al., 2013; Landerl & Moll, 2010; Moll et al., 2014; Willcutt et al., 2013). Es ist jedoch weiterhin unklar, ob sich die Profile von Personen mit isolierter Rechenstörung sowie kombinierter Rechen- und Lesestörung unterscheiden. Als Kompromiss wurde daher eine stark ausgeprägte Lesestörung (d. h. Prozentrang kleiner oder

gleich 16) ausgeschlossen und komorbide Fälle einer Rechenstörung mit einer schwach ausgeprägten Lesestörung (d. h. Prozentrang größer 16 und kleiner oder gleich 25) akzeptiert.

- (j) Die Gruppen mit und ohne Rechenstörung waren gematcht nach Alter (alternativ: Jahrgangsstufe) und dem Geschlecht. War keine Altersangabe gegeben, so war das Kriterium dennoch erfüllt, wenn die Testpersonen beider Gruppen in der gleichen Jahrgangsstufe waren oder die Verteilung der Testpersonen über verschiedene Jahrgangsstufen hinweg sich in beiden Gruppen nicht signifikant unterschied. Ziel des Kriteriums war es, altersbedingte Entwicklungsunterschiede sowie eine unterschiedliche Besuchsdauer zwischen den Gruppen auszuschließen. Hinsichtlich des Geschlechts ist weiterhin unklar, ob Rechenstörung häufiger bei Jungen als bei Mädchen auftritt (Fischbach et al., 2013; Moll et al., 2014) oder ob dieser Unterschied abhängig von den jeweiligen Diagnosekriterien ist (Devine, Soltész, Nobes, Goswami & Szűcs, 2013). In Bezug auf kognitive Fähigkeiten zeigte sich bisher keiner oder nur ein geringer Geschlechtsunterschied (Hyde, 2016; Voyer & Voyer, 2014). Da die Rolle des Geschlechts für die Entwicklung einer Rechenstörung sowie damit verbundene Defizite in kognitiven Fähigkeiten noch nicht abschließend geklärt ist, mussten beide Gruppen nach dem Geschlecht gematcht sein oder sich in der Verteilung an männlichen und weiblichen Testpersonen zwischen beiden Gruppen nicht unterscheiden.
- (k) Studien berichteten mindestens einen Outcome zu einer kognitiven Fähigkeit, der über die Richtigkeit oder benötigte Zeit in einem Leistungstest erfasst wurde. Wurden mehrere solcher Outcomes je Studie berichtet, wurden alle eingeschlossen. Im Zuge der Kodierung und Auswertung erfolgte eine Differenzierung zwischen den Outcomes sowie die Berücksichtigung statistischer Abhängigkeiten multipler Outcomes je Studie. Outcomes auf Basis von Persönlichkeitsfragebögen waren demzufolge nicht eingeschlossen.
- (l) Studien berichteten genügend Daten zur Berechnung der Effektstärke und des Standardfehlers der Effektstärke auf Basis des standardisierten Mittelwertunterschieds zwischen Gruppen mit und ohne Rechenstörung. Effektstärken auf Basis dichotomer Angaben (z. B. Gruppenzugehörig-

keiten) sowie Korrelationswerten waren daher ausgeschlossen, da deren statistischen Auswertung und Interpretation anders erfolgte. Im Falle fehlender bzw. nicht-berichteter Werte wurden die Studienautoren kontaktiert oder es wurde versucht, die Werte auf Basis anderer berichteter Daten zu berechnen (siehe 3.3.3).

3.2 Suchstrategie

Die systematische Literaturrecherche erfolgte im Oktober 2017 in den folgenden Datenbanken: *PsycINFO*, *MEDLINE*, *ERIC*, *ProQuest*, *PSYINDEX* und *MathEduc*. Die Datenbanken ProQuest und PsycINFO wurden überdies zur Suche nach relevanten Dissertationen sowie Masterabschlussarbeiten verwendet. Zusätzlich wurden die Literaturverzeichnisse der folgenden Übersichtsartikel und Reviews (mit und ohne systematischer Literaturrecherche) nach relevanten Artikeln überprüft: David (2012), Geary (2004, 2010, 2011), Johnson et al. (2010), Peng und Fuchs (2014), Peng, Namkung, Barnes und Sun (2016), Peng et al. (2018), Raghobar, Barnes und Hecht (2010), Shin und Bryant (2015), Swanson und Jerman (2006), Swanson et al. (2009) und Vanbinst und de Smedt (2016).

Je nach Datenbank wurde mit deutschen oder englischen Begriffen gesucht. Um Studien zu finden, die über Unterschiede zwischen Personen mit und ohne Rechenstörung berichten, wurden jeweils für Deutsch und Englisch Suchbegriffe zu „Rechenstörung“ und zu „Unterschiede“ mit *AND* verknüpft und innerhalb des Titels, Abstracts und Schlagwörtern gesucht. Die Suchbegriffe zur Rechenstörung wurden initial durch den deutschen und englischen Thesaurus *MeSH* (Medical Subject Headings) ermittelt (Deutsches Institut für Medizinische Dokumentation und Information, 2016; NCBI, 2016). Anschließend erfolgte eine Pilotsuche in den Datenbanken PsycINFO und PSYINDEX nach entsprechenden Studien. Relevante Schlagwörter sowie Schlüsselbegriffe in diesen Studien, die nicht bereits durch den MeSH-Thesaurus abgedeckt waren, wurden so extrahiert und der Liste an Suchbegriffen hinzugefügt. Folgende Suchbegriffe mit den entsprechenden Suchparametern für anderweitige Schreibweisen (z. B. Pluralformen) wurden verwendet:

Tabelle 1

Verwendete Suchparameter in Deutsch und Englisch

Suchbegriff	Suchparameter
Rechenstörung	
Deutsch	Rechenstörung* OR Dyskalkul* OR Rechenschwäche* OR Zahlenblindheit OR Arithmasthenie OR "Schwierigkeit* im Rechnen" OR "Problem* im Rechnen" OR "Schwierigkeit* beim Rechnen" OR "Problem* beim Rechnen" OR Rechenschwierigkeit* OR Rechenproblem* OR "des Rechnens"
Englisch	dyscalculi* OR "math* disorder*" OR "math* disabilit*" OR "math* difficult*" OR "math* learning disorder*" OR "math* learning disabilit*" OR "math* learning difficult*" OR "arithmetic* disorder*" OR "arithmetic* disabilit*" OR "arithmetic* difficult*" OR "arithmetic* learning disorder*" OR "arithmetic* learning disabilit*" OR "arithmetic* learning difficult*" OR "impairment* in math*" OR "developmental acalculia" OR "disorder* of arithmetic* skill*" OR "low math* achievement*" OR "low arithmetic* achievement*" OR "disorder* in math*" OR "difficul* in math*" OR "disabilit* in math*" OR "disorder* in arithmetic*" OR "difficul* in arithmetic*" OR "disabilit* in arithmetic*" OR "math* LD"
Unterschiede	
Deutsch	Unterschied* OR unterscheid* OR Differenz* OR Diskrepan* OR trennt OR vergleich* OR Versuchsgruppe* OR Kontrollgruppe* OR Experimentalgruppe*
Englisch	differ OR differs OR differed OR difference* OR disting* OR discrim* OR contrast* OR gap* OR compar* OR "experimental group*" OR "control group*" OR "typical* achiev*" OR "normal* achiev*" OR "typical* develop*" OR "normal develop*"

3.3 Kodierung

Das Kodierschema setzte sich aus Variablen zu deskriptiven Studienangaben sowie zur Beschreibung der Studienergebnisse eingeschlossener Studien zusammen.

3.3.1 Deskriptive Studienangaben

Für die deskriptiven Studienangaben wurden *Studien-, Stichproben- und Diagnosemerkmale* der Studien kodiert. Studienmerkmale umfassten den Publikationstyp. Stichprobenmerkmale getrennt nach Gruppen mit und ohne Rechenstörung waren die Stichprobengröße, das Durchschnittsalter in Monaten und die Anzahl männlicher Testpersonen. Diagnosemerkmale beinhalteten die Testaufgaben des Intelligenz-, Lese- und Mathematiktests und die Kriterien zum

Ausschluss einer Intelligenzminderung und Lesestörung sowie zur Diagnose einer Rechenstörung. Das Kodierschema für die deskriptiven Studienangaben ist in Tabelle 2 dargestellt.

Tabelle 2
Kodierschema zu deskriptiven Studienangaben

Variable	Beschreibung
Studienmerkmale	
Publikationstyp	
A	Studie als Artikel in einem Journal erschienen
D/M	Studie im Rahmen einer Dissertation oder Masterarbeit
Stichprobenmerkmale (jeweils für Gruppe mit und ohne Rechenstörung)	
n	Stichprobengröße
\bar{X}_{Alter}	Mittelwert des Alters in Monaten
% männl.	Anteil männlicher Testpersonen in Prozent
Diagnosemerkmale	
IQ	
Inhalt	Testaufgaben des Intelligenztests
NV	nonverbale Testaufgaben (z. B. Matrizen)
V	verbale Testaufgaben (z. B. Zahlen, Buchstaben)
Kriterium	verwendetes Kriterium zum Ausschluss einer Intelligenzminderung
Mindest-IQ	erreichter IQ-Wert im Intelligenztest
durchschnittliche Intelligenz	„durchschnittliche Intelligenz“ wird angegeben, ohne dass ein Intelligenztest durchgeführt wurde
Lesen	
Inhalt	Testaufgaben des Lesetests
LG	Testaufgaben zur Lesegenauigkeit (z. B. Wörter-, Nichtwörter- oder Pseudowörter lesen, Buchstaben erkennen)
LV	Testaufgaben zum Leseverständnis (z. B. Fragen zu einem Text beantworten)
Kriterium	verwendetes Kriterium zum Ausschluss einer Lesestörung
Höhe PR	erreichter Prozentrang im Lesetest
Rückstand Lebens- / Unterrichtsmonate	Länge des Rückstands in der Lesefähigkeit in Lebens- oder Unterrichtsmonaten
DSM/ICD-Diagnose	Diagnose „Lesestörung“ gemäß DSM-IV/5 oder ICD-10 ausgeschlossen
Mathematik	
Inhalt	Testaufgaben des Mathematiktests

Variable	Beschreibung
B	Testaufgaben zu den Basiskompetenzen (z. B. Zahlenvergleich, Mengenvergleich, Subitizing, Zahlenstrahl)
R	Testaufgaben zum Rechnen und Faktenwissen (z. B. Einmaleins, schriftliche Aufgaben zu den Grundrechenarten)
S	Testaufgaben zum mathematisches Schlussfolgern (z. B. Textaufgaben, Diagramme interpretieren) (sowie Kombinationen daraus)
Kriterium	verwendetes Kriterium zur Diagnose einer Rechenstörung
Höhe PR	erreichter Prozentrang im Mathematiktest
Rückstand Lebens- / Unterrichtsmonate	Länge des Rückstands in der Mathematikfähigkeit in Lebens- oder Unterrichtsmonaten
DSM/ICD-Diagnose	Diagnose „Rechenstörung“ gemäß DSM-IV/5 oder ICD-10 vorhanden

3.3.2 Studienergebnisse

Für die Studienergebnisse wurden die Outcomes (d. h. die Art der Leistung, die Studien in beiden Gruppen erfassten) sowie die dazugehörigen Effektstärken kodiert. Dabei wurde ein Kodierschema nach Brown et al. (2003) entwickelt. Bei dieser Bottom-up-Methode werden die kodierbaren Outcomes direkt aus dem Pool aller eingeschlossenen Studien gewonnen, anstatt bereits vor der Kodierung festgelegt zu werden. Für die Studienergebnisse bedeutet dies, dass nur Outcomes kodiert werden, zu denen die eingeschlossenen Studien auch Daten berichten.

Um zudem besser zwischen Personen mit und ohne Rechenstörung unterscheiden zu können, wurde jeder Outcome auf die folgenden drei Levels kodiert, die einen zunehmenden Grad an Verallgemeinerung aufwiesen:

Testaufgaben (TA), *untergeordnete Fähigkeit (UF)* und *übergeordnete Fähigkeit (ÜF)*. Übergeordnete Fähigkeiten waren allgemeine kognitive Fähigkeiten (z. B. Kurzzeitgedächtnis). Jede übergeordnete Fähigkeit umfasste dabei mehrere untergeordnete Fähigkeiten, die deutlich spezifischer waren und einen besseren Grad an Differenzierung boten (z. B. visuell-räumliches Kurzzeitgedächtnis, auditives Kurzzeitgedächtnis). Testaufgaben stellten schließlich die konkrete Operationalisierung dieser unter- und übergeordneten Fähigkeiten dar. Die Entwicklung des Kodierschemas erfolgte in 3 Schritten:

- (1) Outcomes wurden Testaufgaben zugewiesen: Outcomes von Studien, die ähnlichen operationalisiert waren, wurden der gleichen Testaufgabe zugewiesen (z. B. Outcome „Vorwärts zählen von 1 bis 10“ und Outcome „Rückwärts zählen von 8 bis 2“ gehörten zu Testaufgabe „Zählen“).
- (2) Testaufgaben wurden untergeordneten Fähigkeiten zugewiesen: Testaufgaben, die die gleiche untergeordnete Fähigkeit erfassten, wurden dieser untergeordneten Fähigkeit zugewiesen (z. B. Testaufgabe „Subitizing“ und Testaufgabe „Zählen“ gehörten zur untergeordneten Fähigkeit „Basiskompetenzen“).
- (3) Untergeordneten Fähigkeiten wurden übergeordneten Fähigkeiten zugewiesen: Untergeordnete Fähigkeiten, die die gleiche übergeordnete Fähigkeit erfassten, wurden dieser übergeordneten Fähigkeit zugewiesen (z. B. untergeordnete Fähigkeit „Basiskompetenzen“ und untergeordnete Fähigkeit „Rechnen“ gehörten zur übergeordneten Fähigkeit „Mathematik“).

Konnte ein Outcome keiner konkreten Testaufgabe zugewiesen werden (z. B. Faktorwerte, die mehrere Testaufgaben zusammenfassten), so wurde der Outcome, wenn möglich, nur seiner entsprechenden untergeordneten oder übergeordneten Fähigkeiten zugewiesen (z. B. Outcome erfasste Leistung in Testaufgaben „Zahlenvergleich“ und „Mengenvergleich“ wurde untergeordneter Fähigkeit „Basiskompetenzen“ und übergeordneter Fähigkeit „Mathematik“ zugeordnet).

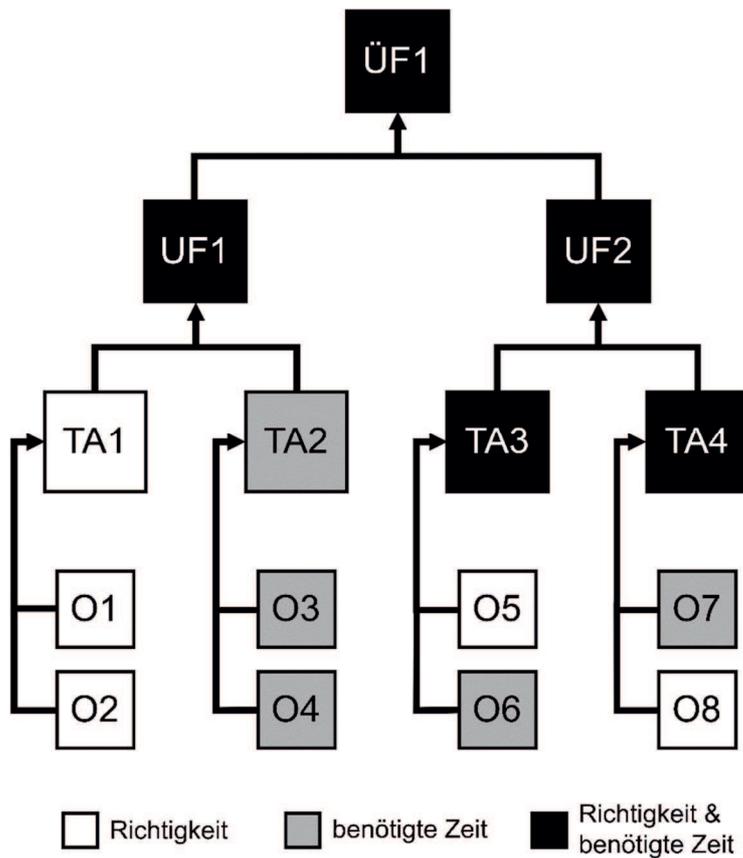


Abbildung 2. Schematische Darstellung des Kodiervorgehens.
 O = Outcome; TA = Testaufgabe; UF = untergeordnete Fähigkeit; ÜF = übergeordnete Fähigkeit.

Nach Zuordnung aller Outcomes zu den entsprechenden Levels Testaufgaben, untergeordnete Fähigkeit und übergeordnete Fähigkeit wurde differenziert, ob der Outcome die Skala *Richtigkeit* (RI) oder *benötigte Zeit* (BZ) erfasste. Outcomes, deren Effektstärken zum Beispiel auf der Anzahl an Lösungen oder Fehler basierten, erfassten die Skala Richtigkeit. Outcomes, deren Effektstärken zum Beispiel auf der durchschnittlichen Zeit zur Lösung von Aufgaben basierten, erfassten die Skala benötigte Zeit. Abbildung 2 zeigt eine schematische Darstellung des Kodierprozesses unter Berücksichtigung der verschiedenen Levels und Skalen.

Je Studie wurden alle Outcomes eingeschlossen, die Werte zur Richtigkeit oder benötigten Zeit berichteten und mindestens einer der 3 Level übergeordnete Fähigkeit, untergeordnete Fähigkeit oder Testaufgabe zugewiesen werden konnten. Dieses Vorgehen resultierte schließlich in einem gemeinsamen Ko-

dierschema, welches zu 6 voneinander unabhängigen Datensätzen kodierter Outcomes führte: (1) Richtigkeit / übergeordnete Fähigkeit (RI/ÜF), (2) Richtigkeit / untergeordnete Fähigkeit (RI /UF), (3) Richtigkeit / Testaufgabe (RI/TA), (4) benötigte Zeit / übergeordnete Fähigkeit (BZ/ÜF), (5) benötigte Zeit / untergeordnete Fähigkeit (BZ/UF) und (6) benötigte Zeit / Testaufgabe (BZ/TA). Das finale Kodierschema zu den Studienergebnissen mit Beschreibungen der einzelnen Testaufgaben, unter- und übergeordneten Fähigkeiten ist in Tabelle 3 dargestellt.

Durch diese Vorgehensweise, verschiedene Levels wie Skalen von Outcomes zu berücksichtigen und separat zu kodieren, war es letztlich möglich, differenzierte Aussagen zu den Unterschieden zwischen Personen mit und ohne Rechenstörung zu treffen und somit die Ziele dieser Arbeit zu erreichen.

Tabelle 3

Kodierschema zu Outcomes der Studienergebnisse und deren Operationalisierung

Outcome (ÜF/UF/TA) ^a	Beschreibung	Operationalisierung
	domänenspezifische Fähigkeiten	
Mathematik	Mathematische oder mathematisch-bezogene Aufgaben durch Rechnen, Schlussfolgern oder Zahlen- und Mengenverarbeitung lösen.	Beinhaltet alle Outcomes der zugehörigen UF und TA. Keine weiteren Outcomes, die keiner UF bzw. TA zugeordnet werden konnten.
Rechnen	Rechenaufgaben lösen.	Beinhaltet alle Outcomes der zugehörigen TA. Keine weiteren Outcomes, die keiner TA zugeordnet werden konnten.
Grundrechenarten	Rechenaufgaben durch Anwendung der Grundrechenarten lösen.	818 + 104
Schätzen	Rechenaufgaben durch Schätzen lösen. Rechenaufgaben sind zu komplex, um das Ergebnis aus dem Gedächtnis abzurufen oder sie innerhalb des vorgegeben Zeitlimits zu berechnen.	497 / 27. Welche Zahl ist näher am Ergebnis: 20 oder 40?
Faktenwissen	Rechenaufgaben lösen, indem das Ergebnis direkt aus dem Gedächtnis abgerufen wird (z. B. Einmaleins).	3 * 5
Mathematisches Schlussfolgern	Komplexe, nicht-vertraute oder mehrstufige mathematische Aufgaben durch Anwendung des gesamten und erlernten mathematischen Wissens lösen.	Beinhaltet alle Outcomes der zugehörigen TA. Weitere Outcomes, die keiner TA zugeordnet werden konnten: Diagramme, Tabellen und Grafiken interpretieren. Geometrische Aufgaben lösen. Zahlenfolgen (z. B. 3 5 7 9... Was ist die nächste Zahl?)

Outcome (ÜF/UF/TA) ^a	Beschreibung	Operationalisierung
Textaufgaben	Rechenaufgaben lösen, die als Text dargestellt sind, durch Anwendung der korrekten Rechenregel.	Peter hat 5 Äpfel. Er gibt Anna 1 Apfel und Paul 2 Äpfel. Wie viele Äpfel hat Peter noch?
Basiskompetenzen	Zahlen- und Mengenverarbeitung.	Beinhaltet alle Outcomes der zugehörigen TA. Keine weiteren Outcomes, die keiner TA zugeordnet werden konnten.
Simultanerfassung (bis 3)	Eine Menge von höchstens 3 Objekten schnell und ohne zu zählen erfassen (d. h. Subitizing).	2 schwarze Punkte sind für 500 ms vor einem weißen Hintergrund zu sehen. Die Anzahl an Punkten wird sofort nach ihrer Vorgabe genannt.
Simultanerfassung (ab 4)	Eine Menge von mindestens 4 Objekten schnell und ohne zu zählen erfassen (d. h. Dot enumeration).	5 schwarze Punkte sind für 500 ms vor einem weißen Hintergrund zu sehen. Die Anzahl an Punkten wird sofort nach ihrer Vorgabe genannt.
Mengenvergleich	Zwei Mengen (z. B. Punktmengen) hinsichtlich ihrer Mächtigkeit vergleichen.	Eine Menge von 4 Punkten wird neben einer Menge von 7 Punkten gezeigt. Welche Punktmenge ist größer?
Zahlenvergleich	Zwei Zahlen hinsichtlich ihrer Mächtigkeit vergleichen.	Die Zahl 4 wird neben der Zahl 7 gezeigt. Welche Zahl ist größer?
Zahlen-Mengen-Vergleich	Eine Zahl und ein Menge oder Zahl-Mengen-Kombination hinsichtlich ihrer Mächtigkeit vergleichen.	Eine Menge von 2 Punkten und die Zahl 2 werden neben der Zahl 7 gezeigt. Welche Menge ist größer?
Zahlenstrahl	Die Position einer Zahl auf dem Zahlenstrahl schätzen oder einer Zahl eine Markierung auf dem Zahlenstrahl zuordnen,	Ein Zahlenstrahl von 0 bis 100 wird vorgegeben. Bestimme die Zahl 40 auf dem Zahlenstrahl.
Zählen	Laut vorwärts- / rückwärts- oder weiterzählen oder Objekten abzählen.	Zähle laut von 10 bis 20.
Transkodieren	Visuell vorgegebene arabische Zahlen vorlesen oder die arabische Form einer auditiv vorgegebenen Zahl wiedergeben	Die Zahl 17 wird laut vorgelesen (d. h. "siebzehn"). Schreibe die Zahl in arabischer Form (d. h. 17).

Outcome (ÜF/UF/TA) ^a	Beschreibung	Operationalisierung
	(z. B. auf Zahl deuten, Zahl niederschreiben).	
	domänenübergreifende Fähigkeiten	
Arbeitsgedächtnis	Komplexe kognitive Prozesse überwachen und regulieren.	Beinhaltet alle Outcomes der zugehörigen UF und TA. Keine weiteren Outcomes, die keiner UF bzw. TA zugeordnet werden konnten.
Arbeitsgedächtniskapazität	Mehrere Informationen (zusammenhängend oder nicht-zusammenhängend) gleichzeitig speichern und verarbeiten.	Beinhaltet alle Outcomes der zugehörigen TA. Weitere Outcomes, die keiner TA zugeordnet werden konnten: Doppelaufgabe: Jeweils eine Menge an Zahlen und Buchstaben in falscher Abfolge werden gezeigt (z. B. 3-1-4-2 und C-B-A-D). Wiederhole die Zahlen und Buchstaben jeweils in richtiger Abfolge (d. h. 1-2-3-4 und A-B-C-D)
Spanne rückwärts	Auditive oder visuell-räumliche Informationen speichern und in umgekehrter Vorgabereihenfolge abrufen.	Eine Sequenz von Buchstaben wird vorgegeben (z. B. F-T-C-A). Wiederhole die Buchstaben in umgekehrter Reihenfolge (d.h. A-C-T-F).
Spanne komplex	Auditive oder visuell-räumliche Informationen speichern und in richtiger Reihenfolge abrufen, während gleichzeitig aufgabenirrelevante Informationen verarbeitet werden.	Ein Set von Sätzen wird vorgelesen (z. B. Der Himmel ist grün. Ein Ball ist rund. Das Eis ist kalt.). Entscheide nach jedem Satz, ob dessen Aussage wahr oder falsch ist. Wiederhole am Ende des Sets das letzte Wort jeden Satzes in der vorgegeben Reihenfolge (d. h. grün-rund-kalt).
Kurzzeitgedächtnis (KZG)	Informationen kurzzeitig speichern.	Beinhaltet alle Outcomes der zugehörigen UF und TA. Keine weiteren Outcomes, die keiner UF bzw. TA zugeordnet werden konnten.

Outcome (ÜF/UF/TA) ^a	Beschreibung	Operationalisierung
Auditives KZG	Auditive Informationen kurzzeitig speichern und unmittelbar abrufen.	Beinhaltet alle Outcomes der zugehörigen TA. Keine weiteren Outcomes, die keiner TA zugeordnet werden konnten.
Spanne vorwärts (auditiv)	Auditive Informationen speichern und unmittelbar in der vorgegebenen Reihenfolge abrufen.	Eine Sequenz von Buchstaben wird vorgegeben (z. B. F-T-C-A). Wiederhole die Buchstaben in der vorgegebenen Reihenfolge (d.h. F-T-C-A).
Abrufen (auditiv)	Auditive Informationen speichern und unmittelbar in irgendeiner Reihenfolge abrufen.	Eine Sequenz von Buchstaben wird vorgegeben (z. B. F-T-C-A). Wiederhole die Buchstaben in irgendeiner Reihenfolge (d.h. T-A-F-C).
Visuell-räumliches KZG	Visuell-räumliche Informationen kurzzeitig speichern und unmittelbar abrufen.	Beinhaltet alle Outcomes der zugehörigen TA. Keine weiteren Outcomes, die keinem TA zugeordnet werden konnten.
Spanne vorwärts (visuell-räumlich)	Visuell-räumliche Informationen speichern und unmittelbar in der vorgegebenen Reihenfolge abrufen.	Corsi block-tapping test: Eine Menge an unregelmäßig angeordneten Blöcken wird vorgegeben. Mehrere Blöcke werden nacheinander angetippt. Tippe die gleichen Blöcke in der gleichen Reihenfolge an.
Abrufen (visuell-räumlich)	Visuell-räumliche Informationen speichern und unmittelbar in irgendeiner Reihenfolge abrufen.	Eine Menge an unregelmäßig angeordneten Blöcken wird vorgegeben. Mehrere Blöcke werden nacheinander angetippt. Tippe die gleichen Blöcke in irgendeiner Reihenfolge an.
Aufmerksamkeit	Die kognitive Leistung und kognitiven Ressourcen überwachen, anpassen und regulieren als Reaktion auf sich ändernde Aufgabenstellungen.	Beinhaltet alle Outcomes der zugehörigen UF und TA. Keine weiteren Outcomes, die keiner UF bzw. TA zugeordnet werden konnten.
Inhibition	Dominante, jedoch falsche Antworttendenz bewusst unterdrücken.	Beinhaltet alle Outcomes der zugehörigen TA. Weitere Outcomes, die keiner TA zugeordnet werden konnten: Flanker Aufgabe (z. B. 4 Pfeile zeigen nach links. Ein

Outcome (ÜF/UF/TA) ^a	Beschreibung	Operationalisierung
Zufallsgenerierung	Zufällige Zahlen- oder Buchstabenspanne generieren und dabei bekannte Sequenzen (z. B. „1, 2, 3“, „a, b, c“) bewusst unterdrücken.	<p>fünfter Pfeil in der Mitte dieser 4 Pfeile zeigt nach rechts. Drücke den Knopf nach links oder rechts, je nachdem in welche Richtung der Pfeil in der Mitte zeigt.)</p> <p>Go/Nogo (z. B. Ein Pfeil nach links oder rechts wird vorgegeben. Ist nach der Vorgabe des Pfeils ein Ton zu hören, drücke den Knopf nach links oder rechts, je nachdem in welche Richtung der Pfeil zeigt. Ist nach der Vorgabe kein Ton zu hören, drücke keinen Knopf.)</p> <p>Schreibe die Zahlen von 1 bis 10 in zufälliger Reihenfolge (z. B. 4-1-9-7).</p>
Stroop-Effekt	Merkmal eines Objekts benennen und dabei ein dominantes Merkmal des oder eines anderen Objekts bewusst unterdrücken.	Das Wort „blau“ wird in gelber Schrift gezeigt. Benenne die Schriftfarbe (d. h. gelb) und nicht den Inhalt des Wortes (d. h. blau).
Shifting	Flexibel zwischen verschiedenen Aufgabenstellungen innerhalb einer Aufgabe wechseln.	<p>Beinhaltet alle Outcomes der zugehörigen TA.</p> <p>Weitere Outcomes, die keiner TA zugeordnet werden konnten:</p> <p>Eine Zahl (z. B. 2) wird innerhalb eines Objekts (z. B. Dreieck) gezeigt. Ist die Farbe der Zahl und des Objekts rot, nenne die Zahl. Ist die Farbe der Zahl und des Objekts blau, benenne das Objekt.</p>
Buchstaben-und-Zahlen-Folgen	Buchstaben und Zahlen in korrekter Reihenfolge abwechselnd verbinden.	Buchstaben und Zahlen sind zufällig auf einer Seite verteilt. Verbinde mit Linien abwechselnd die Zahlen und Buchstaben in der jeweils richtigen Abfolge (d. h. 1-A-2-B-3-C usw.)

Outcome (ÜF/UF/TA) ^a	Beschreibung	Operationalisierung
Updating	Neue Informationen bei der Aufgabebearbeitung einbeziehen, die alte Informationen ersetzen.	Weitere Outcomes, die keiner TA zugeordnet werden konnten: Sequenzen unterschiedlicher Länge von Zahlen werden vorgelesen. Benenne für jede Sequenz die letzten 3 Zahlen. Zahlen werden nacheinander gezeigt. Drücke einen Knopf, wenn die gezeigte Zahl die gleiche ist, die zwei Zahlen zuvor gezeigt wurde.
Verarbeitungsgeschwindigkeit (VG)	Informationen schnell dekodieren und darauf basierende einfache kognitive Aufgaben schnell bearbeiten.	Beinhaltet alle Outcomes der zugehörigen UF und TA. Keine weiteren Outcomes, die keiner UF bzw. TA zugeordnet werden konnten.
Visuelle VG	Einfache visuelle Stimuli schnell in einem visuellen Feld suchen und/oder visuelle Stimuli auf Unterschiede und Ähnlichkeiten vergleichen.	Beinhaltet alle Outcomes der zugehörigen TA. Keine weiteren Outcomes, die keiner TA zugeordnet werden konnten.
Buchstaben-oder-Zahlen-Folgen	Buchstaben oder Zahlen in korrekter Reihenfolge verbinden	Buchstaben oder Zahlen sind zufällig auf einer Seite verteilt. Verbinde mit Linien die Buchstaben oder Zahlen in der richtigen Abfolge (d. h. A-B-C-D bzw. 1-2-3-4 usw.).
Kodieren	Unvertraute Symbole schnell den zugehörigen Zahlen zuordnen.	Die Zahlen 1 bis 9 sind jeweils mit einem unvertrauten Symbol verknüpft. Mehrere Zahlen von 1 bis 9 werden vorgegeben. Weise den Zahlen so schnell wie möglich die korrespondierenden Symbole zu.
Visueller Vergleich	Identische Objekte oder spezifische Zielobjekte unter mehreren Objekten finden.	Finde das Zielobjekt in einer Reihe von 6 verschiedenen Objekten und kreise es ein.
Reaktions-/Entscheidungsgeschwindigkeit	Schnell reagieren, wenn eine bestimmte Bedingung erfüllt ist oder schnell ent-	Beinhaltet alle Outcomes der zugehörigen TA. Weitere Outcomes, die keiner TA zugeordnet werden

Outcome (ÜF/UF/TA) ^a	Beschreibung	Operationalisierung
	scheiden, welche Bedingung erfüllt ist und entsprechend reagieren.	konnten: Reagieren je nach Stimulus (Ein leerer Bildschirm wird gezeigt. Ein Stimulus erscheint links oder rechts. Drücke so schnell wie möglich dementsprechend auf den linken oder rechten Knopf.)
Reagieren auf Stimulus	Auf das Erscheinen eines Stimulus möglichst schnell reagieren.	Ein leerer Bildschirm wird gezeigt. Drücke so schnell wie möglich einen Knopf, sobald ein Punkt erscheint.
Psychomotorische Geschwindigkeit	Schnell und flüssig vertraute Körperbewegungen im Rahmen der Aufgabenbearbeitung durchführen.	Beinhaltet alle Outcomes der zugehörigen TA. Weitere Outcomes, die keiner TA zugeordnet werden konnten: Sprechgeschwindigkeit (z. B. Spreche innerhalb von 90 Sekunden so oft wie möglich die Buchstaben von A bis Z in der richtigen Abfolge.)
Schreibgeschwindigkeit	Schnell und flüssig bekannte Sequenzen aufschreiben oder Inhalte abschreiben.	Schreibe innerhalb von 90 Sekunden so oft wie möglich die Buchstaben von A bis Z in der richtigen Abfolge.
Raumvorstellung	Komplexe visuelle Stimuli bzw. Objekte im Raum wahrnehmen und korrekte mentale Manipulationen durchführen.	Beinhaltet alle Outcomes der zugehörigen UF. Weitere Outcomes, die keiner UF bzw. TA zugeordnet werden konnten: Spatial scanning: Löse ein Labyrinth.
Visualisierung	Mehrere gedankliche Operationen bei Bildern und Objekten durchführen.	Ein Zielobjekt und mehrere relevante und irrelevante Einzelteile werden vorgegeben. Wähle die Einzelteile aus, die zusammen das Zielobjekt ergeben. Ein Zielobjekt und 5 andere Objekte werden vorgegeben. Finde das Objekt, welches nach einer Rotation dem Zielobjekt entspricht.
Phonologische Verarbeitung	Phonologische Informationen bei der	Beinhaltet alle Outcomes der zugehörigen UF. Keine

Outcome (ÜF/UF/TA) ^a	Beschreibung	Operationalisierung
Phonologische Bewusstheit	Verarbeitung geschriebener oder gesprochener Sprache verwenden. Phoneme, Silben und Anlaute erkennen und verwenden.	weiteren Outcomes, die keiner UF bzw. TA zugeordnet werden konnten. Sage ein Wort ohne einen bestimmten Laut (z. B. „Katze“ ohne [k]).
Schnelles Benennen	Vertraute Bilder oder Symbole schnell benennen (RAN bzw. rapid automatized naming)	Mehrere Buchstaben werden nacheinander vorgegeben. Benenne jeden Buchstaben so schnell wie möglich.

^aDie Unterordnung der Outcomes entspricht den Levels der Kodierung, d. h. kein Einzug links = übergeordnete Fähigkeit (ÜF), mittlerer Einzug links = untergeordnete Fähigkeit (UF), großer Einzug links = Testaufgabe (TA).

3.3.3 Ablauf der Kodierung

Alle Variablen wurden vom Autor dieser Arbeit sowie zum Teil von einer studentischen Hilfskraft kodiert. Die Überprüfung der Interrater-Reliabilität erfolgte an einer Teilstichprobe von 10 Studien, die von beiden kodiert wurden. Bei nominalen Variablen des Kodierschemas (z. B. Testaufgaben des Mathematiktests) wurde Cohens (1960) Kappa als Reliabilitätskoeffizient bestimmt, der die tatsächliche Übereinstimmung der kodierten Werte einer zufallsbedingten Übereinstimmung gegenüberstellt. Für Variablen höheren Skalenniveaus (v. a. Ordinal- und Intervallskala) diente die Intraklassenkorrelation als Grad der Übereinstimmung (z. B. Stichprobengröße). Hierbei wurde als entsprechende Modellvariante ein *two-way random effects model* mit absoluter Übereinstimmung (*absolute agreement*) der Einzelwerte (*single-measures*) berechnet (Hallgren, 2012; Orwin & Vevea, 2009). Die Interrater-Reliabilität aller Variablen lag zwischen 88 und 97 %. Die geringeren Übereinstimmungsraten gab es bei den deskriptiven Studienangaben, wo von Seiten der eingeschlossenen Studien die verwendeten Diagnosekriterien nicht eindeutig genug beschrieben wurden (z. B. unklar, ob Diagnose Rechenstörung bei Werten *kleiner* oder *kleiner gleich* erfolgte).

Die Qualität jeder Studie wurde vom Autor dieser Arbeit mit dem *Appraisal Tool For Cross-Sectional Studies (AXIS)* (Downes, Brennan, Williams & Dean, 2016) beurteilt. Diese Checkliste bewertet anhand von 20 Ja/Nein-Fragen das Studiendesign (z. B. Auswahl der Stichprobe in Bezug zur Fragestellung), die Gefahr eines Bias (z. B. nicht gegebene Vergleichbarkeit der Gruppen, selektive Berichterstattung) und die inhaltliche Vollständigkeit (z. B. fehlende Daten, nicht replizierbar beschriebene Methodik) bei Querschnittsstudien. Die finale Beurteilung der Studienqualität erfolgt schließlich durch Einschätzung des Raters unter Berücksichtigung aller Items. Eine zusammenfassende Ratingskala zur Studienqualität liegt bei dieser Checkliste nicht vor. Dies war dahingehend begründet, dass numerische Gesamtwerte für die Studienqualität nicht als Prädiktoren für das Studienergebnis gelten (Jüni, Witschi, Bloch & Egger, 1999). Ebenso erfahren die Items der Checkliste je nach Fragestellung bei der Bewertung eine unterschiedliche Gewichtung, die durch numerische Gesamtwerte unberücksichtigt bleibt.

Im Falle von fehlenden oder nicht berichteten Werten wurde zwischen deskriptiven Studienangaben und Studienergebnissen unterschieden. Fehlende Werte bei deskriptiven Studienangaben wurden nicht weiter berücksichtigt, außer sie bezogen sich auf die Einschlusskriterien dieser Arbeit (siehe 3.1). Hiernach wurden die Korrespondenzautoren/innen der Studien kontaktiert. Erfolgte keine Rückmeldung bzw. Bereitstellung der Daten und konnten die Werte nicht aus anderen Angaben gewonnen werden, so erfüllte die Studie die Einschlusskriterien nicht und wurde ausgeschlossen. Fehlende Werte hinsichtlich der Studienergebnisse betrafen vor allem den Mittelwert und die Standardabweichung der Gruppen mit und ohne Rechenstörung, die zur Berechnung der Effektstärke notwendig waren. Hierbei wurden ebenso zuerst die Korrespondenzautoren/innen kontaktiert. Wurden die Daten nicht zur Verfügung gestellt, so wurde versucht, die Effektstärke aus den vorhandenen statistischen Daten (z. B. Gruppengrößen in Zusammenhang mit berichteten *F*- oder *t*-Werten, Signifikanzwerte oder die Mittelwertdifferenz) zu ermitteln. Die entsprechenden Formeln finden sich in Borenstein (2009) und Higgins und Green (2011). War eine Umrechnung nicht möglich, so wurden Studien ausgeschlossen, bei denen gemäß AXIS ein hohes Risiko einer selektiven Berichterstattung bestand (d. h. Daten wurden bewusst nicht berichtet).

3.4 Auswertung

Die Auswertung der Daten erfolgte unter R (R Core Team, 2017) mit dem Paket *metafor* (Viechtbauer, 2010). Für jeden Datensatz bzw. Level-Skala-Kombination (d. h. RI/ÜF, RI/UF, RI/TA, BZ/ÜF, BZ/UF, BZ/TA) erfolgte eine separate Auswertung. Insgesamt wurden somit sechs identische Analysen durchgeführt. Jede Analyse umfasste eine multivariate Mehrebenen-Metaanalyse sowie Bias- und Heterogenitätsprüfung. Eine gleichzeitige Auswertung aller drei Levels (d. h. ÜF, UF, TA) mittels eines metaanalytischen Strukturgleichungsmodells (M. W. L. Cheung, 2013; MacCallum & Austin, 2000) war aufgrund der unterschiedlichen Outcomes zwischen den Studien nicht möglich. Einen Überblick über das gesamte Auswertungsverfahren der Datensätze zeigt Abbildung 3.

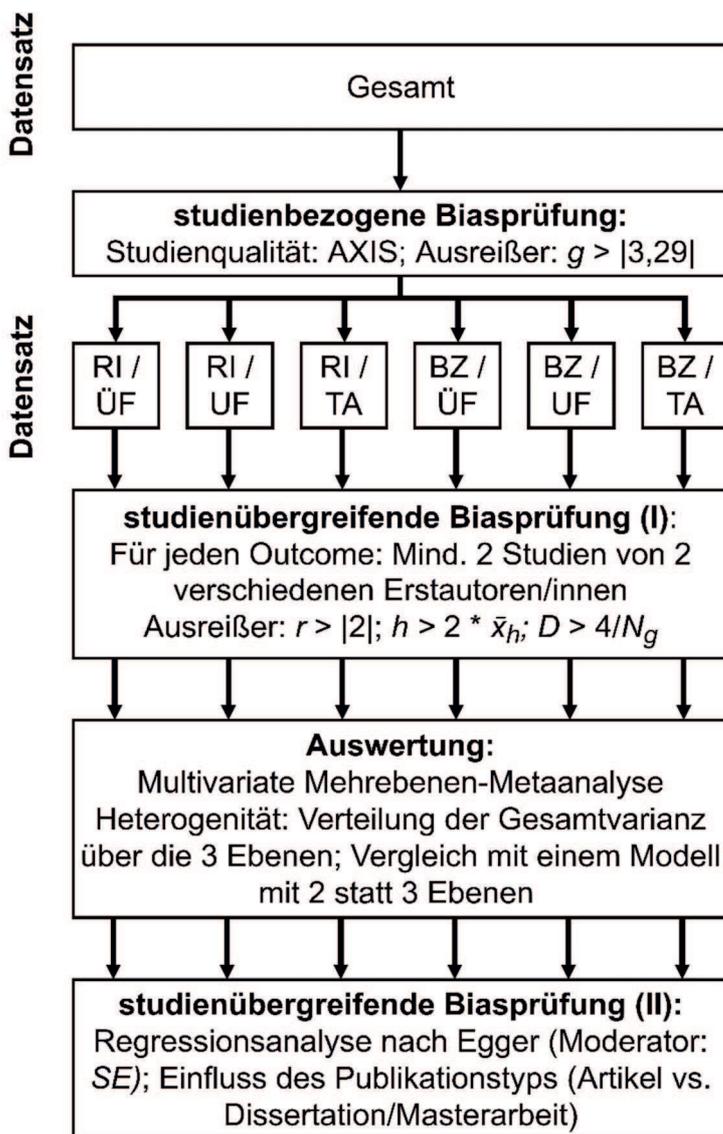


Abbildung 3. Überblick über das Auswertungsprocedere.

RI = Richtigkeit; BZ = benötigte Zeit; ÜF = übergeordnete Fähigkeit; UF = untergeordnete Fähigkeit; TA = Testaufgabe; g = Hedges' g ; r = standardisierte Residuen, h = hat values, D = Cook's distance; SE = Standardfehler.

3.4.1 Ermittlung der Effektstärken

Als Effektstärke wurde der standardisierte Mittelwertunterschied zwischen der Gruppe der Personen mit Rechenstörung (mR) und der Gruppe der Personen ohne Rechenstörung (oR), wiedergegeben durch Hedges' g (Hedges, 1981), berechnet. Hedges' g basiert auf Cohen's d (Cohen, 1969, 1988) und korrigiert es um den Faktor J , da d bei geringen Stichprobengrößen zur Überschätzung der Effektstärke neigt (Hedges, 1981). Die Berechnung erfolgte durch:

$$g = J \times d$$

mit

$$d = \frac{\bar{X}_{oR} - \bar{X}_{mR}}{SD_{pool}}$$

,

$$SD_{pool} = \sqrt{\frac{(N_{oR} - 1) * SD_{oR}^2 + (N_{mR} - 1) * SD_{mR}^2}{N_{oR} + N_{mR} - 2}}$$

und

$$J = 1 - \frac{3}{4(N_{oR} + N_{mR} - 2) - 1}$$

N_{mR} bzw. N_{oR} ist die Anzahl an Personen in der Gruppe mit bzw. ohne Rechenstörung. \bar{X}_{mR} bzw. \bar{X}_{oR} ist der Mittelwert und SD_{mR}^2 bzw. SD_{oR}^2 die Varianz der Personen in der Gruppe mit bzw. ohne Rechenstörung in Bezug auf den jeweiligen Outcome. Größere Effektstärken bedeuteten gemäß dieser Berechnung einen Gruppenunterschied zugunsten der Personen ohne Rechenstörung (d. h. höhere Richtigkeit oder geringere benötigte Zeit).

3.4.2 Multivariate Mehrebenen-Metaanalyse

Metaanalysen fassen die Effekte einzelner Studien zusammen und schätzen einen gewichteten Durchschnittseffekt, der dem wahren Effekt möglichst nahe ist. Im Rahmen dieser Arbeit basierte die Schätzung auf einem Modell mit zufälligen Effekten (d. h. *random-effects model*). Bei diesem wird angenommen, dass die einzelnen Effekte innerhalb und zwischen den Studien nicht identisch sind bzw. komplett dem wahren Effekt entsprechen, sondern studienbezogen variieren können (z. B. aufgrund unterschiedlicher Stichprobenszusammensetzung) (Raudenbush, 2009).

Die meisten der eingeschlossenen Studien berichteten mehrere Outcomes sowie multiple Effektstärken zu einzelnen Outcomes. Dies führte dazu, dass Effektstärken innerhalb sowie zwischen den Studien korrelierten, weswegen sie auch als *abhängige Effektstärken* bezeichnet werden (Gleser & Olkin, 2009). Innerhalb einer Studie treten abhängige Effektstärken auf, wenn bei derselben

Stichprobe entweder mehrere Effektstärken zum gleichen Outcome berichtet werden oder verschiedene Outcomes von denselben Personen stammen (Scammacca, Roberts & Stuebing, 2014). Diese Effektstärken und die darauf basierenden Outcomes sind sich folglich aufgrund ihrer identischen Stichproben im Vergleich zu den Effektstärken und Outcomes anderer Studien ähnlicher. Zwischen den Studien kann sich trotz unterschiedlicher Stichproben ebenso eine Abhängigkeit der Effektstärken bei einem Outcome ergeben, wenn die Rahmenbedingungen und Operationalisierungen einiger Studien sich im Vergleich zu anderen mehr ähneln. Dies ist bei gleichen Autoren/innengruppen oder dem gleichen Setting (z. B. Land der Studiendurchführung) der Fall (Becker, 2000). Auch können die Höhen der Effektstärken zu einem Outcome von den Einschlusskriterien der Studien abhängen oder höhere Werte der Effektstärken in einem Outcome gehen mit höheren Werten der Effektstärken in einem anderen Outcome einher (Jackson, Riley & White, 2011). Für die Auswertung war daher zu berücksichtigen, dass abhängige Effektstärken aufgrund ihrer Korrelation gemeinsame Anteile an der Gesamtvarianz der Daten erklärten (Hedges, Tipton & Johnson, 2010). Eine Nichtbeachtung der Korrelation hätte demnach in der Metaanalyse zu einer Überschätzung der Standardfehler geführt und das Risiko eines Fehlers 2. Art erhöht (Moeyaert et al., 2017; Riley, 2009). Da die Korrelationen der Effektstärken innerhalb einer Studie selten berichtet werden und die Korrelationen zwischen Effektstärken verschiedener Studien nicht berechenbar sind, war es notwendig, diese näherungsweise zu bestimmen.

Eine Möglichkeit hierzu stellte die Mehrebenen-Metaanalyse dar (M. W. L. Cheung, 2014). Diese wurde zusätzlich um ein multivariates Vorgehen erweitert (M. W. L. Cheung, 2015), sodass mehrere Outcomes gleichzeitig analysiert und deren Korrelationen innerhalb sowie zwischen den Studien berücksichtigt werden konnten (Mavridis & Salanti, 2013; van Houwelingen, Arends & Stijnen, 2002). Die so durchgeführte multivariate Mehrebenen-Metaanalyse operationalisiert die Abhängigkeit der Effektstärken durch drei Arten von Varianzen gemäß ihrer Ebenen: die Varianz der jeweiligen Effektstärke auf der ersten Ebene wiedergegeben durch den Standardfehler der Effektstärke, die Varianz zwischen den Effektstärken innerhalb einer Studie auf der zweiten Ebene und die Varianz zwischen den Studien auf der dritten Ebene. Kurz gefasst können die Effekt-

stärken somit zwischen den Testpersonen einer Studie (Ebene 1), den Outcomes einer Studie (Ebene 2) und den Studien selbst (Ebene 3) variieren (Assink & Wibbelink, 2016). Die geschätzte Varianz auf der zweiten Ebene erklärt hierbei die Korrelation zwischen den Effektstärken einer Studie, ohne dass hierzu eine genaue Angabe oder Schätzung der Korrelation notwendig ist (van den Noortgate, Lopez-Lopez, Marin-Martinez & Sanchez-Meca, 2013, 2014). Eine stringente Grundannahme des Modells ist eine gleiche Heterogenität der Varianzen der Effektstärken bzw. Outcomes. Wenngleich dies nicht für jeden Outcome der Fall war, so erweist sich das Modell als robust bei Verletzungen dieser Prämisse (S. F. Cheung & Chan, 2014; van den Noortgate et al., 2013).

Als Schätzverfahren für die Parameter des Modells wurde die *restricted maximum likelihood (REML)* Methode verwendet, die zusätzlich um eine Anpassung nach Knapp und Hartung (2003) korrigiert wurden, welche das Risiko eines Fehlers 1. Art bei Metaregressionen reduziert (Higgins & Thompson, 2004).

3.4.3 Biasprüfung

Vor und nach Auswertung der Daten erfolgte eine Überprüfung des Risikos eines Bias innerhalb der einzelnen Studien sowie über alle Studien hinweg. Im Fokus standen dabei ein *reporting bias* sowie ein Bias hinsichtlich der Studienqualität. Ein reporting bias geht davon aus, dass sich Ergebnisse durch die Art ihrer Publikation (z. B. Veröffentlichung in einem Journal, Ausschluss nicht-signifikanter Ergebnisse) und mit der Publikation indirekt involvierter Faktoren (z. B. Autoren/innen, Sprache der Publikation) voneinander unterscheiden. In der Folge kommt es zum Beispiel zu einer Überrepräsentation von ähnlichen und / oder signifikanten Studien mit hohen Effekten (Dwan, Gamble, Williamson, Kirkham & Reporting Bias Group, 2013; Pigott, Valentine, Polanin, Williams & Canada, 2013). Ein Bias hinsichtlich der Studienqualität meint, dass Mängel im Studiendesign und / oder in der Studiendurchführung zu verzerrten Ergebnissen führen können. Ebenso wurde im Rahmen der Biasprüfung auf das Vorhandensein nicht plausibler Effektstärken getestet, die, auch im Vergleich zu den restlichen Effektstärken, als Ausreißer zu bewerten sind (siehe Abbildung 3).

Vor der Auswertung erfolgte zum einen eine studienbezogene Biasprüfung am gesamten Datensatz. Hierbei wurden Studien ausgeschlossen, die gemäß der Checkliste AXIS eine geringe Studienqualität und ein hohes Risiko einer selektiven Berichterstattung aufwiesen. Ebenso wurden Effektstärken ausgeschlossen, deren Absolutwert über 3,29 lag und gemäß Tabachnick und Fidell (2007) als Ausreißer gelten. Effektstärken dieser Höhe entsprechen bei zweiseitiger Testung einem Signifikanzwert von kleiner 0,001, weswegen methodische Artefakte, die die Höhe der Effektstärke bedingen, nicht vollständig auszuschließen sind. Artefakte sind zum Beispiel Tippfehler im Studientext oder bei der Kodierung (Schmidt, Le & Oh, 2009). Diese weisen nicht direkt auf Mängel in der Studienqualität hin, können jedoch den Gesamteffekt einer Metaanalyse verzerren.

Zum anderen wurde vor der Auswertung die erste von zwei studienübergreifenden Biasprüfungen durchgeführt. Dies geschah separat für jeden der 6 Datensätze (d. h. RI/ÜF, RI/UF, RI/TA, BZ/ÜF, BZ/UF, BZ/TA). Hierbei wurden in einem ersten Schritt die Testaufgaben, untergeordneten Fähigkeiten und übergeordneten Fähigkeiten ausgeschlossen, deren Effektstärken nicht von mindestens zwei Studien stammten, die überdies unterschiedliche Erstautoren/innen hatten. Zwei Studien sind zur Durchführung einer Metaanalyse akzeptabel, wenngleich die Ergebnisse der einzelnen Outcomes in Bezug zur eingeschlossenen Studienanzahl zu berücksichtigen waren (Valentine, Pigott & Rothstein, 2010). Ein möglicher Publikationsbias, bedingt durch ähnliche Studiendurchführung und -ergebnisse, wurde durch die Bedingung unterschiedlicher Erstautoren/innen reduziert (Ebrahim et al., 2016; von Elm, Poglia, Walder & Tramer, 2004). Als zweiter Schritt der studienübergreifenden Biasprüfung vor der Auswertung wurde für jeden Datensatz Ausreißer bei den Effektstärken ermittelt, deren Einfluss auf den ermittelten Gesamteffekt der Metaanalyse zu hoch war. Dies geschah für jede Effektstärke anhand der folgenden 3 Kriterien:

- (1) standardisierte (intern studentisierte) Residuen r jeder Effektstärke kleiner -2 bzw. größer 2: $r > |2|$
- (2) hat values h jeder Effektstärke größer als der zweifache Durchschnitt aller hat values: $h > 2 * \bar{X}_h$

- (3) Cook's distance D jeder Effektstärke größer als 4 geteilt durch die Anzahl an Effektstärken g : $D > 4/N_g$

Standardisierte Residuen geben auf Basis eines Regressionsmodells den standardisierten Unterschied zwischen den beobachteten und erwarteten Werten wieder. Beispielsweise bedeutet ein standardisiertes Residuum von 0,9, dass sich der erwartete Wert einer Effektstärke vom beobachteten Wert um 0,9 Standardabweichungen unterscheidet. Werte kleiner als -2 bzw. größer als 2 fallen außerhalb eines 95 % Konfidenzintervall und gelten demnach als Ausreißer (Stevens, 1984). Hat values (Belsley, Kuh & Welsch, 1980) sowie Cook's distance (Cook & Weisberg, 1982) ermitteln den Einfluss einzelner Effektstärken auf den Gesamteffekt (Viechtbauer & Cheung, 2010). Hat values beziehen sich auf die Distanz zwischen einer Effektstärke und dem Mittelwert aller Effektstärken. Je weiter eine Effektstärke vom Mittelwert aller Effektstärken entfernt liegt, umso größer ist im Rahmen eines Regressionsmodells ihr Einfluss darauf. Hat values können sich somit (als slope) nur zwischen 0 und 1 bewegen und werden, wenn sie größer als der zweifache Durchschnitt aller hat values sind, als kritisch betrachtet (Belsley et al., 1980). Cook's distance gibt an, wie sehr sich in einem Regressionsmodell die Vorhersage der Effektstärken ändert, wenn einzelne Effektstärken vom Modell entfernt werden. Für Cook's distance gibt es verschiedene kritische Werte (z. B. pauschal $D > 1$). Verwendet wurde ein Wert von Kenneth und Robert (1985), der sich an der Anzahl eingeschlossener Parameter orientiert. So wurden Effektstärken als kritisch eingestuft, wenn Cook's distance größer als 4 geteilt durch die Anzahl aller Effektstärken war.

Nach der Auswertung erfolgte die zweite studienübergreifende Biasprüfung. Hierbei wurde eine Regressionsanalyse nach Egger angewandt, die den Einfluss der Standardfehler auf die Effektstärken testet (Egger, Smith, Schneider & Minder, 1997). Als Nullhypothese hierbei gilt, dass sich Studien mit kleinen und großen Stichproben bzw. großen und kleinen Standardfehlern nicht systematisch hinsichtlich der Höhe ihrer Effektstärken unterscheiden. Ein signifikantes Ergebnis bedeutet demzufolge, dass Studien mit einer ähnlich hohen Effektstärke überrepräsentiert sind bzw. Studien mit einer entsprechend niedrigen Effektstärke fehlen. Im Fokus steht dabei insbesondere ein Publikationsbias,

wonach Studien mit signifikanten und hohen Effektstärken häufiger publiziert werden als Studien mit niedrigen und nicht-signifikanten Effektstärken (Rücker, Carpenter & Schwarzer, 2011). Eggers Regressionsanalyse neigt jedoch, insbesondere bei einer geringen Studienanzahl, zu Fehlern 1. Art, weswegen ein signifikantes Ergebnis nicht zwangsläufig einen Publikationsbias bedeutete und in Bezug zur Datengrundlage bewertet werden musste (Sutton, 2009). Grafische Methoden der Biasprüfung bei Metaanalysen, insbesondere *Funnel-Plots*, wurden nicht verwendet, da sie die statistische Abhängigkeit von Effektstärken nicht mitberücksichtigen und sich interpretativ als nicht eindeutig erwiesen haben (Jackson et al., 2011; Terrin, Schmid & Lau, 2005). Abschließend wurde bei der zweiten studienübergreifenden Biasprüfung anhand einer Subgruppenanalyse der moderierende Effekt des Publikationstyps auf die ermittelten Effektstärken der Metaanalyse getestet. Dabei wurde geprüft, ob der Publikationstyp (Artikel vs. Dissertation/Masterarbeit) der zugrundeliegenden Studien mit einer systematisch anderen Höhe der Effektstärken einherging. Die Grundannahme hierzu war, dass Studien, die als Artikel in einem Journal erschienen waren, weniger nicht-signifikante Ergebnisse berichteten, um den Publikationsprozess zu vereinfachen (Pigott et al., 2013).

3.4.4 Heterogenität

Heterogenität gibt das Ausmaß an, wie sehr die beobachteten bzw. eingeschlossenen Effektstärken voneinander abweichen (Borenstein, Hedges, Higgins & Rothstein, 2009). Im Rahmen einer Mehrebenenanalyse wurde hierzu die Verteilung der Gesamtvarianz σ^2 auf die drei Ebenen (d. h. Varianz der Effektstärken, Varianz innerhalb bzw. zwischen den Studien) überprüft (M. W. L. Cheung, 2014). Der Anteil der Varianz jeder Ebene an der Gesamtvarianz war dabei durch I^2 gegeben (Higgins & Thompson, 2002). Im Zuge dessen wurde getestet, ob die vorliegende Heterogenität alleine durch die Varianz der Effektstärken (d. h. Ebene 1) erklärt werden konnte oder ob ein Teil der Gesamtvarianz auf Faktoren auf der zweiten und dritten Ebene zurückzuführen war. Die Prüfung erfolgte jeweils für die zweite und dritte Ebene mittels eines einseitigen Log-Likelihood-Ratio-Tests. Bei diesem wurde das generelle Modell, bei dem beide Varianzen der zweiten und dritten Ebene frei variieren konnten, mit einem restriktiven Modell, bei dem die Varianz der zweiten oder der dritten Ebene auf 0 fixiert war, verglichen. Dabei wurde getestet, ob das Modell mit

fixierter Varianz die Daten nicht genauso gut erklärte (Assink & Wibbelink, 2016). Ein signifikantes Ergebnis bedeutete demzufolge einen besseren statistischen Fit des Modells mit drei Ebenen, was auf eine vorliegende Heterogenität der Daten auf der zweiten und dritten Ebene schließen ließ, die nicht alleine durch die Varianz der einzelnen Effektstärken erklärt werden konnte. In einem solchen Fall kamen Faktoren für die vorhandene Varianz in Betracht, die mit der jeweiligen Stichprobe (z. B. Alter) oder Methodik (z. B. verwendete Tests zur Erfassung der Mathematikleistung) der Studien zusammenhängen.

Der moderierende Einfluss solcher Faktoren auf die Effektstärken kann für kategoriale Variablen durch Subgruppenanalysen und für kontinuierliche Variablen durch Metaregressionen getestet werden. Im Rahmen dieser Arbeit erfolgte jedoch keine weitergehende Überprüfung der Heterogenität aus drei Gründen: Erstens hängt die Reliabilität und Validität von Subgruppenanalysen und Metaregressionen stark von der Qualität der jeweiligen Moderatorvariable und der Anzahl der Studien, die Daten zu dieser Variable berichten, ab. Die *Cochrance Collaboration* empfiehlt zum Beispiel pro Moderator bzw. pro Gruppe mindestens 10 Studien um eine Metaregression bzw. eine Subgruppenanalyse durchzuführen (Higgins & Green, 2011). Insbesondere Subgruppenanalysen zeigen hierbei eine geringe statistische Power, wenn die verglichenen Gruppen eine unterschiedliche Anzahl an Studien aufweisen, die sich zudem in ihrer Qualität unterscheiden (Oxman & Guyatt, 1992; Walker, Hernandez & Kattan, 2008; Wood & Eagly, 2009). Zweitens neigen Gruppen von Studien dazu, zu überlappen, wenn sie nicht auf unverwechselbaren und eindeutigen Kriterien fußen (z. B. Altersgruppen, Geschlecht). Beispielsweise überlappen im Rahmen der Rechenstörung Gruppen von Studien, wenn sie lediglich danach getrennt werden, welchen Schwerpunkt der Test hat, mit dem die Rechenstörung diagnostiziert wurde (z. B. Grundrechenarten, mathematisches Schlussfolgern). Da ein Test somit auch zu einem geringeren Teil andere Inhalte erfasst, ist es anhand dieses Kriteriums nicht möglich, Gruppen von Studien eindeutig voneinander zu trennen, wodurch Unterschiede nicht entdeckt werden bzw. die Gefahr eines Fehler 2. Art besteht. Und drittens ist das Risiko eines Bias und damit falscher Schlussfolgerungen bei Metaregressionen und Subgruppenanalysen deutlich erhöht. Obwohl ihre Datenlage auf einer systematischen Literaturrecherche basiert, handelt es sich bei beiden Methoden nur um Beobachtungsstudien bzw.

Querschnittsstudien, da die Zuordnung der Studien zu den Gruppen bzw. Moderatoren nicht zufällig erfolgt. Es ist daher nicht auszuschließen, dass ein möglicher Gruppenunterschied durch andere, unbekannte Variablen bedingt ist als der eigentlichen Moderatorvariable (Higgins & Green, 2011).

Die Heterogenität der Effekte basierend auf den Gruppenunterschieden zwischen Personen mit und ohne Rechenstörung wurde bisher nur in einer Metaanalyse von Peng et al. (2018) untersucht. Dabei wurden Subgruppenanalysen und Metaregressionen angewandt, um unter anderem den moderierenden Einfluss einer vorhandenen Lesestörung, des Inhalt des Mathematiktests zur Rechenstörungsdiagnose sowie des Schweregrads einer Rechenstörung auf die ermittelten Effektstärken verschiedener Outcomes zu überprüfen. Nur eine vorhandene Lesestörung konnte hierbei einen Teil der Heterogenität dahingehend erklären, dass Personen mit komorbider Rechen- und Lesestörung größere Defizite zeigten als Personen mit isolierter Rechenstörung. Die Ergebnisse für die restlichen Moderatoren waren nicht eindeutig, auch weil die gewählten Gruppen inhaltlich überlappten (z. B. Inhalt des Mathematiktests) oder Analysen aufgrund unzureichender Daten entweder nicht für jeden Outcome durchgeführt werden konnten oder unterpoweriert waren.

Da auch diese Metaanalyse mit vielfach unterpowerierten Analysen aufgrund sich überlappender Gruppen und, bedingt durch ihre inhaltlich breite Konzeption, einer zu geringen Studienanzahl je Moderatorvariable konfrontiert gewesen wäre, wurden keine Metaregressionen oder Subgruppenanalysen mit Ausnahme derer im Rahmen der Biasprüfung (d. h. Eggers Regressionsanalyse, Publikationstyp) berechnet. Stattdessen wurde versucht, die Heterogenität zwischen den Studien bereits vor der Auswertung bestmöglich zu reduzieren, indem spezifische Einschlusskriterien definiert (z. B. Ausschluss Lesestörung) sowie ein studienbasiertes Bottom-up-Kodiervorgehen und eine umfassende Biasprüfung durchgeführt wurden.

4 Ergebnisse

4.1 Suchverlauf und Studienbeschreibung

Abbildung 4 zeigt das Flussdiagramm zur Literaturrecherche und Studienauswahl gemäß Moher, Liberati, Tetzlaff, Altman und The PRISMA Group (2009). Insgesamt konnten in der Literaturrecherche 2808 potenziell relevante Studien identifiziert werden. Nach Löschung aller Duplikate und Anwendung der Einschlusskriterien reduzierte sich die Studienanzahl auf 56. Bei der folgenden studienübergreifende Biasprüfung zeigten 2 Studien eine geringe Studienqualität und wurden für die weitere Auswertung ausgeschlossen. Der so verbliebene Gesamtpool an 54 Studien teilte sich daraufhin in die 6 verschiedenen Datensätze auf (d. h. RI/ÜF, RI/UF, RI/TA, BZ/ÜF, BZ/UF, BZ/TA). Für jeden Datensatz erfolgte die studienbezogene Biasprüfung (I) und, je nach Datensatz, wurden 1 bis 3 Studien ausgeschlossen. In die Auswertung eingeschlossen wurden schließlich je nach Datensatz 47 Studien für RI/ÜF, 46 Studien für RI/UF, 45 Studien für RI/TA, 31 Studien für BZ/ÜF, 30 Studien für BZ/UF und 24 Studien für BZ/TA.

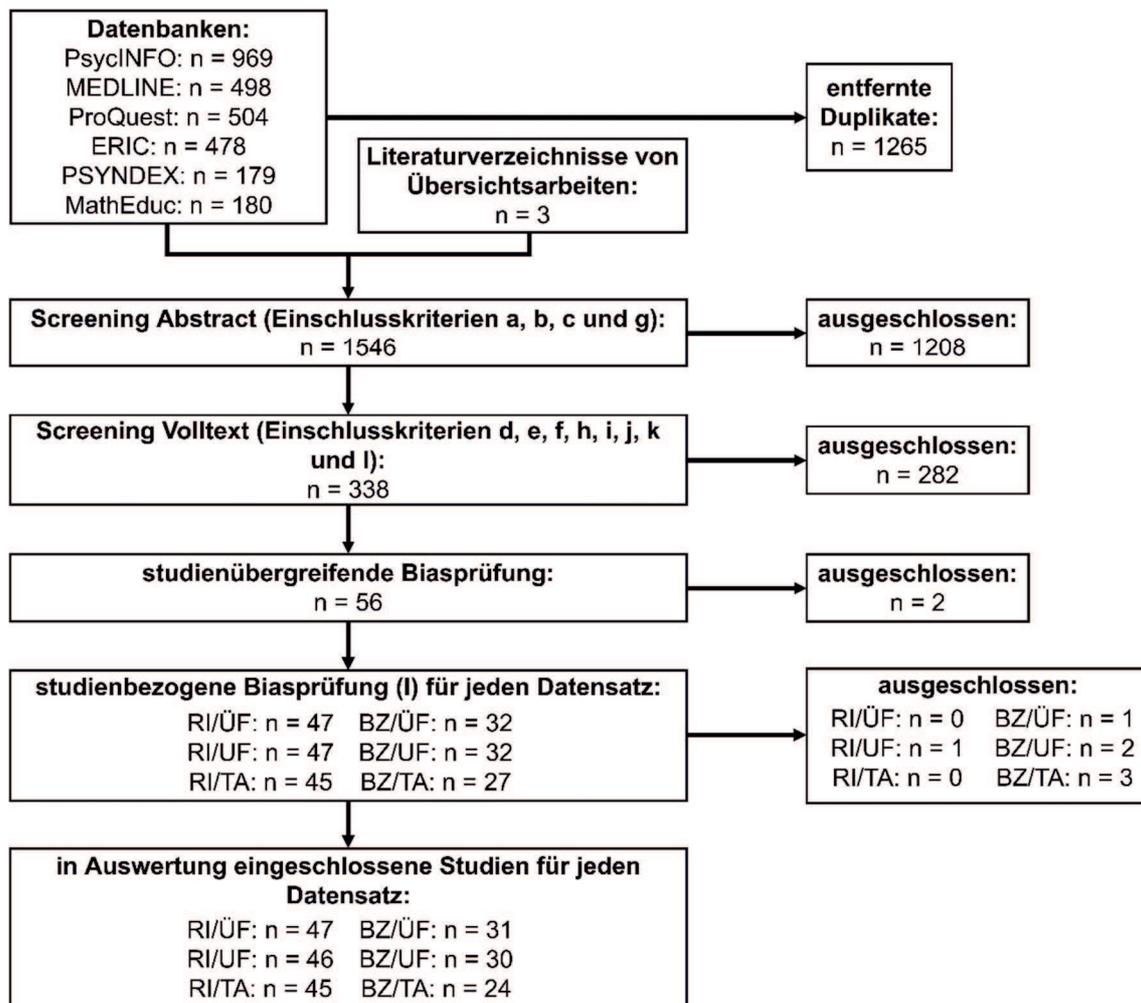


Abbildung 4. Flussdiagramm zum Suchverlauf und Anwendung der Einschlusskriterien.

RI = Richtigkeit; BZ = benötigte Zeit; ÜF = übergeordnete Fähigkeit; UF = untergeordnete Fähigkeit; TA = Testaufgabe.

Über alle Datensätze hinweg wurden 54 Studien mit insgesamt 471 Effektstärken eingeschlossen. 44 Studien waren Journalartikel und bei 10 Studien handelte es sich um Dissertationen oder Masterarbeiten. Die Gesamtgruppe der Personen mit Rechenstörung umfasste 1530 Testpersonen (668 männlich) mit einem Durchschnittsalter von 128,28 Monaten. Die Gesamtgruppe der Personen ohne Rechenstörung umfasste 2922 Testpersonen (1365 männlich) mit einem Durchschnittsalter von 126,81 Monaten. Die homogene Verteilung der Gruppen hinsichtlich Alter und Geschlecht war bedingt durch die Einschlusskriterien dieser Metaanalyse, die ein Matching dieser Variablen innerhalb der Studien vorgaben. Der Unterschied in den Stichprobengrößen war auf die beiden Studien von Compton et al. (2012) zurückzuführen, bei der 9- bzw. 17-mal mehr

Personen ohne Rechenstörung eingeschlossen waren als Personen mit Rechenstörung. Die deskriptiven Studienangaben zu den Studien- und Stichprobenmerkmalen sowie die genaue Zuordnung der Studien zu den jeweiligen Datensätzen ist Tabelle 4 zu entnehmen.

Tabelle 4

Deskriptive Studienangaben zu Studien- und Stichprobenmerkmalen

Studie	PT	mit Rechenstörung			ohne Rechenstörung			Datensatz					
		<i>n</i>	\bar{X}_{Alter}	% männl.	<i>n</i>	\bar{X}_{Alter}	% männl.	RI/ ÜF	RI/ UF	RI/ TA	BZ/ ÜF	BZ/ UF	BZ/ TA
Ashkenazi, Mark-Zigdon und Henik (2013)	A	11	114,1	2	11	113,8	2	X	X	X			
Attout und Majerus (2015)	A	16	113,7	6	16	120,5	6	X	X	X	X	X	X
Censabella und Noël (2008)	A	20	125,9	10	20	124,0	10	X	X	X	X	X	X
Ceulemans et al. (2014)	A	18	172,0	5	28	173,2	8	X	X	X	X	X	X
Cirino et al. (2015)	A	105	92,9	47	403	89,8	186	X	X	X	X	X	X
Compton et al. (2012) (Studie 1)	A	56	NA	30	476	NA	239	X	X	X			
Compton et al. (2012) (Studie 2)	A	32	NA	17	537	NA	270	X	X	X			
de Weerd, Desoete und Roeyers (2013)	A	22	117,6	6	45	120,9	19	X	X	X			
Donker et al. (2016)	A	31	107,4	7	34	102,8	12				X	X	
Evans (2008)	D/M	65	122,5	31	72	116,6	29	X	X	X	X	X	
Gold et al. (2013)	A	17	112,4	13	23	111,4	13				X	X	
Heine et al. (2013)	A	20	99,2	10	20	98,4	9	X	X	X	X	X	X
Jolles et al. (2016)	A	19	106,9	6	19	105,0	10	X	X	X			
Koontz und Berch (1996)	A	16	124,0	7	16	123,0	7	X	X	X	X	X	
Kucian (2005)	D/M	9	121,2	4	10	110,4	5	X	X	X	X	X	X
MacKinnon McQuarrie, Siegel, Perry und Weinberg (2014)	A	39	79,2	18	44	78,7	23	X	X	X	X	X	
Mähler und Schuchardt (2016)	A	18	102,5	5	31	107,1	15	X	X	X			
Mammarella, Caviola, Giofrè und Szűcs (2018)	A	24	117,4	14	24	117,6	10	X	X	X			
Mammarella, Hill, Devine, Ca-	A	24	NA	12	23	NA	12	X	X	X			

Studie	PT	mit Rechenstörung			ohne Rechenstörung			Datensatz						
		<i>n</i>	\bar{X}_{Alter}	% männl.	<i>n</i>	\bar{X}_{Alter}	% männl.	RI/ ÜF	RI/ UF	RI/ TA	BZ/ ÜF	BZ/ UF	BZ/ TA	
viola und Szűcs (2015)														
McCall (1999)	D/M	11	157,2	5	30	158,4	14	X	X	X	X	X	X	X
McDonald (2016) (Studie 1)	D/M	20	116,1	7	20	117,8	11	X	X	X	X	X	X	X
McDonald (2016) (Studie 2)	D/M	25	111,1	10	27	112,7	15	X	X	X	X	X	X	X
Mejias, Mussolin, Rousselle, Gregoire und Noël (2012)	A	23	118,4	9	23	118,1	14	X	X					
Micallef und Prior (2004)	A	15	132,2	6	28	140,4	17	X	X	X	X	X	X	X
Moll et al. (2015)	A	17	112,8	6	32	107,5	15	X	X	X	X	X	X	X
Mussolin, Mejias und Noël (2010)	A	15	121,8	9	15	123,7	7	X	X	X	X	X	X	X
Peng et al. (2012)	A	18	132,1	11	30	131,7	13	X	X	X	X	X	X	X
Powell, Fuchs, Fuchs, Cirino und Fletcher (2009)	A	109	NA	59	109	NA	55	X	X	X				
Raddatz et al. (2017)	A	20	112,6	6	40	111,6	9	X	X	X	X	X	X	X
Rosenberg-Lee et al. (2014)	A	16	100,1	6	20	101,3	9	X	X	X	X	X	X	X
Rousselle und Noël (2008)	A	18	102,7	6	22	105,3	8	X	X	X	X	X	X	X
Schleifer und Landerl (2011) (Studie 1)	A	12	99,0	6	12	99,3	6				X	X	X	
Schleifer und Landerl (2011) (Studie 2)	A	19	111,1	5	19	111,0	10				X	X	X	
Schleifer und Landerl (2011) (Studie 3)	A	21	125,3	10	21	121,3	12				X	X	X	
Schuchardt und Mähler (2010)	A	22	102,0	6	30	108,0	15	X	X	X				
Schuchardt et al. (2008)	A	17	103,4	5	30	108,8	15	X	X	X				
Silver, Elder und DeBolt (1999)	A	14	142,8	11	13	138,0	11	X						
Swanson (2012)	A	12	193,2	8	15	188,8	8	X	X	X				

Studie	PT	mit Rechenstörung			ohne Rechenstörung			Datensatz					
		<i>n</i>	\bar{X}_{Alter}	% männl.	<i>n</i>	\bar{X}_{Alter}	% männl.	RI/ ÜF	RI/ UF	RI/ TA	BZ/ ÜF	BZ/ UF	BZ/ TA
Swanson (2014)	A	59	105,5	34	88	105,4	38	X	X	X			
Swanson, Lussier und Orosco (2013a)	A	71	103,2	39	49	104,4	26	X	X	X	X	X	
Swanson, Lussier und Orosco (2013b)	A	100	106,53		92	106,1		X	X	X			
Szanto (1997)	D/M	24	370,56	12	24	332,0	12	X	X	X	X		
Szűcs et al. (2013)	A	12	110	8	12	109,0	5	X	X	X	X	X	X
Torbeyns, Verschaffel und Ghesquiere (2004)	A	25	88,52	8	26	88,2	14	X	X	X	X	X	X
van Daal, van der Leij und Adèr (2013)	A	34	168,12	10	40	164,9	9	X	X	X			
van der Sluis et al. (2004)	A	18	128,72	5	19	127,4	9				X	X	X
van der Sluis, de Jong und van der Leij (2005)	A	17	128,29	4	18	127,8	9	X	X	X			
Wang et al. (2012)	A	45	129	33	45	132,0	28				X	X	X
Willburger et al. (2008)	A	19	110,7	4	42	109,5	17	X	X	X			
A. J. Wilson et al. (2015)	A	22	348,0	10	19	336,0	7	X	X	X	X	X	X
Winegar (2013) (Studie 1)	D/M	31	NA	15	15	NA	9	X	X	X			
Winegar (2013) (Studie 2)	D/M	40	NA	21	15	NA	9	X	X	X			
Yaghoubzadeh (1995) (Studie 1)	D/M	15	131,33	13	15	123,8	12	X	X	X			
Yaghoubzadeh (1995) (Studie 2)	D/M	12	106,16	11	15	103,5	12	X	X	X			
RI/ÜF		1367	130,35	589	2749	128,82	1275						
RI/UF		1353	130,04	578	2736	128,59	1264						
RI/TA		1330	130,34	569	2713	128,86	1250						

Studie	PT	mit Rechenstörung			ohne Rechenstörung			Datensatz					
		<i>n</i>	\bar{X}_{Alter}	% männl.	<i>n</i>	\bar{X}_{Alter}	% männl.	RI/ ÜF	RI/ UF	RI/ TA	BZ/ ÜF	BZ/ UF	BZ/ TA
BZ/ÜF		800	131,69	366	1196	129,67	555						
BZ/UF		776	123,73	354	1172	122,93	543						
BZ/TA		537	127,64	239	934	127,12	433						
Gesamt		1530	128,28	668	2922	126,81	1365						

Anmerkung. PT = Publikationstyp; RI = Richtigkeit; BZ = benötigte Zeit; ÜF = übergeordnete Fähigkeit; UF = untergeordnete Fähigkeit; TA = Testaufgabe; A = Studie als Artikel in einem Journal erschienen; D/M = Studie als Dissertation oder Masterarbeit erschienen; NA = keine Angabe.

Hinsichtlich der Intelligenzleistung verwendeten 89 % aller Studien Tests, die den nonverbalen IQ oder eine Kombination aus nonverbalen und verbalen IQ erfassten. Die Kriterien zum Ausschluss einer Intelligenzminderung bewegten sich zwischen einem Mindest-IQ von 70 bis 110. 54 % aller Studien wandten dabei einen Mindest-IQ zwischen 80 und 85 an. Hinsichtlich der Leseleistung verwendeten 87 % aller Studien Tests, die die Lesegenauigkeit oder eine Kombination aus Lesegenauigkeit und Leseverständnis erfassten. Die Kriterien zum Ausschluss einer Lesestörung bewegten sich zwischen einem Prozentrang von größer oder gleich 16 (dem vorgegebenen kleinstmöglichen Prozentrang) bis größer 40. 31 % aller Studien wandten einen Prozentrang von größer oder gleich 16, 28 % von größer oder gleich 25 und 19 % von größer oder gleich 30 an. Hinsichtlich der Mathematikleistung verwendeten 91 % aller Studien Tests, die das Rechnen und Faktenwissen oder eine Kombination aus Rechnen und Faktenwissen mit den Basiskompetenzen und / oder dem mathematischen Problemlösen erfassten. Die Kriterien zur Diagnose einer Rechenstörung bewegten sich zwischen einem Prozentrang von kleiner oder gleich 2 bis kleiner oder gleich 25 (dem vorgegebenen größtmöglichen Prozentrang). 43 % aller Studien wandten einen Prozentrang von kleiner oder gleich 25 und 28 % von kleiner oder gleich 16 an. Die deskriptiven Studienangaben zu den Diagnosemerkmalen finden sich in Tabelle 5.

Tabelle 5
Deskriptive Studienangaben zu Diagnosemerkmalen

Studie	IQ		Lesen		Mathematik	
	Inhalt	Krit. ^a	Inhalt	Krit. ^b	Inhalt	Krit. ^b
Ashkenazi et al. (2013)	NV	≥ 90	LG/LV	≥ 25	B/R/S	> 24 U. Monate
Attout und Majerus (2015)	NV	> 85	LG/LV	< 12 L. Monate	R	≤ 2
Censabella und Noël (2008)	NV/V	> 85	LV	> 30	R/S	> 24 L. Monate
Ceulemans et al. (2014)	NV/V	dschn.	LG	> 25	B/R/S	< 25
Cirino et al. (2015)	NV/V	> 70	LG	≥ 25	B/R	< 25
Compton et al. (2012) (Studie 1)	NV/V	≥ 80	LG/LV	> 39	B/R	< 16
Compton et al. (2012) (Studie 2)	NV/V	≥ 80	LG	> 39	B/R	< 16
de Weerd et al. (2013)	NV/V	> 80	LG	> 25	B/S	< 11
Donker et al. (2016)	NV/V	> 70	LG/LV	≥ 16	R/S	< 16

Studie	IQ		Lesen		Mathematik	
	Inhalt	Krit. ^a	Inhalt	Krit. ^b	Inhalt	Krit. ^b
Evans (2008)	NV/V	≥ 90	LG	≥ 16	R	≤ 16
Gold et al. (2013)	NV	> 70	LG	> 30	B/R	≤ 25
Heine et al. (2013)	NV/V	≥ 85	LV	> 21	R	≤ 5
Jolles et al. (2016)	NV/V	> 80	LG/LV	> 25	B/R	≤ 25
Koontz und Berch (1996)	NV/V	dschn.	LV	≥ 30	R/S	≤ 25
Kucian (2005)	NA	dschn.	LG/LV	DSM/ ICD	B/R	DSM/ ICD
Mackinnon McQuarrie et al. (2014)	V	> 70	LG	> 40	R/S	≤ 25
Mähler und Schuchardt (2016)	NV/V	≥ 80	LG	≥ 16	B/R/S	< 16
Mammarella et al. (2018)	V	> 70	LG	> 30	B/R	≤ 16
Mammarella et al. (2015)	NV/V	≥ 85	LV	≥ 16	B/R	< 16
McCall (1999)	NV/V	≥ 80	LG	≥ 30	B/R	≤ 23
McDonald (2016) (Studie 1)	NV	> 90	LG	> 30	B/R	< 25
McDonald (2016) (Studie 2)	NV	> 90	LG	> 30	B/R	< 25
Mejias et al. (2012)	NV/V	≥ 80	LG	≥ 16	R/S	< 15
Micallef und Prior (2004)	NV/V	dschn.	LG	> 25	B/R	≤ 25
Moll et al. (2015)	NV/V	dschn.	LG	≥ 16	B/R	< 16
Mussolin et al. (2010)	NV/V	> 85	LG/LV	≥ 16	R	< 15
Peng et al. (2012)	NV	> 110	LG/LV	> 25	B/R/S	< 25
Powell et al. (2009)	NV/V	> 70	LG	> 39	B/R	≤ 25
Raddatz et al. (2017)	NV	≥ 85	LV	≥ 25	R	≤ 16
Rosenberg-Lee et al. (2014)	NV/V	> 80	LG	> 25	B/R	≤ 25
Rousselle und Noël (2008)	NV/V	≥ 85	LG/LV	≥ 16	B/R	< 15
Schleifer und Landerl (2011) (Studie 1)	NV/V	≥ 85	LG	≥ 16	R	≤ 7
Schleifer und Landerl (2011) (Studie 2)	NV/V	≥ 85	LG	≥ 16	R	≤ 7
Schleifer und Landerl (2011) (Studie 3)	NV/V	≥ 85	LG	≥ 16	R	≤ 7
Schuchardt und Mähler (2010)	NV/V	> 80	LG	> 16	B/R/S	≤ 16
Schuchardt et al. (2008)	NV/V	≥ 80	LG	≥ 16	B/R/S	< 16

Studie	IQ		Lesen		Mathematik	
	Inhalt	Krit. ^a	Inhalt	Krit. ^b	Inhalt	Krit. ^b
Silver et al. (1999)	NV/V	> 80	LG/LV	> 25	R/S	< 25
Swanson (2012)	V	> 85	LG/LV	> 25	B/R/S	< 16
Swanson (2014)	NV	≥ 94	LG/LV	≥ 35	S	≤ 25
Swanson et al. (2013a)	NV	> 90	LG/LV	> 25	S	≤ 25
Swanson et al. (2013b)	NV	≥ 94	LG/LV	≥ 35	S	≤ 25
Szanto (1997)	NV/V	≥ 80	LG	≥ 30	B/R	≤ 25
Szűcs et al. (2013)	NV	≥ 85	LG/LV	≥ 16	R	< 16
Torbeyns et al. (2004)	NV	≥ 85	LG	≥ 16	R	≤ 25
van Daal et al. (2013)	NV/V	≥ 70	LG/LV	> 16	R	≤ 16
van der Sluis et al. (2004)	V	dschn.	LG	< 15 U. Monate	R	≥ 15 U. Monate
van der Sluis et al. (2005)	V	> 94	LG	< 15 U. Monate	R	≥ 15 U. Monate
Wang et al. (2012)	NV/V	≥ 90	LV	≥ 16	NA	< 15
Willburger et al. (2008)	NV/V	≥ 85	LV	≥ 24	R	≤ 16
A. J. Wilson et al. (2015)	NV/V	≥ 85	LG	> 25	B/R	≤ 25
Winegar (2013) (Studie 1)	NV	≥ 90	LG	> 25	B/R/S	< 25
Winegar (2013) (Studie 2)	NV	≥ 90	LG	> 25	B/R/S	< 25
Yaghoubzadeh (1995) (Studie 1)	NV/V	> 80	LG	> 30	B/R	< 25
Yaghoubzadeh (1995) (Studie 2)	NV/V	> 80	LG	> 30	B/R	< 25

Anmerkung. Inhalt = Testaufgaben des Intelligenz-/Lese-/Mathematiktests; Krit. = Kriterium zum Ausschluss einer Intelligenzminderung bzw. Lesestörung und zur Diagnose einer Rechenstörung, NV = nonverbal; V = verbal; LG = Lesegenauigkeit; LV = Leseverständnis; B = Basiskompetenzen; R = Rechnen und Faktenwissen; S = mathematisches Schlussfolgern; dschn. = durchschnittliche Intelligenz; L. Monate = Lebensmonate; U. Monate = Unterrichtsmonate; DSM/ICD = Diagnose nach DSM-IV/5 oder ICD-10 liegt vor bzw. liegt nicht vor; NA = keine Angabe.

^aIQ-Wert, außer anders angegeben.

^bProzentrang, außer anders angegeben.

4.2 Ergebnisse der multivariaten Mehrebenen-Metaanalysen

Die Interpretation der Größe der Effektstärke erfolgte gemäß Cohen (1969, 1988) (d. h. kein Effekt: $g < 0,2$; kleiner Effekt: $0,2 \leq g < 0,5$; mittlerer Effekt: $0,5$

$\leq g < 0,8$; großer Effekt: $g \geq 0,8$). Positive Effektstärken bedeuteten eine bessere Leistung in der Richtigkeit bzw. der benötigten Zeit der Personen ohne Rechenstörung. Statistische Signifikanz war bei $p \leq 0,05$ gegeben. Forest-Plots wurden keine erzeugt, da sie abhängige Effektstärken nicht berücksichtigten und demzufolge keine Aussagekraft für diese Arbeit hatten. Tabelle 6 listet die ermittelten Effektstärken für sämtliche Levels (d. h. ÜF, UF, TA) sowie Skalen (d. h. RI, BZ).

Tabelle 6

Ergebnisse der multivariaten Mehrebenen-Metaanalysen aller Datensätze

Outcome (ÜF/UF/TA) ^a	Richtigkeit						benötigte Zeit					
	ST	ES	<i>g</i>	<i>SE</i>	<i>t</i>	95% KI	ST	ES	<i>g</i>	<i>SE</i>	<i>t</i>	95% KI
domänenspezifische Fähigkeiten												
Mathematik	28	133	0,75***	0,05	14,91	[0,65; 0,85]	19	74	0,68***	0,09	7,27	[0,50; 0,87]
Rechnen	13	39	0,73***	0,07	9,87	[0,58; 0,88]	8	14	0,62***	0,14	4,28	[0,33; 0,90]
Grundrechenarten	6	12	0,95***	0,12	8,13	[0,72; 1,17]	4	7	1,09***	0,22	5,02	[0,66; 1,52]
Schätzen	2	2	1,32***	0,36	3,68	[0,61; 2,03]						
Faktenwissen	8	25	0,59***	0,08	7,31	[0,43; 0,75]	4	6	0,29	0,2	1,43	[-0,11; 0,7]
Mathematisches Schlussfolgern	10	16	1,03***	0,10	10,48	[0,84; 1,22]						
Textaufgaben	7	10	0,98***	0,11	9,27	[0,77; 1,18]						
Basiskompetenzen	18	78	0,64***	0,06	10,51	[0,52; 0,76]	14	59	0,68***	0,10	6,48	[0,47; 0,89]
Simultanerfassung (bis 3)							6	6	0,77***	0,20	3,74	[0,36; 1,17]
Simultanerfassung (ab 4)	4	4	0,81***	0,22	3,78	[0,39; 1,24]	6	6	0,97***	0,21	4,69	[0,56; 1,38]
Mengenvergleich	8	31	0,49***	0,09	5,28	[0,31; 0,68]	5	24	0,51**	0,16	3,23	[0,20; 0,83]
Zahlenvergleich	5	14	0,26*	0,13	2,06	[0,01; 0,51]	5	15	0,73***	0,16	4,58	[0,41; 1,05]
Zahlen-Mengen-Vergleich	2	2	0,88***	0,23	3,76	[0,42; 1,34]						
Zahlenstrahl	6	8	0,67***	0,14	4,98	[0,41; 0,94]						
Zählen	2	3	0,33	0,23	1,44	[-0,12; 0,78]	2	2	0,54	0,33	1,62	[-0,12; 1,2]
Transkodieren	3	5	1,00***	0,19	5,39	[0,63; 1,36]						
domänenübergreifende Fähigkeiten												
Arbeitsgedächtnis	26	59	0,55***	0,06	8,92	[0,43; 0,67]						
Arbeitsgedächtniskapazität	26	59	0,56***	0,06	9,65	[0,44; 0,67]						
Spanne rückwärts	14	21	0,43***	0,09	4,97	[0,26; 0,60]						

Outcome (ÜF/UF/TA) ^a	Richtigkeit						benötigte Zeit					
	ST	ES	<i>g</i>	<i>SE</i>	<i>t</i>	95% KI	ST	ES	<i>g</i>	<i>SE</i>	<i>t</i>	95% KI
Spanne komplex	18	34	0,58***	0,07	8,35	[0,45; 0,72]						
Kurzzeitgedächtnis (KZG)	23	81	0,50***	0,06	8,34	[0,38; 0,62]						
Auditives KZG	22	53	0,37***	0,06	5,85	[0,25; 0,50]						
Spanne vorwärts (auditiv)	21	44	0,35***	0,06	5,50	[0,23; 0,48]						
Abrufen (auditiv)	5	5	0,40*	0,18	2,23	[0,05; 0,75]						
Visuell-räumliches KZG	11	28	0,79***	0,09	8,75	[0,61; 0,96]						
Spanne vorwärts (visuell-räumlich)	8	10	0,77***	0,13	5,76	[0,51; 1,04]						
Abrufen (visuell-räumlich)	6	13	0,89***	0,12	7,42	[0,66; 1,13]						
Aufmerksamkeit	9	25	0,44***	0,10	4,33	[0,24; 0,63]	8	26	0,69***	0,13	5,23	[0,43; 0,96]
Inhibition	7	20	0,41***	0,11	3,71	[0,19; 0,63]	7	19	0,65***	0,15	4,39	[0,35; 0,94]
Zufallsgenerierung	4	7	0,38*	0,16	2,39	[0,07; 0,70]						
Stroop-Effekt	3	11	0,30*	0,15	2,03	[0,01; 0,60]	6	17	0,71***	0,16	4,54	[0,40; 1,03]
Shifting							3	5	0,97***	0,22	4,36	[0,53; 1,41]
Buchstaben-und-Zahlen-Folgen							4	5	0,80**	0,24	3,35	[0,33; 1,28]
Updating	2	3	0,34	0,23	1,51	[-0,10; 0,78]						
Verarbeitungsgeschwindigkeit (VG)	11	18	0,50***	0,11	4,46	[0,28; 0,73]	7	11	0,28	0,15	1,89	[-0,01; 0,57]
Visuelle VG	8	10	0,65***	0,13	4,90	[0,39; 0,91]	4	5	0,51*	0,21	2,44	[0,10; 0,93]
Buchstaben-oder-Zahlen-Folgen							3	4	0,74**	0,27	2,80	[0,22; 1,27]
Kodieren	3	3	0,64*	0,27	2,40	[0,11; 1,17]						
Visueller Vergleich	6	7	0,73***	0,14	5,07	[0,44; 1,01]						

Outcome (ÜF/UF/TA) ^a	Richtigkeit						benötigte Zeit					
	ST	ES	<i>g</i>	<i>SE</i>	<i>t</i>	95% KI	ST	ES	<i>g</i>	<i>SE</i>	<i>t</i>	95% KI
Reaktions-/Entscheidungs- geschwindigkeit	2	2	-0,37	0,31	-1,18	[-0,98; 0,24]	4	5	0,11	0,20	0,55	[-0,29; 0,94]
Reagieren auf Stimulus							3	4	0,18	0,25	0,72	[-0,31; 0,67]
Psychomotorische Ge- schwindigkeit	4	7	0,33	0,19	1,76	[-0,04; 0,70]						
Schreibgeschwindigkeit	3	6	0,24	0,20	1,23	[-0,14; 0,63]						
Raumvorstellung	6	7	0,57***	0,17	3,35	[0,23; 0,90]						
Visualisierung	4	4	0,64**	0,22	2,91	[0,21; 1,07]						
Phonologische Verarbeitung	7	18	0,33**	0,10	3,17	[0,13; 0,53]	10	18	0,35**	0,13	2,79	[0,10; 0,60]
Phonologische Bewusstheit	4	5	0,31	0,17	1,86	[-0,02; 0,64]						
Schnelles Benennen	5	12	0,28*	0,12	2,33	[0,04; 0,53]	10	18	0,39**	0,13	3,06	[0,14; 0,64]

Anmerkung. ÜF = übergeordnete Fähigkeit; UF = untergeordnete Fähigkeit; TA = Testaufgabe; ST = Anzahl an Studien; ES = Anzahl an Effektstärken; *g* = Hedges' *g*; *SE* = Standardfehler; *t* = *t*-Wert; KI = Konfidenzintervall.

^aDie Unterordnung der Outcomes entspricht den Levels der Kodierung, d. h. kein Einzug links = übergeordnete Fähigkeit (ÜF), mittlerer Einzug links = untergeordnete Fähigkeit (UF), großer Einzug links = Testaufgabe (TA).

p* < 0,05. *p* < 0,01. ****p* < 0,001.

4.2.1 Domänenspezifische Fähigkeiten: Ermittelte Effektstärken

In der übergeordneten Fähigkeit Mathematik zeigten Personen mit Rechenstörung Defizite sowohl in der Richtigkeit (RI: $g = 0,75$) als in der benötigten Zeit (BZ: $g = 0,68$). Für die untergeordneten Fähigkeiten ergaben sich in der Richtigkeit und benötigten Zeit mittlere Effekte beim Rechnen (RI: $g = 0,73$; BZ: $g = 0,62$) und bei den Basiskompetenzen (RI: $g = 0,64$; BZ: $g = 0,68$). Der Effekt in der Richtigkeit beim mathematischen Schlussfolgern war groß (RI: $g = 1,03$). Dieser beruhte fast ausschließlich auf der Testaufgabe Textaufgaben, wo sich dementsprechend eine hohe Effektstärke zeigte (RI: $g = 0,98$). Hinsichtlich der Testaufgaben, die der untergeordneten Fähigkeit Rechnen zugeordnete waren, ergaben sich große Effekte bei den Grundrechenarten (RI: $g = 0,95$; BZ: $g = 1,09$) sowie beim Schätzen (RI: $g = 1,32$). Beim Faktenwissen zeigte sich nur ein signifikanter Effekt in der Richtigkeit, der auch größer war im Vergleich zur benötigten Zeit (RI: $g = 0,59$; BZ: $g = 0,29$). Für die Testaufgaben, die der untergeordneten Fähigkeit Basiskompetenzen zugewiesen waren, waren die Effektstärken in der benötigten Zeit für die Simultanerfassung (ab 4) (BZ: $g = 0,97$) und dem Zahlenvergleich (BZ: $g = 0,73$) höher verglichen mit der Simultanerfassung (bis 3) (BZ: $g = 0,77$) und dem Mengenvergleich (BZ: $g = 0,51$). In der Richtigkeit zeigten Personen mit Rechenstörung hingegen größere Defizite beim Mengenvergleich (RI: $g = 0,49$) als beim Zahlenvergleich (RI: $g = 0,26$). Ebenso in der Richtigkeit wurden große Effekte für die Simultanerfassung (ab 4) (RI: $g = 0,81$), dem Zahlen-Mengen-Vergleich (RI: $g = 0,88$) und dem Transkodieren (RI: $g = 1,00$) errechnet, wohingegen die Effektstärke für den Zahlenstrahl mittel war (RI: $g = 0,67$). Die Testaufgabe Zählen war sowohl in der Richtigkeit als in der benötigten Zeit nicht-signifikant mit einem kleinen bzw. mittleren Effekt (RI: $g = 0,33$; BZ: $g = 0,54$). Die ermittelten Effektstärken der domänenspezifischen Fähigkeiten für die Richtigkeit sowie benötigte Zeit sind in Abbildung 5 veranschaulicht.

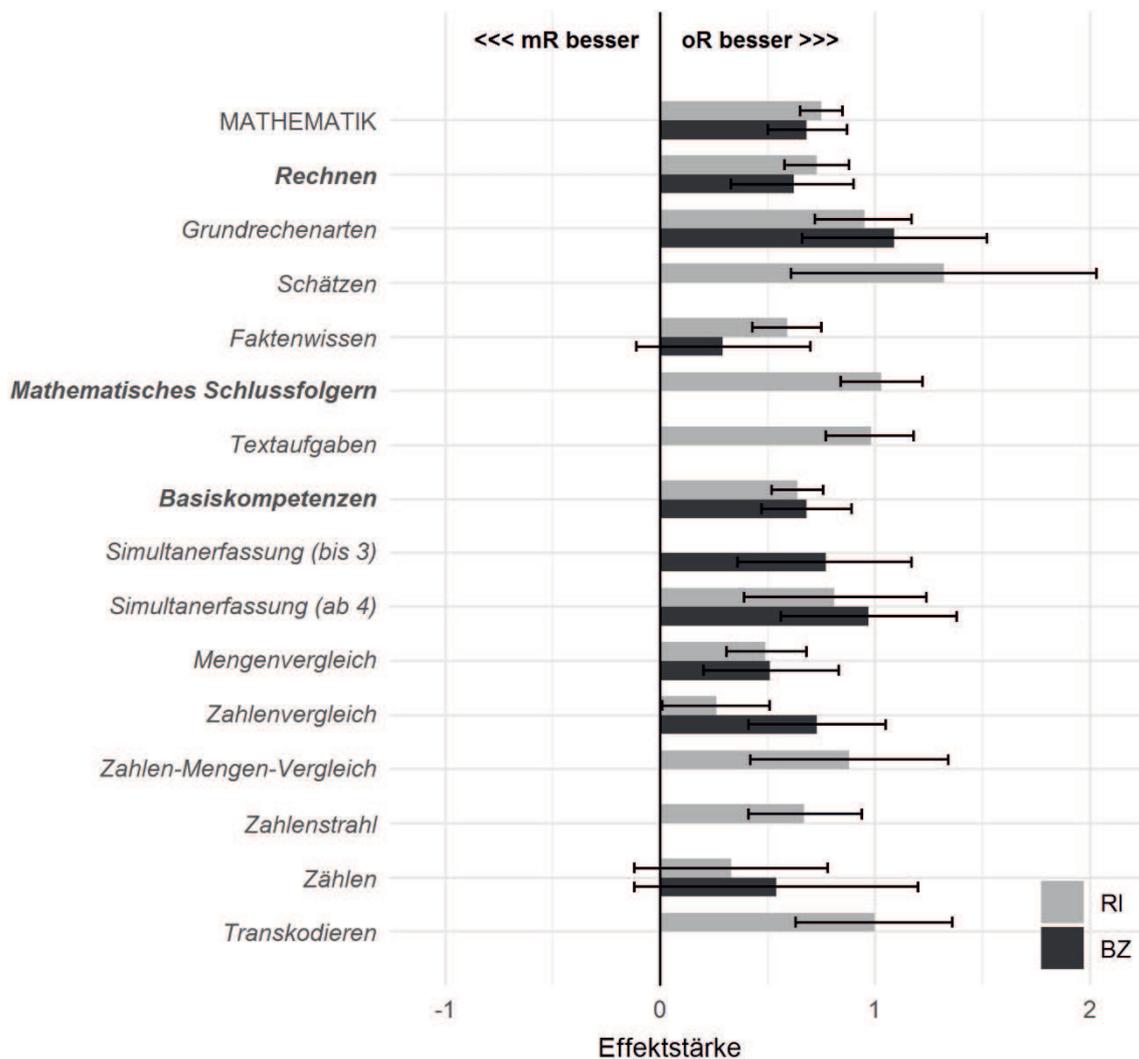


Abbildung 5. Ermittelte Effektstärken für domänenübergreifende Fähigkeiten. mR besser = Effektstärken im negativen Bereich bedeuten eine bessere Leistung von Personen mit Rechenstörung; oR besser = Effektstärken im positiven Bereich bedeuten eine bessere Leistung von Personen ohne Rechenstörung; RI = Richtigkeit; BZ = benötigte Zeit. Formatierung der Schriftart der Outcomes entspricht den Levels der Kodierung, d. h. Großbuchstaben = übergeordnete Fähigkeit, fettkursiv = untergeordnete Fähigkeit, kursiv = Testaufgabe.

4.2.2 Domänenübergreifende Fähigkeiten: Ermittelte Effektstärken

Für das Arbeitsgedächtnis berichteten die eingeschlossenen Studien nur Daten zu Richtigkeit. Bei der übergeordneten Fähigkeit Arbeitsgedächtnis sowie der untergeordneten Fähigkeit Arbeitsgedächtniskapazität ergaben sich jeweils mittlere Effekte (RI: $g = 0,55$ bzw. RI: $g = 0,56$). Entsprechend verhielt es sich bei

den Testaufgaben, wo für Spanne rückwärts sowie Spanne komplex jeweils Effektstärken in ähnlicher Höhe ermittelt wurden (RI: $g = 0,43$; RI: $g = 0,58$).

Für das Kurzzeitgedächtnis berichteten die eingeschlossenen Studien nur Daten zu Richtigkeit. Bei der übergeordneten Fähigkeit Kurzzeitgedächtnis ergab sich ein mittlerer Effekt (RI: $g = 0,50$). Hinsichtlich der untergeordneten Fähigkeiten zeigten Personen mit Rechenstörung deutlich größere Defizite im visuell-räumlichen Kurzzeitgedächtnis (RI: $g = 0,79$) als im auditiven Kurzzeitgedächtnis (RI: $g = 0,37$). Dies betraf ebenso die zu den untergeordneten Fähigkeiten zugewiesenen Testaufgaben. Hierbei waren für die Testaufgaben Spanne vorwärts sowie Abrufen die Effektstärken beim visuell-räumlichen Kurzzeitgedächtnis höher (RI: $g = 0,77$; RI: $g = 0,89$) verglichen mit denen beim auditiven Kurzzeitgedächtnis (RI: $g = 0,35$; AC: $g = 0,40$). Inhaltlich bedeutsame Unterschiede der Effektstärken zwischen den beiden Testaufgaben innerhalb der jeweiligen untergeordneten Fähigkeit zeigten sich nicht.

Für die Aufmerksamkeit zeigten Personen mit Rechenstörung in der übergeordneten Fähigkeit Aufmerksamkeit gravierendere Defizite in der benötigten Zeit (BZ: $g = 0,69$) als in der Richtigkeit (RI: $g = 0,44$). Für die untergeordneten Fähigkeiten ergaben sich in der benötigten Zeit ein mittlerer Effekt bei der Inhibition (BZ: $g = 0,65$) und ein großer Effekt beim Shifting (BZ: $g = 0,97$) verglichen mit kleinen und teilweise nicht-signifikanten Effekten in der Richtigkeit bei der Inhibition (RI: $g = 0,41$) und beim Updating (RI: $g = 0,34$). Hinsichtlich der Testaufgabe Stroop-Effekt der untergeordneten Fähigkeit Inhibition zeigte sich eine mittlere Effektstärke in der benötigten Zeit (BZ: $g = 0,71$) und eine kleine Effektstärke in der Richtigkeit (RI: $g = 0,30$). Analog verhielt es sich auch mit den restlichen Testaufgaben, wonach die Effekte in der benötigten Zeit größer waren als in der Richtigkeit. Für die Testaufgabe Buchstaben-und-Zahlen-Folgen der untergeordneten Fähigkeit Shifting wurde ein großer Effekt in der benötigten Zeit ermittelt (BZ: $g = 0,80$) und für die Testaufgabe Zufallsgenerierung der untergeordneten Fähigkeit Inhibition ein kleiner Effekt in der Richtigkeit (RI: $g = 0,38$).

Für die Verarbeitungsgeschwindigkeit zeigten Personen mit Rechenstörung größere Defizite in der Richtigkeit als in der benötigten Zeit. Bei der übergeordneten Fähigkeit Verarbeitungsgeschwindigkeit ergab sich in der Richtigkeit ein

mittlerer Effekt (RI: $g = 0,50$) verglichen mit einem kleinen Effekt in der benötigten Zeit (BZ: $g = 0,28$). Bei den untergeordneten Fähigkeiten war festzustellen, dass ausschließlich für die visuelle Verarbeitungsgeschwindigkeit bedeutsame Effekte ermittelt wurden. So lagen bei der visuellen Verarbeitungsgeschwindigkeit mittlere Effekte sowohl in der Richtigkeit (RI: $g = 0,65$) als auch in der benötigten Zeit (BZ: $g = 0,51$) vor verglichen mit keinen bis kleinen nicht-signifikanten Effekten in den anderen untergeordneten Fähigkeiten Reaktions- und Entscheidungsgeschwindigkeit (RI: $g = -0,37$; BZ: $g = 0,11$) sowie psychomotorische Geschwindigkeit (RI: $g = 0,33$). Dieses Bild setzte sich bei den Testaufgaben fort. Für die visuelle Verarbeitungsgeschwindigkeit wurden bei den Testaufgaben Buchstaben-oder-Zahlen-Folgen (BZ: $g = 0,74$), Kodieren (RI: $g = 0,64$) und visueller Vergleich (RI: $g = 0,73$) jeweils mittlere Effektstärken in der Richtigkeit oder benötigten Zeit ermittelt. Bei Reagieren auf Stimulus, als Testaufgabe zur Reaktions- und Entscheidungsgeschwindigkeit, ergab sich hingegen kein relevanter Effekt in der benötigten Zeit (BZ: $g = 0,18$). Auch zeigte sich für die Schreibgeschwindigkeit, als Testaufgabe zur psychomotorischen Geschwindigkeit, nur ein kleiner Effekt in der Richtigkeit (RI: $g = 0,24$).

Für die Raumvorstellung berichteten die eingeschlossenen Studien nur Daten zu Richtigkeit. Bei der übergeordneten Fähigkeit Raumvorstellung und der entsprechenden untergeordneten Fähigkeit Visualisierung ergaben sich jeweils mittlere Effekte (RI: $g = 0,57$ bzw. RI: $g = 0,64$).

Für die phonologische Verarbeitung zeigten Personen mit Rechenstörung ähnliche Defizite in der Richtigkeit und in der benötigten Zeit. Bei der übergeordneten Fähigkeit phonologische Verarbeitung ergaben sich in der Richtigkeit sowie in der benötigten Zeit jeweils kleine Effekte (RI: $g = 0,33$; BZ: $g = 0,35$). Hinsichtlich der untergeordneten Fähigkeit schnelles Benennen wurden in der Richtigkeit und in der benötigten Zeit jeweils kleine Effekte ermittelt (RI: $g = 0,28$; BZ: $g = 0,39$). Der Effekt bei der phonologischen Bewusstheit in der Richtigkeit war ebenso klein (RI: $g = 0,31$).

Die ermittelten Effektstärken der domänenübergreifenden Fähigkeiten für die Richtigkeit sowie benötigte Zeit sind in Abbildung 6 veranschaulicht.

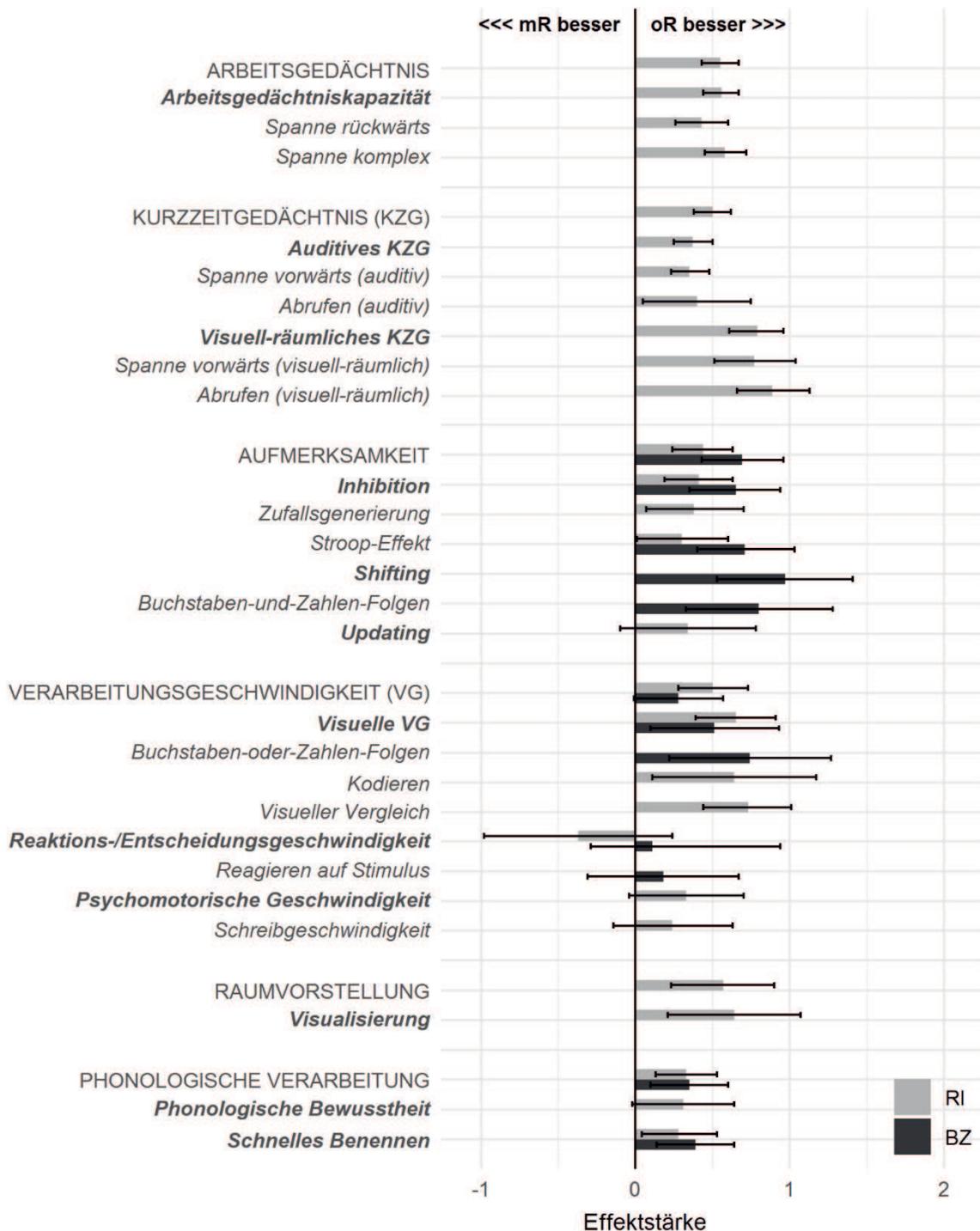


Abbildung 6. Ermittelte Effektstärken für domänenübergreifende Fähigkeiten. mR besser = Effektstärken im negativen Bereich bedeuten eine bessere Leistung von Personen mit Rechenstörung; oR besser = Effektstärken im positiven Bereich bedeuten eine bessere Leistung von Personen ohne Rechenstörung; RI = Richtigkeit; BZ = benötigte Zeit. Formatierung der Schriftart der Outcomes entspricht den Levels der Kodierung, d. h. Großbuchstaben = übergeordnete Fähigkeit, fettkursiv = untergeordnete Fähigkeit, kursiv = Testaufgabe.

4.3 Heterogenität und Publikationsbias

Die Ergebnisse des Modellfits mit fixierter Ebene 2 (d. h. Varianz zwischen den Effektstärken innerhalb einer Studie) bzw. Ebene 3 (d. h. Varianz zwischen den Studien) sowie die prozentuale Verteilung der Varianz über alle Ebenen hinweg sind in Tabelle 7 dargestellt. Hinsichtlich des Modellfits zeigte sich, dass bei allen Datensätzen, mit Ausnahme RI/TA, ein Modell, bei dem die Varianz auf 3 Ebenen variieren konnte, die Daten signifikant besser erklärte, als ein Modell mit fixierter Varianz auf Ebene 2 oder 3. In den Datensätzen lag somit eine signifikante Heterogenität vor, die nicht alleine durch die Stichprobenvarianz (d. h. Ebene 1) erklärt werden konnte. Dies spiegeln auch die prozentualen Anteile P der einzelnen Ebenen an der Gesamtvarianz σ^2 wider. Die Varianz der Ebene 1 erklärte über diese Datensätze hinweg lediglich maximal 39,29 % der Gesamtvarianz bzw. waren bis zu 64,73 % der Gesamtvarianz auf die Heterogenität der Daten innerhalb sowie zwischen den Studien (d. h. Ebene 2 und Ebene 3) zurückzuführen. Eine Ausnahme stelle der Datensatz RI/TA dar, bei dem auch ein Modell mit fixierter Varianz zwischen den Studien (d. h. Ebene 3) eine Passung erbrachte. Ebene 3 wies bei diesem Datensatz den geringsten Anteil an der Gesamtvarianz im Vergleich zu den anderen Datensätzen auf. Ebenso konnten 44,69 % der Gesamtvarianz bereits durch die Stichprobenvarianz (d. h. Ebene 1) erklärt werden. Aus Gründen der Vergleichbarkeit mit den anderen Datensätzen sowie der Tatsache, dass der Modellfit sich nur knapp ergab ($p = 0,085$), wurde bei diesem Datensatz auf eine separate Modellberechnung mit fixierter Ebene 3 verzichtet.

Tabelle 7

Modellfit mit fixierter Ebene 2 bzw. 3 und Anteile der Varianz je Ebene

Heterogenität	Richtigkeit			benötigte Zeit		
	ÜF	UF	TA	ÜF	UF	TA
LRT						
Ebene 2 (f)	81,02***	67,93***	32,93***	13,56***	11,96***	17,84***
Ebene 3 (f)	14,78***	10,13**	2,96	22,25***	20,18***	5,00*
σ^2	0,18	0,16	0,14	0,26	0,26	0,27
P						
Ebene 1	35,27	39,29	44,69	35,84	35,85	38,33
Ebene 2	47,45	46,81	46,49	21,01	19,87	32,22
Ebene 3	17,28	13,90	8,81	43,15	44,28	29,45

Anmerkung. ÜF = übergeordnete Fähigkeit; UF = untergeordnete Fähigkeit; TA = Testaufgabe; LRT = Wert des Likelihood-Ratio-Tests; f = fixiert; σ^2 = Gesamtvarianz; P = Varianzanteil je Ebene in Prozent.

* $p < 0.05$. ** $p < 0,01$. *** $p < 0,001$.

Die Ergebnisse zur studienübergreifenden Biasprüfung (II) sind in Tabelle 8 dargestellt. Bei der Richtigkeit ergab die Regressionsanalyse nach Egger Hinweise auf einen Publikationsbias für die übergeordneten Fähigkeiten Mathematik, Arbeitsgedächtnis, Kurzzeitgedächtnis und Aufmerksamkeit und in Teilen für die entsprechend dazugehörigen untergeordneten Fähigkeiten sowie Testaufgaben. Bei der benötigten Zeit bestand gemäß der Analyse die Gefahr eines Publikationsbias bei allen übergeordneten Fähigkeiten sowie den meisten untergeordneten Fähigkeiten und Testaufgaben.

Der Publikationstyp (d. h. Artikel vs. Dissertation/Masterarbeit) war in der Subgruppenanalyse ein signifikanter Moderator für die ermittelten Effektstärken in der Richtigkeit bei der übergeordneten Fähigkeit Mathematik (RI: $b = 0,30$) sowie der untergeordneten Fähigkeit Basiskompetenzen (RI: $b = 0,72$). In beiden Fällen waren die ermittelten Effektstärken bei Dissertationen/Masterarbeiten größer.

Tabelle 8

Studienübergreifende Biasprüfung (II)

Outcome (ÜF/UF/TA) ^a	Richtigkeit				benötigte Zeit			
	Egger	Publikationstyp (Referenzgruppe: A)			Egger	Publikationstyp (Referenzgruppe: A)		
		<i>t</i>	<i>b</i>	<i>SE</i>		<i>t</i>	<i>b</i>	<i>SE</i>
domänenspezifische Fähigkeiten								
Mathematik	2,23*	0,30*	0,13	2,43	4,91***	-0,32	0,26	-1,21
Rechnen	2,00*	0,13	0,16	0,81	3,41***	-0,92	0,48	-1,91
Grundrechenarten	0,04	0,20	0,25	0,81	1,91			
Schätzen	-1,21							
Faktenwissen	1,02	-0,06	0,18	-0,34	0,19	-0,66	0,62	-1,07
Mathematisches Schlussfolgern	2,45*	0,22	0,22	1,02				
Textaufgaben	1,59	0,09	0,29	0,30				
Basiskompetenzen	1,80	0,73**	0,25	2,96	4,92***	-0,27	0,32	-0,82
Simultanerfassung (bis 3)					2,51*			
Simultanerfassung (ab 4)	0,85				3,71***			
Mengenvergleich	3,52***	0,60	0,56	1,07	0,45	-0,37	0,60	-0,62
Zahlenvergleich	-1,80				0,49			
Zahlen-Mengen-Vergleich	-0,23							
Zahlenstrahl	2,08*							
Zählen	0,11				-0,08	0,23	0,65	0,35
Transkodieren	1,23							
domänenübergreifende Fähigkeiten								
Arbeitsgedächtnis	2,58*	-0,11	0,15	-0,72				
Arbeitsgedächtniskapazität	2,32*	-0,13	0,14	-0,91				

Outcome (ÜF/UF/TA) ^a	Richtigkeit				benötigte Zeit			
	Egger	Publikationstyp (Referenzgruppe: A)			Egger	Publikationstyp (Referenzgruppe: A)		
		<i>t</i>	<i>b</i>	<i>SE</i>		<i>t</i>	<i>b</i>	<i>SE</i>
Spanne rückwärts	0,93	-0,17	0,19	-0,87				
Spanne komplex	2,51*	-0,04	0,20	-0,19				
Kurzzeitgedächtnis (KZG)	3,40***	-0,15	0,16	-0,91				
Auditives KZG	2,57*	-0,04	0,16	-0,23				
Spanne vorwärts (auditiv)	2,12*	-0,04	0,15	-0,28				
Abrufen (auditiv)	0,91							
Visuell-räumliches KZG	1,97							
Spanne vorwärts (visuell-räumlich)	1,45							
Abrufen (visuell-räumlich)	0,14							
Aufmerksamkeit	2,00*	0,25	0,26	0,95	6,27***	0,66	0,34	1,96
Inhibition	1,51	0,24	0,26	0,91	6,11***	0,66	0,44	1,51
Zufallsgenerierung	0,05	0,42	0,33	1,29				
Stroop-Effekt	0,81				7,08***	0,65	0,42	1,56
Shifting					0,78	0,37	0,46	0,80
Buchstaben-und-Zahlen-Folgen					0,72	0,92	0,48	1,92
Updating	-0,25							
Verarbeitungsgeschwindigkeit (VG)	1,32	-0,03	0,23	-0,12	5,64***	-0,08	0,34	-0,25
Visuelle VG	1,74	0,07	0,28	0,27	4,40***	-0,11	0,44	-0,25
Buchstaben-oder-Zahlen-Folgen					4,45***	0,14	0,52	0,27
Kodieren	0,20							
Visueller Vergleich	1,86	0,29	0,37	0,79				
Reaktions-/Entscheidungsgeschwindigkeit	0,37				3,54***			
Reagieren auf Stimulus					3,61***			

Outcome (ÜF/UF/TA) ^a	Richtigkeit				benötigte Zeit			
	Egger	Publikationstyp (Referenzgruppe: A)			Egger	Publikationstyp (Referenzgruppe: A)		
	<i>t</i>	<i>b</i>	<i>SE</i>	<i>t</i>	<i>t</i>	<i>b</i>	<i>SE</i>	<i>t</i>
Psychomotorische Geschwindigkeit	0,88	-0,16	0,53	-0,3				
Schreibgeschwindigkeit	0,86	-0,32	0,53	-0,6				
Raumvorstellung	1,60	-0,28	0,34	-0,82				
Visualisierung	1,05	-0,47	0,46	-1,03				
Phonologische Verarbeitung	1,42	-0,09	0,22	-0,40	2,17*	-0,01	0,25	-0,02
Phonologische Bewusstheit	0,20	-0,06	0,33	-0,19				
Schnelles Benennen	1,16	-0,22	0,29	-0,77	2,20*	-0,05	0,26	-0,19

Anmerkung. Egger = Regressionsanalyse nach Egger; Referenzgruppe = Die Veränderung der Effektstärke zur jeweiligen Referenzgruppe; A = Studie als Artikel in einem Journal erschienen; *t* = *t*-Wert; *b* = Veränderung der Effektstärke; *SE* = Standardfehler.

^aDie Unterordnung der Outcomes entspricht den Levels der Kodierung, d. h. kein Einzug links = übergeordnete Fähigkeit, mittlerer Einzug links = untergeordnete Fähigkeit, großer Einzug links = Testaufgabe.

p* < 0,05. *p* < 0,01. ****p* < 0,001.

4.4 Kognitives Profil der Rechenstörung

Die Anzahl an Studien je Outcome unterschied sich deutlich und lag zwischen 2 und 28 Studien. Obwohl es keine Regel gibt, wie viele Studien für eine möglichst genaue Schätzung der Effektstärken notwendig sind, so werden für gewöhnlich mit größerer Studienanzahl auch genauere Effektstärkenschatzungen erreicht. Es war in dieser Arbeit daher nicht auszuschließen, dass die beobachteten Werte der Effektstärken, die auf Basis der Daten von nur wenigen Studien berechnet wurden, die wahren Werte der zugrundeliegenden Effektstärken nur verzerrt wiedergaben. Ein Profil der Rechenstörung, welches daher lediglich auf die ermittelten Effektstärken der einzelnen Outcomes fußte, hätte somit zu falschen Schlüssen geführt. Aus diesem Grund war es notwendig, die ermittelten Effektstärken danach zu bewerten, wie genau sie der zugrundeliegenden wahren Effektstärke entsprachen. Die Genauigkeit bzw. Reliabilität der ermittelten Effektstärken war in diesem Fall durch den ermittelten Standardfehler gegeben. Dessen Höhe zeigt an, wie sehr die berichteten Effektstärken in den einzelnen Studien um die wahre Effektstärke streuen. Im Rahmen von Metaanalysen ist der Standardfehler für gewöhnlich kleiner, je mehr Studien eingeschlossen sind (Borenstein et al., 2009). Dies besagt letztlich auch der zentrale Grenzwertsatz, wonach mit zunehmender Anzahl an Effektstärken zu einem Outcome deren Verteilung sich sukzessive einer Normalverteilung annähert, dessen Mittelwert eine immer genauere Schätzung der wahren Effektstärke darstellt.

Eine Abwandlung des Standardfehlers ist der *relative Standardfehler (RSE)*. Er wird berechnet als Quotient aus dem Standardfehler und dem Mittelwert einer Studie und gibt an, ob der jeweilige Standardfehler groß oder klein war in Bezug zum jeweiligen Mittelwert. Somit ist es möglich, die Höhe des Standardfehlers auch ohne Kenntnis des wahren Mittelwerts zu interpretieren und Aussagen darüber zu treffen, wie genau der ermittelte Mittelwert dem wahren Mittelwert entspricht. Übertragen auf Metaanalysen berechnet sich der relative Standardfehler durch:

$$RSE = \left(\frac{SE_g}{|g|} \right) * 100$$

Der relative Standardfehler ist somit der Quotient aus dem Standardfehler der ermittelten Effektstärke SE_g geteilt durch den absoluten Betrag dieser Effekt-

stärke g und angegeben als Prozent. Effektstärken mit einem relativen Standardfehler von kleiner oder gleich 25 % werden als *eher reliable* Schätzung der wahren Effektstärke betrachtet. Effektstärken mit einem relativen Standardfehler von größer 25 % und kleiner oder gleich 50 % als *eher nicht-reliable* Schätzung der wahren Effektstärke bewertet, die mit Vorsicht zu interpretieren sind. Bei Effektstärken mit einem relativen Standardfehler von über 50 % sind keine Aussagen möglich, wie nah die Schätzung an der wahren Effektstärke liegt (Harper, 2005; Klein, Proctor, Boudreault & Turczyn, 2002). Die Reliabilität einer Effektstärkenschtzung erlaubt in diesem Fall auch Rückschlüsse auf die Stabilität der jeweiligen zugrundeliegenden Fähigkeit oder Testaufgabe. Bezogen auf diese Metaanalyse bedeutete dies: Umso genauer die ermittelte Effektstärke der wahren Effektstärke entsprach, als umso stabiler konnten die gefunden Unterschiede zwischen Personen mit und ohne Rechenstörung in dieser Fähigkeit und unter Verwendung der zugehörigen Testaufgaben gesehen werden. Dies bedeutete, dass sich Defizite in diesen Fähigkeiten und Testaufgaben in ähnlicher Ausprägung auch bei anderen Personen mit Rechenstörung zeigen sollten, die nicht Teil einer Studienstichprobe dieser Metaanalyse waren.

Auf Basis dieser Methode wurden die ermittelten Effektstärken dieser Metaanalyse, aufgeteilt nach Level und Skala, nach der Höhe ihres relativen Standardfehlers sortiert (von klein zu groß) und ihrer Reliabilität bzw. Stabilität entsprechend bewertet. Das so ermittelte finale kognitive Profil der Rechenstörung zeigt Tabelle 9. Zum besseren Verständnis wurden in der Tabelle die Werte der Effektstärken durch deren Interpretation nach (Cohen, 1969, 1988) ersetzt (d. h. kein, kleiner, mittlerer oder großer Effekt). Eine Rechenstörung kennzeichnete sich demnach durch mittlere bis große stabile domänenspezifische Defizite in der übergeordneten Fähigkeit Mathematik (RI, BZ), den untergeordneten Fähigkeiten Basiskompetenzen (RI, BZ), Rechnen (RI, BZ) und mathematisches Schlussfolgern (RI) sowie den Testaufgaben Textaufgaben (RI), Grundrechenarten (RI, BZ), Faktenwissen (RI), Transkodieren (RI), Zahlenstrahl (RI), Simultanerfassung (ab 4) (BZ) und Zahlenvergleich (BZ). Bezogen auf die domänenübergreifenden Fähigkeiten zeigten sich mittlere bis große stabile Defizite in den übergeordneten Fähigkeiten Arbeitsgedächtnis (RI), Kurzzeitgedächtnis (RI), Aufmerksamkeit (BZ) und Verarbeitungsgeschwindigkeit (RI), den untergeordneten Fähigkeiten Arbeitsgedächtniskapazität (RI), visuell-räumliches

Kurzzeitgedächtnis (RI), visuelle Verarbeitungsgeschwindigkeit (RI), Shifting (BZ) und Inhibition (BZ) sowie den Testaufgaben Spanne komplex (RI), Abrufen (visuell-räumlich) (RI), Spanne vorwärts (visuell-räumlich) (RI), visueller Vergleich (RI) und Stroop-Effekt (BZ). Stabile domänenspezifische und domänenübergreifende Defizite mit keinem oder nur geringem Effekt lagen für die übergeordnete Fähigkeit Aufmerksamkeit (RI), der untergeordneten Fähigkeit auditives Kurzzeitgedächtnis (RI) sowie den Testaufgaben Spanne vorwärts (auditiv) (RI), Mengenvergleich (RI) und Spanne rückwärts (RI) vor.

Tabelle 9

Kognitives Profil der Rechenstörung basierend auf dem relativen Standardfehler (RSE)

übergeordnete Fähigkeit			untergeordnete Fähigkeit			Testaufgabe		
Pos.	Outcome (Skala)	<i>g</i>	Pos.	Outcome (Skala)	<i>g</i>	Pos.	Outcome (Skala)	<i>g</i>
eher reliabel ($RSE \leq 25\%$)								
1	Mathematik (RI)	M	1	Basiskompetenzen (RI)	M	1	Textaufgaben (RI)	G
2	Arbeitsgedächtnis (RI)	M	2	Rechnen (RI)	M	2	Spanne komplex (RI)	M
3	Kurzzeitgedächtnis (RI)	M	3	Mathematisches Schlussfolgern (RI)	G	3	Grundrechenarten (RI)	G
4	Mathematik (BZ)	M	4	Arbeitsgedächtniskapazität (RI)	M	4	Abrufen (visuell-räumlich) (RI)	G
5	Aufmerksamkeit (BZ)	M	5	Visuell-räumliches Kurzzeitgedächtnis (RI)	M	5	Faktenwissen (RI)	M
6	Verarbeitungsgeschwindigkeit (RI)	M	6	Basiskompetenzen (BZ)	M	6	Spanne vorwärts (visuell-räumlich) (RI)	M
7	Aufmerksamkeit (RI)	K	7	Auditives Kurzzeitgedächtnis (RI)	K	7	Spanne vorwärts (auditiv) (RI)	K
			8	Visuelle Verarbeitungsgeschwindigkeit (RI)	M	8	Mengenvergleich (RI)	K
			9	Rechnen (BZ)	M	9	Transkodieren (RI)	G
			10	Shifting (BZ)	G	10	Visueller Vergleich (RI)	M
			11	Inhibition (BZ)	M	11	Grundrechenarten (BZ)	G
						12	Zahlenstrahl (RI)	M
						13	Spanne rückwärts (RI)	K
						14	Simultanerfassung (ab 4) (BZ)	G
						15	Zahlenvergleich (BZ)	M

übergeordnete Fähigkeit			untergeordnete Fähigkeit			Testaufgabe		
Pos.	Outcome (Skala)	<i>g</i>	Pos.	Outcome (Skala)	<i>g</i>	Pos.	Outcome (Skala)	<i>g</i>
						16	Stroop-Effekt (BZ)	M
eher nicht-reliabel ($25\% > RSE \leq 50\%$)								
8	Raumvorstellung (RI)	M	12	Inhibition (RI)	K	17	Simultanerfassung (bis 3) (BZ)	M
9	Phonologische Verarbeitung (RI)	K	13	Schnelles Benennen (BZ)	K	18	Zahlen-Mengen-Vergleich (RI)	G
10	Phonologische Verarbeitung (BZ)	K	14	Visualisierung (RI)	M	19	Simultanerfassung (ab 4) (RI)	G
			15	Visuelle Verarbeitungsgeschwindigkeit (BZ)	M	20	Schätzen (RI)	G
			16	Schnelles Benennen (RI)	K	21	Buchstaben-und-Zahlen-Folgen (BZ)	G
						22	Mengenvergleich (BZ)	M
						23	Buchstaben-oder-Zahlen-Folgen (BZ)	M
						24	Zufallsgenerierung (RI)	K
						25	Kodieren (RI)	M
						26	Abrufen (auditiv) (RI)	K
						27	Stroop-Effekt (RI)	K
						28	Zahlenvergleich (RI)	K
keine Aussage möglich ($RSE > 50\%$)								
11	Verarbeitungsgeschwindigkeit (BZ)	K	17	Phonologische Bewusstheit (RI)	K	29	Zählen (BZ)	M

übergeordnete Fähigkeit			untergeordnete Fähigkeit			Testaufgabe		
Pos.	Outcome (Skala)	<i>g</i>	Pos.	Outcome (Skala)	<i>g</i>	Pos.	Outcome (Skala)	<i>g</i>
			18	Psychomotorische Geschwindigkeit (RI)	K	30	Faktenwissen (BZ)	K
			19	Updating (RI)	K	31	Zählen (RI)	K
			20	Reaktions-/Entscheidungsgeschwindigkeit (RI)	K	32	Schreibgeschwindigkeit (RI)	K
			21	Reaktions-/Entscheidungsgeschwindigkeit (BZ)	N	34	Reagieren auf Stimulus (BZ)	N

Anmerkung. Pos. = Sortierung der Outcomes auf Basis ihres relativen Standardfehlers; *g* = Interpretation der Effektstärke gemäß (Cohen, 1969, 1988); N = kein Effekt ($g < 0,2$); K = kleiner Effekt ($0,2 \leq g < 0,5$); M = mittlerer Effekt ($0,5 \leq g < 0,8$); G = großer Effekt ($g \geq 0,8$); *RSE* = relativer Standardfehler; RI = Richtigkeit; BZ = benötigte Zeit.

5 Diskussion

Im Rahmen dieser Metaanalyse wurden Personen mit und ohne Rechenstörung hinsichtlich ihrer Leistung in der Richtigkeit und der benötigten Zeit bei domänenspezifischen und domänenübergreifenden Fähigkeiten und Testaufgaben untersucht. Ein kognitives Profil stabiler Defizite in den übergeordneten und untergeordneten Fähigkeiten sowie Testaufgaben konnte so ermittelt werden.

5.1 Domänenspezifische Fähigkeiten: Stabile Defizite

In Bezug zu den Fragestellungen dieser Metaanalyse ist eine Rechenstörung charakterisiert durch stabile mittlere bis große Defizite in der übergeordneten Fähigkeit Mathematik (RI, BZ), den untergeordneten Fähigkeiten Basiskompetenzen (RI, BZ), Rechnen (RI, BZ) und mathematisches Schlussfolgern (RI) sowie den Testaufgaben Textaufgaben (RI), Grundrechenarten (RI, BZ), Faktenwissen (RI), Transkodieren (RI), Zahlenstrahl (RI), Simultanerfassung (ab 4) (BZ) und Zahlenvergleich (BZ). Ein stabiles, jedoch kleines Defizit liegt für die Testaufgabe Mengenvergleich (RI) vor.

5.1.1 Rechnen

Eine Rechenstörung betrifft demnach alle Bereiche der Mathematik. Diese Schlussfolgerung ist übereinstimmend mit bisherigen Forschungsergebnissen, die die Rechenstörung als eine heterogene Störung auffassten (Karagiannakis et al., 2014; Kaufmann et al., 2013). Beim Rechnen sind, wie auch in anderen Arbeiten, große Defizite bei Personen mit Rechenstörung festzustellen (Busch et al., 2015). Interessanterweise zeigen sich jedoch nur bei den Grundrechenarten Defizite sowohl in der Richtigkeit als auch in der benötigten Zeit, wohingegen für das Faktenwissen nur Defizite in der Richtigkeit festzustellen sind. Da Defizite im Faktenwissen ein Leitsymptom der Rechenstörung darstellen (Jordan & Hanich, 2003), wäre zu erwarten gewesen, dass Personen mit Rechenstörung hier eine entsprechend langsamere Bearbeitungszeit aufweisen. Vier Studien berichteten im Rahmen dieser Metaanalyse Daten zur Bearbeitungszeit, zwischen denen jedoch methodische Unterschiede festzustellen sind, die die Höhe und Stabilität der Effektstärke beeinträchtigt haben können. Rosenberg-Lee et al. (2014) und Kucian (2005) sind neurobiologische Studien, bei denen Kinder während einer fMRT-Untersuchung Aufgaben zum Faktenwissen lösen muss-

ten. In der Studie von Raddatz et al. (2017) wurde den Testpersonen eine Kombination aus einfachen sowie komplexen Aufgaben zum Faktenwissen vorgelegt. Und bei Cirino et al. (2015) wurden nicht ausreichend Daten zur Berechnung der Effektstärken berichtet, sodass diese durch Umrechnung anderer berichteter Daten ermittelt wurden. Wenngleich alle 4 Studien die Einschlusskriterien erfüllen, so unterscheiden sie sich deutlich hinsichtlich ihres Studiendesigns und ihrer -qualität, wodurch ihre Ergebnisse eventuell durch die jeweiligen Studienbedingungen verzerrt bzw. moderiert sind und damit eine geringere Vergleichbarkeit zwischen den Studien vorliegt. Der relative Standardfehler (RSE) beim Faktenwissen in der benötigten Zeit ist demzufolge auch über 50 %, was auf eine zu geringe Reliabilität der Effektstärke hinweist und somit keine Schlussfolgerungen zulässt. Es ist daher weiterhin unklar und es besteht Forschungsbedarf in der Frage, in welchem Ausmaß auf die benötigte Zeit beim Faktenwissen als valides Maß für die Diagnostik einer Rechenstörung zurückgegriffen werden kann.

5.1.2 Mathematisches Schlussfolgern

Hinsichtlich des mathematischen Schlussfolgerns verwenden die meisten der eingeschlossenen Studien als Testaufgabe Textaufgaben oder sie berichten zusammengefasste Werte, die neben Textaufgaben auch Aufgaben beinhalten, bei denen die Testpersonen Tabellen oder Grafiken interpretieren müssen. Für Textaufgaben zeigt sich ein großes Defizit bei Personen mit Rechenstörung. Keine Aussagen können jedoch getroffen werden für andere Testaufgaben zum mathematischen Schlussfolgern. Lediglich eine Studie von McCall (1999) gab den Testpersonen Zahlenfolgen als Testaufgabe vor. Hierbei wurden ebenso große Defizite bei Personen mit Rechenstörung gefunden. Jedoch wies die Gruppe der Personen mit Rechenstörung ebenso einen um ca. 8 Punkte geringeren nonverbalen IQ auf. Da die nonverbale Intelligenzleistung hoch mit Zahlenfolgen korreliert (W. J. Schneider & McGrew, 2018), können demnach die gefundenen Defizite auch durch eine geringere Leistung im allgemeinen logischen Schlussfolgern bedingt sein. Vor diesem Hintergrund und der geringen Datenlage von einer Studie können somit keine Schlussfolgerungen zur Validität von Zahlenfolgen für die Diagnostik einer Rechenstörung gezogen werden. Da sowohl DSM-5 als auch ICD-11 Defizite im mathematischen Schlussfolgern als eines der Leitsymptome einer Rechenstörung definieren,

besteht zur Validität der Testaufgaben Klärungsbedarf. Es sind daher mehr Studien notwendig, die die Leistung von Personen mit und ohne Rechenstörung beim mathematischen Schlussfolgern anhand verschiedener Testaufgaben vergleichen und dabei entweder die nonverbale Intelligenz in der Auswertung als Kovariate einfließen lassen oder die Gruppen nach ihrem nonverbalen IQ matchen.

5.1.3 Basiskompetenzen

Die Basiskompetenzen des Rechnens erweisen sich übergreifend als ein stabiler Indikator für eine Rechenstörung. Auffällig ist jedoch, dass, unabhängig der Effektstärkenhöhe, nur 5 von 12 erfassten Outcome-Skalen-Kombinationen in der Lage sind, genau genug zwischen Personen mit und ohne Rechenstörung zu differenzieren. Diese sind Transkodieren (RI), Mengenvergleich (RI), Zahlenstrahl (RI), Simultanerfassung (ab 4) (BZ) und Zahlenvergleich (BZ). Von methodischer Seite her betrachtet kann dies auf die geringere Anzahl an Studien zurückzuführen sein, die sich je nach Outcome zwischen 2 und 8 bewegte. Eine geringere Studienanzahl geht tendenziell mit einem größeren Standardfehler einher. Dieser wird umso größer, je unähnlicher die Effektstärken der Studien sind (Raudenbush, 2009), was wiederum zu weniger genauen Ergebnissen führt. Bezogen auf die ermittelten Effektstärken für die einzelnen Testaufgaben bestätigt diese Metaanalyse allerdings die Ergebnisse ähnlicher Reviews (de Smedt et al., 2013; Sasanguie, Göbel, Moll, Smets & Reynvoet, 2013; M. Schneider et al., 2017; Schwenk et al., 2017; Vanbinst & de Smedt, 2016), wonach bei der benötigten Zeit Testaufgaben zum Zahlenvergleich besser zwischen Personen mit und ohne Rechenstörung unterscheiden als Testaufgaben zum Mengenvergleich. Interessant dabei ist jedoch, dass bei nahezu allen Testaufgaben zur Mengenverarbeitung (d. h. Simultanerfassung (bis 3 bzw. ab 4), Mengenvergleich) mittlere bis hohe Effektstärken erzielt werden, die sich überdies als mehrheitlich reliabel erweisen. Eine Ausnahme stellt die Testaufgabe Mengenvergleich (RI) dar, deren Effektstärke von 0,49 per Definition noch als klein gilt. Die genaue Repräsentation von Mengen spielt zwar eine wichtige Rolle bei der Entwicklung der Zahlenverarbeitung im Vorschulalter und zu Beginn der Grundschule, korreliert jedoch mit den späteren mathematischen Fähigkeiten eher gering (Fazio, Bailey, Thompson & Siegler, 2014). Auch zeigt sich, wie eingangs erwähnt, die Studienlage zur Rolle der Mengenverarbeitung

bei der Rechenstörung hierzu nicht eindeutig. Es wäre daher für diese Metaanalyse zu erwarten gewesen, dass sich die unklare Forschungslage auch in den Testaufgaben zur Mengenverarbeitung widerspiegelt und diese sich als eher nicht-reliabel erweisen und mit geringeren Effektstärken einhergehen. Eine mögliche Erklärung für die dennoch gefunden stabilen Defizite bieten neurobiologische Befunde von fMRT-Studien. Die besondere Bedeutung der parietalen Regionen rund um den intraparietalen Sulcus (IPS) für die Zahlen- und Mengenverarbeitung wurde in mehreren Studien bestätigt (Peters & de Smedt, 2018). Mit zunehmendem Alter erfolgt überdies eine stärkere funktionale Ausdifferenzierung sämtlicher Gehirnregionen, die am Rechnen beteiligt sind. Es werden weniger kognitive Ressourcen des Arbeitsgedächtnisses und der Aufmerksamkeit benötigt, was sich in einer geringeren Aktivierung präfrontaler Regionen zeigt. Stattdessen erfolgt eine funktionale Aktivierung im Bereich des IPS (Kucian, 2016). McCaskey et al. (2018) konnten dabei zeigen, dass bei Personen mit Rechenstörung dieser Reifungsprozess erst verzögert stattfindet. Bei Testpersonen mit und ohne Rechenstörung wurden hierzu zum ersten Mal mit 9,6 Jahren und anschließend mit 13,8 Jahren zum zweiten Mal fMRT-Aufnahmen bei der Bearbeitung von Zahlenvergleichsaufgaben gemacht. Einerseits wiesen dabei Personen mit Rechenstörung bei der zweiten Testung eine höhere Aktivierung des IPS im Vergleich zur ersten Testung auf, andererseits war im Vergleich zu gleichaltrigen Personen ohne Rechenstörung weiterhin eine erhöhte präfrontale Aktivierung feststellbar. Bezogen auf diese Metaanalyse liegt das Durchschnittsalter der Personen mit Rechenstörung bei 10,7 Jahren. Überdies sind in sämtlichen Studien die Gruppen der Personen mit und ohne Rechenstörung nach dem Alter gematcht. Es ist daher hinsichtlich der ermittelten mittleren bis hohen Effektstärken bei Testaufgaben zur Mengenverarbeitung nicht auszuschließen, dass die Höhe der Effektstärke auch durch den Reifestatus des Gehirns moderiert wurde und der Unterschied zwischen den Gruppen mit zunehmendem Alter geringer geworden wäre. Zusammenfassend bleibt daher festzuhalten, dass im Rahmen dieser Metaanalyse Personen mit Rechenstörung Defizite sowohl in Testaufgaben zur Zahlenverarbeitung wie in Testaufgaben zur Mengenverarbeitung aufweisen. Beide Arten von Testaufgaben sind demzufolge bei der Diagnostik zu berücksichtigen, wenngleich zu beachten ist, dass Defizite in der Mengenverarbeitung sich mit zunehmendem Al-

ter und zunehmender Reifung des Gehirns nicht mehr als stabiler Marker erweisen können.

5.2 Domänenübergreifende Fähigkeiten: Stabile Defizite

In Bezug zu den Fragestellungen dieser Metaanalyse ist eine Rechenstörung charakterisiert durch stabile mittlere bis große Defizite in der übergeordneten Fähigkeiten Arbeitsgedächtnis (RI), Kurzzeitgedächtnis (RI), Aufmerksamkeit (BZ) und Verarbeitungsgeschwindigkeit (RI), den untergeordneten Fähigkeiten Arbeitsgedächtniskapazität (RI), visuell-räumliches Kurzzeitgedächtnis (RI), visuelle Verarbeitungsgeschwindigkeit (RI), Shifting (BZ) und Inhibition (BZ) sowie den Testaufgaben Spanne komplex (RI), Abrufen (visuell-räumlich) (RI), Spanne vorwärts (visuell-räumlich) (RI), visueller Vergleich (RI) und Stroop-Effekt (BZ). Stabile, jedoch kleine Defizite liegen für die übergeordnete Fähigkeit Aufmerksamkeit (RI), die untergeordnete Fähigkeit auditives Kurzzeitgedächtnis (RI) sowie die Testaufgaben Spanne vorwärts (auditiv) (RI) und Spanne rückwärts (RI) vor.

5.2.1 Arbeitsgedächtnis und Kurzzeitgedächtnis

Die gefundenen Defizite beim Arbeitsgedächtnis sind in Übereinstimmung mit den Ergebnissen anderer Reviews (z. B. David, 2012; Johnson et al., 2010; Peng et al., 2018; Peng & Fuchs, 2014; Swanson & Jerman, 2006). Die Testaufgabe Spanne komplex erweist sich dabei für Personen mit Rechenstörung als schwieriger zum Lösen als die Testaufgabe Spanne rückwärts. Der Unterschied zwischen den Effektstärken ist jedoch gering (Differenz von 0,15 Standardabweichungen), sodass er ebenso durch die Stichprobenszusammensetzung oder das methodische Vorgehen innerhalb der jeweiligen Studien begründet sein kann. Für das Kurzzeitgedächtnis wird übergreifend ein stabiles mittleres Defizit bei Personen mit Rechenstörung festgestellt, was ebenso in Einklang mit den oben genannten Reviews zum Arbeitsgedächtnis ist, die vergleichbare Ergebnisse zum Kurzzeitgedächtnis berichten. Bezogen auf die beiden untergeordneten Fähigkeiten, auditives und visuell-räumliches Kurzzeitgedächtnis, werden ähnlich wie bei Studien von David (2012) und Szűcs et al. (2013) größere Defizite für das visuell-räumliche Kurzzeitgedächtnis ermittelt. Dies kann als weiterer Beleg für das Konzept störungsspezifischer Defizite im Kurzzeitgedächtnis bei den Lernstörungen gelten, wonach eine isolierte Rechenstörung,

die auch bei David (2012) und Szűcs et al. (2013) erfasst wurde, mehrheitlich mit Defiziten im visuell-räumlichen Kurzzeitgedächtnis einhergeht, während bei einer Lesestörung Defizite im auditiven Kurzzeitgedächtnis charakteristisch sind. Im Falle einer komorbiden Rechen- und Lesestörung ist ein additiver Effekt festzustellen und betroffene Personen weisen Defizite in beiden Bereichen auf. Dies zeigte sich auch in einem Review von Peng et al. (2018), in dem bei Personen mit Rechen- und Lesestörung ähnlich große Effektstärken für das auditive und visuell-räumliche Kurzzeitgedächtnis ermittelt wurden. Hinsichtlich der Testaufgaben ergibt sich ein ähnliches Bild wie beim Arbeitsgedächtnis. So sind sich die Effektstärken der Testaufgaben des auditiven sowie des visuell-räumlichen Kurzzeitgedächtnisses untereinander sehr ähnlich (Differenz von 0,05 bzw. 0,12 Standardabweichungen). Auch hier ist, analog zu den Testaufgaben des Arbeitsgedächtnisses, nicht auszuschließen, dass die Unterschiede auf die Stichprobenzusammensetzung und Methodik der jeweiligen Studien zurückzuführen sind, weswegen die unterschiedlichen Effektstärkenhöhen nicht im Sinne unterschiedlich großer Defizite interpretiert werden können. Festzuhalten bleibt demnach, dass Personen mit Rechenstörung ein Defizit im Arbeitsgedächtnis sowie insbesondere im visuell-räumlichen Kurzzeitgedächtnis aufweisen und die Auswahl der Testaufgaben keinen bzw. nur einen geringen Einfluss auf das ermittelte Defizit hat.

5.2.2 Aufmerksamkeit

Hinsichtlich der Aufmerksamkeit zeigen sich in der übergeordneten Fähigkeit Aufmerksamkeit größere und stabilere Defizite, wenn die benötigte Zeit anstatt die Richtigkeit erfasst wird, was in Übereinstimmung mit mehreren Studien ist (z. B. Peng et al., 2012; Wang et al., 2012). Die untergeordneten Fähigkeiten Shifting und Updating wurden nur von wenigen Studien erfasst. Für Shifting berichten die eingeschlossenen Studien allesamt mittlere bis hohe Effektstärken, sodass sich diese untergeordnete Fähigkeit als stabil erweist. Es sind hierzu dennoch mehr Studien notwendig, um Aussagen darüber zu treffen, ob Defizite im Shifting explizit Teil des Profils einer Rechenstörung sind. Die dritte untergeordnete Fähigkeit zur Aufmerksamkeit, Inhibition, sowie deren zugeordnete Testaufgabe Stroop-Effekt kann hingegen als Defizit einer Rechenstörung gesehen werden, sofern die Daten auf der benötigten Zeit basieren. Für andere Testaufgaben zur Inhibition lagen für die benötigte Zeit zu wenige Daten vor,

um separate Effektstärken zu berechnen. Diese sind jedoch in der Gesamteffektstärke zur Inhibition eingeschlossen, dessen Defizit sich dadurch dennoch als stabil erweist. Ob sich demnach bei anderen Testaufgaben ähnlich stabile Defizite wie beim Stroop-Effekt bei Personen mit Rechenstörung in der benötigten Zeit zeigen, ist zu klären, wenngleich auf Basis der Daten dieser Studie die Vermutung naheliegt. Für die Richtigkeit ergeben sich hingegen keine stabilen Defizite. Hinsichtlich der Interpretation der Defizite zur Inhibition muss jedoch in Betracht gezogen werden, dass Schwierigkeiten in der Aufmerksamkeit und insbesondere in der Inhibition auch ein Leitsymptom einer Aufmerksamkeitsdefizit-/Hyperaktivitätsstörung (ADHS) sind (Barkley, 1997). Bei ungefähr 10 bis 20 % aller Personen mit Rechenstörung liegt zusätzlich eine ADHS vor (Fortes et al., 2016; Gross-Tsur et al., 1996; Schuchardt, Fischbach, Balke-Melcher & Mähler, 2015), wobei bis zu 40 % Probleme in der Aufmerksamkeitsleistung zeigen (Fischbach et al., 2010). ADHS war kein explizites Ausschlusskriterium dieser Metaanalyse und wurde nur von 3 der 7 Studien ausgeschlossen, die Daten zur Inhibition berichteten. Es ist daher zu berücksichtigen, dass das gefundene Defizit in der Aufmerksamkeit durch komorbide Fälle von Personen mit Rechenstörung und ADHS konfundiert ist.

Peng et al. (2018) berichten in ihrem Review lediglich ein kleines Defizit in der Aufmerksamkeit bei Personen mit isolierter Rechenstörung. Die Autoren des Reviews verglichen zwar auch die Defizite der untergeordneten Fähigkeiten Inhibition, Shifting sowie Updating, nutzten für den Vergleich jedoch eine Stichprobe, die ebenso komorbide Fälle von Personen mit einer Rechen- und Lesestörung beinhaltete. Hierbei zeigten sich im Gegensatz zu dieser Metaanalyse ein kleines Defizit bei der Inhibition und große Defizite beim Updating und Shifting. Zwei Begründungen für den Unterschied zwischen den Ergebnissen von Peng et al. (2018) und dieser Metaanalyse sind möglich. Zum einen ist unklar, ob die Daten von Peng et al. (2018) auf der benötigten Zeit, der Richtigkeit oder auf beiden Skalen basieren. Da diese Metaanalyse große Defizite bei der benötigten Zeit und kleine Defizite bei der Richtigkeit fand, ist nicht auszuschließen, dass die verwendete Skala die Schwere des Defizits beeinflusst. Zum anderen, und wie bereits erwähnt, schlossen Peng et al. (2018) komorbide Fälle von Personen mit Rechen- und Lesestörung für den Vergleich bei den untergeordneten Fähigkeiten nicht aus. Defizite bei der Inhibition, beim Updating und Shifting

wurden in anderen Studien ebenso bei Personen mit isolierter Lesestörung sowie komorbider Rechen- und Lesestörung gefunden (Booth, Boyle & Kelly, 2010; van der Sluis et al., 2004). Insbesondere Shifting und Updating scheinen dabei enger mit einer Lesestörung zusammenzuhängen als Inhibition (O. Moura, Simões & Pereira, 2014; Peng, Tao & Li, 2013). Festzuhalten bleibt, dass Inhibition, erfasst durch die Testaufgabe Stroop-Effekt, durchaus als ein Defizit der Rechenstörung zu sehen ist, wenngleich eine mögliche Konfundierung der Leistung durch eine komorbide ADHS in Betracht gezogen werden muss. Für Shifting und Updating können hingegen aufgrund der geringen Datenlage keine klaren Aussagen getroffen werden.

5.2.3 Verarbeitungsgeschwindigkeit

Bei der Verarbeitungsgeschwindigkeit stimmen die ermittelten Effektstärken mit denen anderer Reviews überein, die von einem Defizit bei Personen mit Rechenstörung berichten (Johnson et al., 2010; Peng et al., 2018). Hierbei ist zu berücksichtigen, dass alle bis auf eine Studie, die Daten zur Verarbeitungsgeschwindigkeit berichten, eine Rechenstörung unter anderem durch Defizite im Rechnen feststellen. Rechnen beruht, wie eingangs erwähnt, auch auf dem Abruf von Fakten aus dem Langzeitgedächtnis. Dieser kann bei einer defizitären Verarbeitungsgeschwindigkeit verlangsamt erfolgen, was wiederum eine geringere Rechenleistung bedingt (Bull & Johnston, 1997). Wenngleich sich die Verarbeitungsgeschwindigkeit in dieser Metaanalyse als stabiler Indikator einer Rechenstörung erweist, ist daher nicht auszuschließen, dass dies durch den Inhalt des verwendeten Mathematiktests in den eingeschlossenen Studien bedingt ist. So zeigten in einer Studie von Fuchs et al. (2008) Personen mit Rechenstörung geringere Defizite in der Verarbeitungsgeschwindigkeit, wenn die Diagnose einer Rechenstörung über die Leistung im mathematischen Schlussfolgern und nicht im Rechnen gestellt wurde. Zur Verarbeitungsgeschwindigkeit in dieser Metaanalyse erfasste ausschließlich Winegar (2013) eine Rechenstörung nicht über das Rechnen, sondern über das mathematische Schlussfolgern. Personen mit Rechenstörung zeigten dabei keine bis kleine Defizite in der psychometrischen Geschwindigkeit. Hierbei ist allerdings zu berücksichtigen, dass die visuelle Verarbeitungsgeschwindigkeit und nicht die psychometrische Geschwindigkeit die zentrale Komponente der Verarbeitungsgeschwindigkeit darstellt. Es kann daher nicht abschließend geklärt werden, inwiefern der Inhalt

des Mathematiktests zur Diagnosestellung die, in dieser Metaanalyse gefunden, Defizite in der Verarbeitungsgeschwindigkeit beeinflusst hat. Auf die der Verarbeitungsgeschwindigkeit untergeordneten Fähigkeiten eingehend, sind die Defizite dabei größer und stabiler bei der visuellen Verarbeitungsgeschwindigkeit als bei der Reaktions- und Entscheidungsgeschwindigkeit sowie der psychometrischen Geschwindigkeit. Hierbei zu berücksichtigen ist jedoch auch die Anzahl an Studien je Outcome, die für die Reaktions- und Entscheidungsgeschwindigkeit, der psychometrischen Geschwindigkeit sowie ihren jeweiligen Testaufgaben gering ist und zwischen 2 und 4 Studien liegt. Ähnliche Ergebnisse, insbesondere zur Reaktions- und Entscheidungsgeschwindigkeit, berichteten jedoch auch Raddatz et al. (2017) und Szűcs et al. (2013). In ihren Studien zeigte sich bei Testaufgaben, bei denen Testpersonen schnell reagieren oder eine Entscheidung treffen mussten, kein Unterschied zwischen Personen mit und ohne Rechenstörung. Im Gegensatz dazu erweisen sich in dieser Metaanalyse die Testaufgaben der visuellen Verarbeitungsgeschwindigkeit, bei denen visuelle Stimuli schnell erfasst, dekodiert und verglichen werden müssen, als besserer Indikator einer Rechenstörung. Dieses Ergebnis ist in Einklang mit den gefundenen Defiziten bei Personen mit Rechenstörung im visuell-räumlichen Kurzzeitgedächtnis sowie der Raumvorstellung (siehe 5.2.4), was auf ein allgemeines Defizit bei der Rechenstörung im Umgang mit Aufgaben schließen lässt, die ein hohes Maß an der Verarbeitung von visuellen sowie visuell-räumlichen Informationen beinhalten (Skagerlund & Träff, 2016; Szűcs et al., 2013). Eine Erklärung für den Zusammenhang zwischen der visuellen Verarbeitungsgeschwindigkeit, der Raumvorstellung und dem visuell-räumlichen Kurzzeitgedächtnis bietet das CHC-Modell der Intelligenz, welches eine mehrdimensionale, hierarchische Taxonomie von Fähigkeiten bzw. Intelligenzfaktoren ist, die über Faktorenanalysen gewonnen wurden. Gemäß dem Modell sind das visuell-räumliche Kurzzeitgedächtnis sowie die visuelle Verarbeitungsgeschwindigkeit Intelligenzfaktoren erster Ordnung, die am höchsten auf deren inhaltlich entsprechenden und hierarchisch höher gestellten Intelligenzfaktoren zweiter Ordnung laden, nämlich das Arbeitsgedächtnis und die Verarbeitungsgeschwindigkeit. Allerdings werden im Rahmen des Modells auch hohe Ladungen zur Raumvorstellung festgestellt, insbesondere dann, wenn die Items der jeweiligen Testaufgaben komplexer sind und demzufolge ein höheres Ausmaß

an visueller Verarbeitung und Raumvorstellung bedürfen (W. J. Schneider & McGrew, 2018). Es ist daher nicht auszuschließen, dass die visuelle und visuell-räumliche Komplexität der Items und das individuelle Leistungsniveau in der Raumvorstellung die Schwere der Defizite in anderen domänenübergreifenden Fähigkeiten, wie in diesem Fall Arbeitsgedächtnis oder Verarbeitungsgeschwindigkeit, beeinflussen. Festzuhalten bleibt, dass die Verarbeitungsgeschwindigkeit und dabei insbesondere die visuelle Verarbeitungsgeschwindigkeit als stabile Indikatoren einer Rechenstörung gelten. Möglich erscheint jedoch, dass das gefundene Defizit durch den Inhalt des verwendeten Mathematiktests zur Diagnosestellung sowie der visuellen und visuell-räumlichen Komplexität der Items beeinflusst wird.

5.2.4 Raumvorstellung

Bei der Raumvorstellung findet diese Metaanalyse zwar mittlere, jedoch nicht stabile Defizite. Von einem mittleren Defizit berichtet auch Swanson (2012). Hinsichtlich der untergeordneten Fähigkeiten lagen nur für die Visualisierung ausreichend Daten zur Schätzung von Effektstärken vor. Es zeigt sich hier ein mittleres Defizit, welches auch bei Skagerlund und Träff (2014) und Swanson (2012) gefunden wurde. Unter Berücksichtigung der Defizite, die für das visuell-räumliche Kurzzeitgedächtnis sowie der visuellen Verarbeitungsgeschwindigkeit ermittelt wurden, deutet diese Metaanalyse auf allgemeine Schwierigkeiten von Personen mit Rechenstörung bei der Verarbeitung visueller und visuell-räumlicher Informationen hin. Auch berichten die fMRT-Studien von McCaskey et al. (2017) und Skagerlund, Karlsson und Träff (2016) von einer höheren neuronalen Aktivität beim Bearbeiten von Längenvergleichs- sowie Rotationsaufgaben in den Regionen des Okzipital- und Parietallappen, die unter anderem auch beim Lösen von Zahlen- und Mengenvergleichsaufgaben aktiv sind. Ein neuraler Zusammenhang zwischen der Verarbeitung visueller und insbesondere visuell-räumlicher Informationen sowie der Zahlen- und Mengenverarbeitung, der bereits auf kognitiver Ebene auf Basis von Verhaltensdaten vermutet wurde (Cheng & Mix, 2014; Mussolin et al., 2011), scheint daher möglich. Zusätzlich ist zu klären, welche Wechselwirkungen zwischen diesen Fähigkeiten vorliegen und wie sich diese im Laufe der Entwicklung und mit zunehmender Ausdifferenzierung des Gehirns ändern. Anzunehmen ist, wie auch schon McCaskey et al. (2017) erwägen, dass die Aktivität domänenübergreifender Fähigkeiten, wie die

visuelle Verarbeitungsgeschwindigkeit, das visuell-räumliche Kurzzeitgedächtnis und die Raumvorstellung, mit zunehmender Herausbildung des IPS und angrenzenden Regionen, die vor allem für die Zahlen- und Mengenverarbeitung zuständig sind, abnimmt. Aufgaben zur Zahlen- und Mengenverarbeitung mit wenig komplexen visuellen und visuell-räumlichen Items können daher größtenteils effizient in diesen Regionen bearbeitet werden, während Aufgaben mit komplexeren visuellen und visuell-räumlichen Items eine zusätzliche entsprechende Aktivierung der Raumvorstellung, der visuellen Verarbeitungsgeschwindigkeit sowie des visuell-räumlichen Kurzzeitgedächtnisses bedürfen. Deren Aktivierung erfolgt ebenso okzipito-parietal, jedoch nicht ausschließlich im IPS (siehe 1.1) (McCaskey et al., 2018; Peters & de Smedt, 2018; Peters, De Smedt & Op de Beeck, 2015).

Bezogen auf die Rechenstörung zeigt diese Metaanalyse, dass Personen mit Rechenstörung Defizite bei der Verarbeitung visueller und visuell-räumlicher Informationen aufweisen. Es existieren jedoch keine Langzeitstudien über das Zusammenspiel dieser Fähigkeiten und ihrem Einfluss auf die Zahlen- und Mengenverarbeitung, sodass keine weiterführenden Annahmen getroffen werden können. Interessanterweise berichtet das Review von Peng et al. (2018) nur kleine Defizite bei der Raumvorstellung bei Personen mit Rechenstörung und geben dafür zwei Erklärungen an. Zum einen wurde Rechenstörung in ihrem Review als Defizit beim Rechnen sowie bei Textaufgaben definiert, wohingegen Schwierigkeiten bei der Raumvorstellung mehr, laut Aussage der Autoren/innen, mit komplexerer Mathematik einhergehen. Zum anderen wurde der Mindest-IQ, der als Ausschlusskriterium fungierte, in den meisten der eingeschlossenen Studien mit nonverbalen IQ-Tests (z. B. Matrizen- oder Mosaiktest) erfasst. Diese korrelieren hoch mit der Raumvorstellung und könnten, nach Sicht der Autoren/innen, die Ergebnisse dahingehend konfundiert haben, dass Personen mit niedrigem IQ und demzufolge auch niedrigerer Raumvorstellung nicht in die Analyse miteinbezogen wurden. Zur ersten Erklärung ist zu erwähnen, dass eine Rechenstörung für gewöhnlich mit Defiziten bei Aufgaben mit inhaltlich simplen bis normalen mathematischen Anforderungsniveau einhergehen (d. h. Rechnen, Basiskompetenzen, Textaufgaben) und es zweifelhaft ist, ob Schwierigkeiten in darüberhinausgehender Mathematik noch als Rechenstörung gelten (American Psychiatric Association, 2013). Die zweite Erklärung

steht im Widerspruch mit den Ergebnissen dieser Metaanalyse, bei der alle bis auf 5 Studien nonverbale IQ-Tests zur Erfassung des Mindest-IQ vorgegeben haben und dennoch Defizite in der Raumvorstellung vorliegen. Wohingegen diese Metaanalyse ein deutliches Defizitmuster bei der Verarbeitung visueller und visuell-räumlicher Informationen findet, so zeigen sich bei Peng et al. (2018) deutlich größere Effektstärken bei allen Fähigkeiten, wenn Personen mit Rechenstörung Aufgaben mit numerischen oder verbalen Informationen lösen mussten. Jedoch, wie oben bereits erwähnt, verwendeten Peng et al. (2018) für diese Subgruppenanalysen eine Stichprobe von Personen mit isolierter Rechenstörung sowie komorbider Rechen- und Lesestörung. Da das Verarbeiten von numerischen und verbalen Informationen stark mit der allgemeinen Lesekompetenz und der Fähigkeit zur phonologischen Verarbeitung zusammenhängt (Moll et al., 2015), kann das Vorhandensein einer Lesestörung in der Stichprobe das Defizit von Personen mit Rechenstörung beim Verarbeiten visueller und visuell-räumlicher Informationen überdeckt haben.

Festzuhalten bleibt daher, dass die Raumvorstellung sowie ihre untergeordnete Fähigkeit Visualisierung zwar keine stabilen Indikatoren einer Rechenstörung darstellen, aber dennoch in Zusammenhang mit den gefundenen Defiziten zum visuell-räumlichen Arbeitsgedächtnis und der visuellen Verarbeitungsgeschwindigkeit ein mögliches allgemeines Defizit bei Personen mit Rechenstörung im Umgang mit visuellen und visuell-räumlichen Informationen vorliegen kann, das weiterer Klärung bedarf.

5.2.5 Phonologische Verarbeitung

Hinsichtlich der phonologischen Verarbeitung werden in dieser Metaanalyse für die übergeordnete Fähigkeit sowohl für die Richtigkeit als auch die benötigte Zeit geringe und eher nicht stabile Defizite gefunden. Analog verhält es sich mit den entsprechenden untergeordneten Fähigkeiten phonologische Bewusstheit und schnelles Benennen. Da die phonologische Verarbeitung eine Basiskompetenz der Lesefähigkeit darstellt und eine Lesestörung vorab ausgeschlossen wurde, waren kleine Effektstärken zu erwarten. Dies ist ebenso in Übereinstimmung mit Studien von Moll et al. (2015) und Willburger et al. (2008), die von ähnlichen Ergebnissen bei Personen mit isolierter Rechenstörung berichten. Die phonologische Verarbeitung spielt dennoch, wie eingangs erwähnt, eine

wichtige Rolle beim Aufbau des Zahlen- und Faktenwissens. Die Zuordnung und Abspeicherung der symbolischen, ausgesprochenen und geschriebenen Form einer Zahl zu ihrer entsprechenden Menge (z. B. 3 = drei = ●●●) stellt auch einen sprachlichen Prozess dar, der insbesondere im Vorschulalter relevant ist (Peng et al., 2018). Die Stichprobe der meisten Studien in der Metaanalyse besuchte bereits die 3. Jahrgangsstufe, weswegen sie möglicherweise schon zu alt war, um größere Defizite in der phonologischen Verarbeitung unter gleichzeitigem Ausschluss einer Lesestörung zu finden. Vor dem Hintergrund, dass die Diagnose einer Rechenstörung erst ab der Grundschule möglich ist (Haberstroh & Schulte-Körne, 2019), bleibt daher festzuhalten, dass die phonologische Verarbeitung sowie ihre untergeordneten Fähigkeiten keine stabilen Indikatoren einer Rechenstörung sind. Es bedarf jedoch weiterer Klärung, ob sie mögliche Indikatoren zur Risikoidentifikation einer Rechenstörung im Vorschulalter darstellen.

5.3 Einschränkungen

Erstens lag für die Outcomes eine unterschiedliche Anzahl an Studien vor. Überdies reduzierte sich diese Anzahl mit zunehmender Spezifität der Outcomes (d. h. von übergeordneten Fähigkeiten hin zu Testaufgaben) und es wurden weniger Studien zur benötigten Zeit als zur Richtigkeit gefunden. Obwohl Metaanalysen bereits mit 2 Studien durchgeführt werden können, so erweisen sich die Ergebnisse bei größerer Studienanzahl als stabiler und weniger anfälliger für einen Bias (z. B. Publikationsbias). Effektstärken auf der Basis von nur wenigen Studien müssen daher mit Bedacht interpretiert werden und insbesondere dann, wenn zusätzlich von zufälligen Effekten sowie einem bestimmten Grad an Heterogenität zwischen sowie innerhalb der Studien auszugehen ist (Valentine et al., 2010). Aus diesem Grund basiert das kognitive Profil der Rechenstörung nicht ausschließlich auf der reinen Stärke der Effekte (d. h. klein, mittel oder groß), sondern wurde über den relativen Standardfehler (*RSE*) gewonnen. So war es möglich, genaue und stabile Defizite zu ermitteln, die die Rechenstörung charakterisieren.

Zweitens wurde über alle 6 Datensätze hinweg 61,79 % der Gesamtvarianz nicht durch die Stichprobenvarianz erklärt (durchschnittliche Varianz innerhalb der Studien: 35,64 %; durchschnittliche Varianz zwischen den Studien: 26,15

%). Obwohl dies die Verwendung eines Mehrebenenmodells statistisch rechtfertigt, so weist es auch auf einen bestimmten Grad an Heterogenität im Datensatz hin, der mit der studienspezifischen Stichprobenzusammensetzung und verwendeten Methodik zusammenhängt. Aufgrund gewichtiger methodischer Gründe wurden jedoch keine Subgruppenanalyse oder Metaregressionen durchgeführt mit der Absicht, die vorliegende Heterogenität zu erklären (siehe 3.4.4). Auch deuteten die deskriptiven Studienangaben aller eingeschlossenen Studien hinsichtlich Stichproben- und Diagnosemerkmalen auf einen homogenen Datensatz hin. Nichtsdestotrotz muss festgehalten werden, dass die Stichprobenvarianz im Durchschnitt und über alle Datensätze hinweg nur 38,21 % der Gesamtvarianz erklärte. Ein beträchtlicher Anteil an Heterogenität zwischen sowie innerhalb der Studien liegt somit vor, der nicht kontrolliert wurde bzw. werden konnte und folglich die Höhe der Effekte beeinflusst haben kann. Eine andere Auffassung zur Heterogenität bieten jedoch Kaufmann et al. (2013). Diese sehen Heterogenität als Merkmal einer Rechenstörung an. Danach zeigen Personen mit Rechenstörung ein Kerndefizit in den domänenspezifischen Fähigkeiten, welches von Defiziten in den domänenübergreifenden Defiziten begleitet sein kann. Die genaue Ausprägung und Schwere dieser Defizite hängt jedoch von der individuellen neurobiologischen und kognitiven Entwicklung jeder Person mit Rechenstörung ab und variiert dementsprechend von Person zu Person. Eine Schlussfolgerung dieser Annahme ist, dass Unterschiede zwischen Personen mit und ohne Rechenstörung somit weniger vom Studiendesign abhängen. Stattdessen ist davon auszugehen, dass aufgrund intraindividuelle neurobiologischer Entwicklungsunterschiede unterschiedliche Ausprägungen einer Rechenstörung innerhalb jeder Stichprobe von Personen mit Rechenstörungen vorliegen (McCaskey et al., 2018; Zhang et al., 2017). Diese Sichtweise findet auch ihre Bestätigung in der umfangreichen Forschungsliteratur zur Frage verschiedener Subtypen der Rechenstörung (Bartelet et al., 2014; Cirino et al., 2015; Gold et al., 2013; Landerl et al., 2004; Landerl et al., 2009; Shin & Bryant, 2015; Träff et al., 2017). Allerdings muss erwähnt werden, dass solche Forschungsfragen nur durch kontrollierte Studien beantwortet werden können. Diese Metaanalyse verneint daher nicht die Möglichkeit, dass verschiedene Subtypen der Rechenstörung dem Gesamtdatensatz zugrunde liegen und auf die Heterogenität einwirken. Jedoch ist es einerseits nicht möglich,

dies methodisch zu kontrollieren, noch ist es andererseits der Fokus dieser Arbeit, der darauf liegt, die stabilsten Defizite einer Rechenstörung zu ermitteln und diese in ein allgemeines kognitives Profil der Rechenstörung zu integrieren.

Drittens wurde ein Publikationsbias festgestellt, der die Ergebnisse dieser Metaanalyse beeinflusst haben kann. Obwohl mehrere Maßnahmen zur Kontrolle eines Biasrisikos vor der Datenauswertung durchgeführt wurden, weist Eggers Regressionsanalyse bei mehreren Outcomes auf einen Publikationsbias hin. Hierzu sei jedoch erwähnt, dass Eggers Regressionsanalyse anfällig für einen Fehler 1. Art ist, wenn die Studienanzahl gering ist, wie es in dieser Metaanalyse bei mehreren Outcomes der Fall ist (Sutton, 2009). Auch zeigte sich der Publikationstyp teilweise als signifikanter Moderator des Gesamteffekts. Überraschenderweise sind dabei die berichteten Effektstärken in den eingeschlossenen Dissertationen und Masterarbeiten größer als diejenigen in den Journalartikeln. Zusammengefasst sind jedoch nur 10 der insgesamt 54 eingeschlossenen Studien Dissertationen bzw. Masterarbeiten, von denen nicht jede Daten zu jedem Outcome dieser Metaanalyse erhob, weswegen es nicht möglich ist, valide Schlussfolgerungen über den moderierenden Effekt des Publikationstyps zu treffen.

Viertens wurde im Rahmen dieser Metaanalyse ADHS nicht ausgeschlossen. Ungefähr 10 bis 20 % aller Personen mit Rechenstörung weisen eine ADHS auf oder zeigen entsprechende Symptome (Fortes et al., 2016; Gross-Tsur et al., 1996). Kuhn et al. (2016) konnten dabei zeigen, dass das kognitive Profil von Personen mit komorbider Rechenstörung und ADHS lediglich eine Kombination der typischen Symptomatik einer Rechenstörung und einer ADHS ist und nicht zu einem neuen Muster an Symptomen führt. Da eine ADHS meistens mit Defiziten in der Aufmerksamkeit sowie Verarbeitungsgeschwindigkeit einhergeht, können die Ergebnisse dieser Metaanalyse zu diesen beiden Outcomes durch das Vorhandensein einer ADHS in den Stichproben der jeweiligen Studien verzerrt sein. Es sei jedoch erwähnt, dass im Gesamtdatensatz dieser Metaanalyse kaum eine Studie auf ADHS in den Stichproben kontrollierte. Daher war es nicht möglich, Studien mit komorbiden Fällen von Personen mit Rechenstörung und ADHS auszuschließen.

5.4 Implikationen für Theorie und Praxis

Dies ist die erste Metaanalyse mit systematischer Literaturrecherche, die die domänenspezifischen und domänenübergreifenden Defizite von Personen mit Rechenstörung sowohl bei den Fähigkeiten als auch bei den Testaufgaben untersucht und als Skala sowohl die Richtigkeit als auch die benötigte Zeit miteinander bezieht. Ein kognitives Profil stabiler Defizite konnte so ermittelt werden, aus dem sich mehrere Implikationen für die theoretische Betrachtung der Rechenstörung sowie ihrer praktischen Handhabung ableiten lassen.

Hinsichtlich der theoretischen Implikationen wurden stabile Defizite über mehrere Fähigkeiten und Testaufgaben hinweg gefunden, wenn die jeweiligen Items einen hohen Grad an visueller und visuell-räumlicher Verarbeitung bedurften (z. B. visuelle Verarbeitungsgeschwindigkeit, visuell-räumliches Kurzzeitgedächtnis, Testaufgaben zur Mengenverarbeitung). Obwohl die Effekte auch durch die übergeordnete Fähigkeit zur Raumvorstellung moderiert sein können, scheint ein funktioneller Zusammenhang zwischen der Verarbeitung visueller und visuell-räumlicher Informationen und den Defiziten bei der Mengen- und in weiterer Folge der Zahlenverarbeitung möglich. Da die Raumvorstellung jedoch auch eine essenzielle menschliche Fähigkeit und einen wichtigen Intelligenzfaktor darstellt (W. J. Schneider & McGrew, 2018), ist ihre Rolle bei der Herausbildung der domänenspezifischen Fähigkeiten, dabei insbesondere der Basiskompetenzen, weiter zu untersuchen.

Ein weiteres wichtiges Ergebnis ist, dass ebenso stabile Defizite in der benötigten Zeit ermittelt wurden, die sich insbesondere bei den Testaufgaben zur Aufmerksamkeit sowie zu den Basiskompetenzen als stabiler erweisen als ihre jeweils korrespondierenden Defizite in der Richtigkeit. Die benötigte Zeit wird jedoch im Vergleich zur Richtigkeit stärker von den allgemeinen Testbedingungen (z. B. lauter Raum) und der Stimmung der jeweiligen Testperson (z. B. Müdigkeit) beeinflusst. Zukünftige Studien sollten daher untersuchen, wie die Daten auf Basis der benötigten Zeit systematisch zur Diagnostik einer Rechenstörung verwendet werden können (z. B. Cut-Off-Werte, Normen). Derzeit werden die meisten der erhältlichen Tests zur Diagnostik einer Rechenstörung noch als Paper-Pencil-Format angeboten. Die Messung der benötigten Zeit wird in diesem Fall durch den/die Testleiter/in per Stoppuhr durchgeführt, was einen höhe-

ren Grad an Ungenauigkeit bedingt und die Vergleichbarkeit von Messwerten bei einer Testperson sowie zwischen mehreren Testpersonen reduziert. Mit zunehmender Digitalisierung der Schulen sowie Praxen und damit einfacheren, mobilen und womöglich kostengünstigeren Zugängen zu Rechenstörungstests ist jedoch mit einer Zunahme an digitalen Testangeboten zu rechnen, die wiederum eine exakte Messung der benötigten Zeit erlauben. Es ist hier Aufgabe der Forschung, für Testwerte auf Basis der benötigten Zeit entsprechende Normen oder Cut-Off-Werte bereitzustellen, die sich zur Diagnostik einer Rechenstörung eignen.

Diese Arbeit stellt zudem augenscheinlich die erste Metaanalyse dar, bei der eine Bottum-up-Kodierung zur Entwicklung des Kodierschemas angewandt wurde. Obwohl dieser Ansatz im Vergleich zu einer Top-down-Kodierung deutlich komplexer und arbeitsintensiver ist sowie eine permanente Adjustierung des Kodierschemas bedarf, reduziert er den Informationsverlust deutlich, indem keine Fähigkeiten oder Testaufgaben vorab ausgeschlossen werden. Ein umfassendes kognitives Profil der Rechenstörung konnte letztlich so gewonnen werden. Ein Nachteil dieser Vorgehensweise war allerdings, dass mehrere Outcomes von weniger als 5 Studien berichtet wurden. Als Folge davon wurde eine umfassende Analyse des Biasrisikos durchgeführt und ein entsprechendes statistisches Modell in Form eines multivariaten Mehrebenenmodells zur Auswertung gewählt. Die statistische Robustheit dieser Vorgehensweise wurde bisher jedoch nicht vollumfänglich überprüft. Dennoch möchte diese Metaanalyse auch andere systematische Reviews ermutigen, stärker datenbasiert vorzugehen (d. h. Bottum-up-Kodierung, Biasprüfung) und statistisch komplexere Modelle zu verwenden, die abhängige Effektstärken innerhalb sowie zwischen den Studien berücksichtigen. Auch wenn diese Modelle mit einem größeren Arbeitsaufwand sowie höheren Anforderungen für die statistische Robustheit einhergehen, so hat sich gezeigt, dass der selektive Ausschluss von Daten (z. B. univariate Auswertung, Nichtberücksichtigung abhängiger Effektstärken) zu anderen Ergebnissen und damit eventuell auch falschen Schlussfolgerungen führt (Riley, 2009).

Hinsichtlich der praktischen Implikationen wurde in dieser Metaanalyse ein Profil reliabler und stabiler kognitiver Defizite bei der Rechenstörung ermittelt, aus

dem sich Konsequenzen für die Diagnostik einer Rechenstörung ableiten lassen. Gemäß DSM-5 sowie ICD-11 müssen bei der Diagnostik der Rechenstörung domänenspezifische Defizite vorliegen. Anhand dieser Metaanalyse, die als erste systematisch die domänenspezifischen Defizite bei Personen mit Rechenstörung untersuchte, ist es nun leichter, den passenden Test auszuwählen, welcher die Testaufgaben beinhaltet und die Fähigkeiten erfasst, die am besten zwischen Personen mit und ohne Rechenstörung unterscheiden. Teile bzw. Vorläufer dieser Arbeit (siehe 3) flossen überdies als praktische Konsequenz bereits in Empfehlungen der im Februar 2018 erschienenen S3-Leitlinie zur „Diagnostik und Behandlung der Rechenstörung“ (Haberstroh & Schulte-Körne, 2019) ein. In dieser wird ein Profil der Rechenstörung definiert, welches über die bloße Bewertung der Rechenleistung, wie es von ICD-10 beschrieben wird, hinausgeht. Zum einen wird der Einbezug weiterer domänenspezifischer Fähigkeiten, insbesondere der Basiskompetenzen, zur Diagnosestellung empfohlen. Zum anderen wird auf die zusätzliche Bedeutung domänenübergreifender Fähigkeiten und dabei vor allem dem visuell-räumlichen Kurzzeitgedächtnis sowie der Inhibition hingewiesen. Eine Gruppe von Forscher/innen aus dem Bereich der Lernstörungen argumentierte überdies, dass nicht die domänenspezifischen Defizite, sondern hauptsächlich die domänenübergreifende Defizite zur Diagnostik einer Rechenstörung verwendet werden sollen (Hale et al., 2010). Ein sogenannter *pattern of strengths and weaknesses approach (PSW)* wurde beschrieben, dessen Ziel es ist, die Rechenstörung sowie andere Lernstörungen über die ihnen zugrundeliegenden charakteristischen Prozesse in allgemeinen kognitiven Fähigkeiten (z. B. Verarbeitungsgeschwindigkeit, Arbeitsgedächtnis) zu identifizieren (Maki, Floyd & Roberson, 2015). Ein solches Profil konnte bisher weder für die Rechenstörung noch für andere Lernstörungen ermittelt werden, weswegen es dem PSW-Ansatz weiterhin an Validität mangelt (McGill, Styck, Palomares & Hass, 2016; Miciak, Fletcher, Stuebing, Vaughn & Tolar, 2014). Nichtsdestotrotz ist diese Vorgehensweise in einigen Bundesstaaten der USA eine legitime Methode zur Diagnostik einer Rechenstörung (Maki et al., 2015). Die in dieser Metaanalyse ermittelten stabilen domänenübergreifenden Defizite bei einer Rechenstörung stellen damit nicht nur ganz konkret für den PSW-Ansatz, sondern auch für die übliche Diagnostik auf Basis domänenspezifischer Defizite eine wichtige Datengrundlage dar, wie und welche domänen-

übergreifenden Defizite bei der Diagnostik zusätzlich einbezogen werden können. Tabelle 10 zeigt zum Abschluss einen fiktiven Rechentest, in dem die Ergebnisse sowie Schlussfolgerungen dieser Arbeit zu den Fragestellungen und zum kognitiven Profil der Rechenstörung zusammengefasst sind und der folglich als Modell für zukünftige Rechentests gelten kann.

Tabelle 10

Inhalte eines fiktiven Rechentests auf Basis des kognitiven Profils der Rechenstörung

übergeordnete Fähigkeit	Richtigkeit		benötigte Zeit	
	untergeordnete Fähigkeit	Testaufgabe	untergeordnete Fähigkeit	Testaufgabe
Mathematik	Basiskompetenzen	Transkodieren	Basiskompetenzen	Simultanerfassung (ab 4) Zahlenvergleich Grundrechenarten
	Rechnen	Zahlenstrahl Grundrechenarten Faktenwissen Textaufgaben	Rechnen	
	Mathematisches Schlussfolgern			
Arbeitsgedächtnis	Arbeitsgedächtnis- kapazität	Spanne komplex		
Kurzzeitgedächtnis	Visuell-räumliches Kurzzeitgedächtnis	Abrufen (visuell- räumlich) Spanne vorwärts (vi- suell-räumlich)		
Aufmerksamkeit			Inhibition	Stroop-Effekt
Verarbeitungsgeschwindigkeit	Visuelle Verarbei- tungsgeschwindigkeit	Visueller Vergleich		

6 Zusammenfassung

3 bis 7 % aller Personen leiden unter einer Rechenstörung und zeigen, trotz normaler Intelligenz und regulärer Beschulung, ausgeprägte und andauernde Schwierigkeiten in spezifisch mathematischen sowie allgemeinen nicht-mathematischen Fähigkeiten und Testaufgaben. Über die genaue Vielfalt, Schwere und Stabilität dieser kognitiven Defizite herrscht jedoch weiterhin Unklarheit, was die Feststellung einer Rechenstörung erschwert und Fehldiagnosen begünstigt.

Ziel dieses systematischen Reviews war es daher, ein kognitives Profil der Rechenstörung zu ermitteln, welches diejenigen kognitiven Fähigkeiten und deren zugrundeliegenden Testaufgaben beschreibt, bei denen Personen mit Rechenstörung stabile Defizite aufweisen und die demzufolge im Rahmen der Diagnostik zu erfassen sind.

Hierzu wurde eine systematische Literaturrecherche nach Studien durchgeführt, die Personen mit und ohne Rechenstörung in ihrer kognitiven Leistung verglichen. 54 Studien wurden so identifiziert, die insgesamt 4452 Testpersonen umfassten und Daten zu 471 Outcomes berichteten. Anhand einer Bottom-Up-Kodierung sowie multivariaten Mehrebenen-Metaanalyse wurden schließlich 7 Fähigkeiten und 13 Testaufgaben ermittelt, bei denen Personen mit Rechenstörung stabile Defizite in der Richtigkeit und / oder der benötigten Zeit bei der Bearbeitung aufwiesen.

Das kognitive Profil der Rechenstörung setzte sich demnach zusammen aus spezifisch mathematischen Defiziten im Rechnen (Richtigkeit / benötigte Zeit), in den Basiskompetenzen (Richtigkeit / benötigte Zeit) und im mathematischen Schlussfolgern (Richtigkeit). Überdies lagen allgemeine nicht-mathematische Defizite in der Arbeitsgedächtniskapazität (Richtigkeit), dem visuell-räumlichen Kurzzeitgedächtnis (Richtigkeit), der Inhibition (benötigte Zeit) sowie der visuellen Verarbeitungsgeschwindigkeit (Richtigkeit) vor. Die Effektstärken aller Fähigkeiten und deren korrespondierenden Testaufgaben waren dabei mittel bis groß. Insbesondere Testaufgaben, bei denen visuelle und visuell-räumliche Informationen verarbeitet wurden, führten zu stabileren Defiziten.

Das so ermittelte kognitive Profil der Rechenstörung kann durch den ausschließlichen Einbezug der ermittelten stabilen Defizite die Diagnosesicherheit erhöhen. Der dezidierte Zusammenhang der Rechenstörung mit Defiziten in der Verarbeitung visueller und visuell-räumlicher Informationen bedarf weiterer Klärung.

Literaturverzeichnis

Studien, die mit einem Asterisk gekennzeichnet sind, waren in die Metaanalyse eingeschlossen.

- Aben, B., Stapert, S. & Blokland, A. (2012). About the distinction between working memory and short-term memory. *Frontiers in Psychology*, 3, 1-9. doi:10.3389/fpsyg.2012.00301
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5. Aufl.). Arlington, VA, US: Author.
- Arsalidou, M. & Taylor, M. J. (2011). Is $2+2=4$? Meta-analyses of brain areas needed for numbers and calculations. *NeuroImage*, 54(3), 2382-2393. doi:10.1016/j.neuroimage.2010.10.009
- *Ashkenazi, S., Mark-Zigdon, N. & Henik, A. (2013). Do subitizing deficits in developmental dyscalculia involve pattern recognition weakness? *Developmental Science*, 16(1), 35-46. doi:10.1111/j.1467-7687.2012.01190.x
- Ashkenazi, S., Rosenberg-Lee, M., Tenison, C. & Menon, V. (2012). Weak task-related modulation and stimulus representations during arithmetic problem solving in children with developmental dyscalculia. *Developmental Cognitive Neuroscience*, 2S, 152-166. doi:10.1016/j.dcn.2011.09.006
- Assink, M. & Wibbelink, C. J. M. (2016). Fitting three-level meta-analytic models in R: A step-by-step tutorial. *Tutorials in Quantitative Methods for Psychology*, 12(3), 154-174. doi:10.20982/tqmp.12.3.p154
- *Attout, L. & Majerus, S. (2015). Working memory deficits in developmental dyscalculia: The importance of serial order. *Child Neuropsychology*, 21(4), 432-450. doi:10.1080/09297049.2014.922170
- Baddeley, A. D. (2012). Working memory: Theories, models, and controversies. *Annual Review of Psychology*, 63, 1-29. doi:10.1146/annurev-psych-120710-100422
- Baddeley, A. D. & Hitch, G. (1974). Working memory. In G. H. Bower (Hrsg.), *Psychology of learning and motivation: Advances in research and theory* (Bd. 8, S. 47-89). New York, NY, US: Academic Press.
- Baker, J. M. & Reiss, A. L. (2016). A meta-analysis of math performance in Turner syndrome. *Developmental Medicine & Child Neurology*, 58(2), 123-130. doi:10.1111/dmcn.12961
- Barkley, R. A. (1997). Behavioral inhibition, sustained attention, and executive functions: Constructing a unifying theory of ADHD. *Psychological Bulletin*, 121(1), 65-94. doi:10.1037/0033-2909.121.1.65
- Bartelet, D., Ansari, D., Vaessen, A. & Blomert, L. (2014). Cognitive subtypes of mathematics learning difficulties in primary education. *Research in Developmental Disabilities*, 35(3), 657-670. doi:10.1016/j.ridd.2013.12.010
- Becker, B. J. (2000). Multivariate meta-analysis. In H. E. A. Tinsley & S. D. Brown (Hrsg.), *Handbook of applied multivariate statistics and mathematical modeling* (S. 499-525). San Diego, CA, US: Academic Press.

- Belsley, D. A., Kuh, E. & Welsch, R. E. (1980). *Regression diagnostics: Identifying influential data and sources of collinearity*. Hoboken, NJ, US: Wiley.
- Berch, D. B. (2005). Making sense of number sense: Implications for children with mathematical disabilities. *Journal of Learning Disabilities, 38*(4), 333-339. doi:10.1177/00222194050380040901
- Booth, J. N., Boyle, J. M. E. & Kelly, S. W. (2010). Do tasks make a difference? Accounting for heterogeneity of performance of children with reading difficulties on tasks of executive function: Findings from a meta-analysis. *British Journal of Developmental Psychology, 28*(1), 133-176. doi:10.1348/026151009X485432
- Borenstein, M. (2009). Effect sizes for continuous data. In H. Cooper, L. V. Hedges & J. C. Valentine (Hrsg.), *The handbook of research synthesis and meta-analysis* (2. Aufl., S. 221-235). New York, NY, US: Russell Sage Foundation.
- Borenstein, M., Hedges, L. V., Higgins, J. P. T. & Rothstein, H. R. (2009). *Introduction to meta-analysis*. Chichester, UK: Wiley.
- Brankaer, C., Ghesquière, P., de Wel, A., Swillen, A. & de Smedt, B. (2016). Numerical magnitude processing impairments in genetic syndromes: A cross-syndrome comparison of Turner and 22q11.2 deletion syndromes. *Developmental Science, 20*(6), 1-14. doi:10.1111/desc.12458
- Brown, S. A., Upchurch, S. L. & Acton, G. J. (2003). A framework for developing a coding scheme for meta-analysis. *Western Journal of Nursing Research, 25*(2), 205-222. doi:10.1177/0193945902250038
- Bull, R. & Johnston, R. S. (1997). Children's arithmetical difficulties: Contributions from processing speed, item identification, and short-term memory. *Journal of Experimental Child Psychology, 65*(1), 1-24. doi:10.1006/jecp.1996.2358
- Busch, J., Schmidt, C. & Grube, D. (2015). Arithmetic fact retrieval: Are there differences between children with developmental dyscalculia and those with mathematical difficulties? *Zeitschrift für Psychologie, 223*(2), 110-119. doi:10.1027/2151-2604/a000209
- Butterworth, B., Varm, S. & Laurillard, D. (2011). Dyscalculia: From brain to education. *Science, 332*(6033), 1049-1053. doi:10.1126/science.1201536
- Casey, B. M., Pezaris, E., Fineman, B., Pollock, A., Demers, L. & Dearing, E. (2015). A longitudinal analysis of early spatial skills compared to arithmetic and verbal skills as predictors of fifth-grade girls' math reasoning. *Learning and Individual Differences, 40*, 90-100. doi:10.1016/j.lindif.2015.03.028
- *Censabella, S. & Noël, M. P. (2008). The inhibition capacities of children with mathematical disabilities. *Child Neuropsychology, 14*(1), 1-20. doi:10.1080/09297040601052318
- *Ceulemans, A., Titeca, D., Loeys, T., Hoppenbrouwers, K., Rousseau, S. & Desoete, A. (2014). Enumeration of small and large numerosities in adolescents with mathematical learning disorders. *Research in Developmental Disabilities, 35*(1), 27-35. doi:10.1016/j.ridd.2013.10.018
- Chaiken, S. R., Kyllonen, P. C. & Tirre, W. C. (2000). Organization and components of psychomotor ability. *Cognitive Psychology, 40*(3), 198-226. doi:10.1006/cogp.1999.0729

- Cheng, Y.-L. & Mix, K. S. (2014). Spatial training improves children's mathematics ability. *Journal of Cognition and Development, 15*(1), 2-11. doi:10.1080/15248372.2012.725186
- Cheung, M. W. L. (2013). Multivariate meta-analysis as structural equation models. *Structural Equation Modeling: A Multidisciplinary Journal, 20*(3), 429-454. doi:10.1080/10705511.2013.797827
- Cheung, M. W. L. (2014). Modeling dependent effect sizes with three-level meta-analyses: A structural equation modeling approach. *Psychological Methods, 19*(2), 211-229. doi:10.1037/a0032968
- Cheung, M. W. L. (2015). *Meta-analysis: A structural equation modeling approach*. Chichester, UK: Wiley.
- Cheung, S. F. & Chan, D. K. S. (2014). Meta-analyzing dependent correlations: An SPSS macro and an R script. *Behavior Research Methods, 46*(2), 331-345. doi:10.3758/s13428-013-0386-2
- *Cirino, P. T., Fuchs, L. S., Elias, J. T., Powell, S. R. & Schumacher, R. F. (2015). Cognitive and mathematical profiles for different forms of learning difficulties. *Journal of Learning Disabilities, 48*(2), 156-175. doi:10.1177/0022219413494239
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20*, 37-46. doi:10.1177/001316446002000104
- Cohen, J. (1969). *Statistical power analysis for the behavioral sciences*. New York, NY, US: Academic Press.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2. Aufl.). Hillsdale, NJ, US: Lawrence Erlbaum Associates.
- *Compton, D. L., Fuchs, L. S., Fuchs, D., Lambert, W. & Hamlett, C. (2012). The cognitive and academic profiles of reading and mathematics learning disabilities. *Journal of Learning Disabilities, 45*(1), 79-95. doi:10.1177/0022219410393012
- Conway, A. R. A., Cowan, N., Bunting, M. F., Theriault, D. J. & Minkoff, S. R. B. (2002). A latent variable analysis of working memory capacity, short-term memory capacity, processing speed, and general fluid intelligence. *Intelligence, 30*(2), 163-183. doi:10.1016/S0160-2896(01)00096-4
- Cook, R. D. & Weisberg, S. (1982). *Residuals and influence in regression*. New York, NY, US: Chapman and Hall.
- Cowan, R. & Powell, D. (2014). The contributions of domain-general and numerical factors to third-grade arithmetic skills and mathematical learning disability. *Journal of Educational Psychology, 106*(1), 214-229. doi:10.1037/a0034097
- D'Amico, A. & Passolunghi, M. C. (2009). Naming speed and effortful and automatic inhibition in children with arithmetic learning disabilities. *Learning and Individual Differences, 19*(2), 170-180. doi:10.1016/j.lindif.2009.01.001
- David, C. V. (2012). Working memory deficits in math learning difficulties: A meta-analysis. *International Journal of Developmental Disabilities, 58*(2), 67-84. doi:10.1179/2047387711Y.0000000007
- Davis, N., Cannistraci, C. J., Rogers, B. P., Gatenby, J. C., Fuchs, L. S., Anderson, A. W. & Gore, J. C. (2009). Aberrant functional activation in school age children at-risk for mathematical disability: A functional imaging study of simple arithmetic skill. *Neuropsychologia, 47*(12), 2470-2479. doi:10.1016/j.neuropsychologia.2009.04.024

- de Smedt, B. & Boets, B. (2010). Phonological processing and arithmetic fact retrieval: Evidence from developmental dyslexia. *Neuropsychologia*, 48(14), 3973-3981. doi:10.1016/j.neuropsychologia.2010.10.018
- de Smedt, B. & Gilmore, C. (2011). Defective number module or impaired access? Numerical magnitude processing in first graders with mathematical difficulties. *Journal of Experimental Child Psychology*, 108(2), 278-292. doi:10.1016/j.jecp.2010.09.003
- de Smedt, B., Noël, M.-P., Gilmore, C. & Ansari, D. (2013). How do symbolic and non-symbolic numerical magnitude processing skills relate to individual differences in children's mathematical skills? A review of evidence from brain and behavior. *Trends in Neuroscience and Education*, 2(2), 48-55. doi:10.1016/j.tine.2013.06.001
- de Smedt, B., Swillen, A., Verschaffel, L. & Ghesquiere, P. (2009). Mathematical learning disabilities in children with 22q11.2 deletion syndrome: A review. *Developmental Disabilities Research Reviews*, 15(1), 4-10. doi:10.1002/ddrr.44
- *de Weerd, F., Desoete, A. & Roeyers, H. (2013). Working memory in children with reading disabilities and/or mathematical disabilities. *Journal of Learning Disabilities*, 46(5), 461-472. doi:10.1177/0022219412455238
- Dehaene, S. (1992). Varieties of numerical abilities. *Cognition*, 44(1), 1-42. doi:10.1016/0010-0277(92)90049-N
- Desoete, A., Ceulemans, A., de Weerd, F. & Pieters, S. (2012). Can we predict mathematical learning disabilities from symbolic and non-symbolic comparison tasks in kindergarten? Findings from a longitudinal study. *British Journal of Educational Psychology*, 82(1), 64-81. doi:10.1348/2044-8279.002002
- Deutsches Institut für Medizinische Dokumentation und Information. (2016). MeSH - Medical Subject Headings. Verfügbar unter <https://www.dimdi.de/dynamic/de/klassifikationen/weitere-klassifikationen-und-standards/mesh/> [19.02.2019].
- Devine, A., Soltész, F., Nobes, A., Goswami, U. & Szűcs, D. (2013). Gender differences in developmental dyscalculia depend on diagnostic criteria. *Learning and Instruction*, 27, 31-39. doi:10.1016/j.learninstruc.2013.02.004
- Dirks, E., Spyer, G., van Lieshout, E. C. D. M. & de Sonneville, L. (2008). Prevalence of combined reading and arithmetic disabilities. *Journal of Learning Disabilities*, 41(5), 460-473. doi:10.1177/0022219408321128
- *Donker, M., Kroesbergen, E. H., Slot, E., van Viersen, S. & de Bree, E. (2016). Alphanumeric and non-alphanumeric rapid automatized naming in children with reading and/or spelling difficulties and mathematical difficulties. *Learning and Individual Differences*, 47, 80-87. doi:10.1016/j.lindif.2015.12.011
- Downes, M. J., Brennan, M. L., Williams, H. C. & Dean, R. S. (2016). Development of a critical appraisal tool to assess the quality of cross-sectional studies (AXIS). *BMJ Open*, 6(12), 1-7. doi:10.1136/bmjopen-2016-011458
- Dwan, K., Gamble, C., Williamson, P. R., Kirkham, J. J. & Reporting Bias Group. (2013). Systematic review of the empirical evidence of study publication bias and outcome reporting bias - An updated review. *Plos One*, 8(7), 1-37. doi:10.1371/journal.pone.0066844

- Ebrahim, S., Montoya, L., el Din, M. K., Sohani, Z. N., Agarwal, A., Bance, S., . . . Ioannidis, J. P. A. (2016). Randomized trials are frequently fragmented in multiple secondary publications. *Journal of Clinical Epidemiology*, *79*, 130-139. doi:10.1016/j.jclinepi.2016.05.016
- Egger, M., Smith, G. D., Schneider, M. & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal*, *315*(7109), 629-634. doi:10.1136/bmj.315.7109.629
- Endlich, D., Dummert, F., Schneider, W. & Schwenck, C. (2014). Verhaltensprobleme bei Kindern mit umschriebener und kombinierter schulischer Minderleistung. *Kindheit und Entwicklung*, *23*(1), 61-69. doi:10.1026/0942-5403/a000128
- *Evans, M. (2008). *Underlying cognitive processes in reading, math, and comorbid reading and math learning disabilities*. (Doctoral dissertation). Verfügbar unter ProQuest Dissertations & Theses Global. (UMI No. 304806797)
- Fazio, L. K., Bailey, D. H., Thompson, C. A. & Siegler, R. S. (2014). Relations of different types of numerical magnitude representations to each other and to mathematics achievement. *Journal of Experimental Child Psychology*, *123*, 53-72. doi:10.1016/j.jecp.2014.01.013
- Fischbach, A., Schuchardt, K., Brandenburg, J., Kleszczewski, J., Balke-Melcher, C., Schmidt, C., . . . Hasselhorn, M. (2013). Prävalenz von Lernschwächen und Lernstörungen: Zur Bedeutung der Diagnosekriterien. [Prevalence of poor learners and children with learning disorders: Investigating the role of diagnostic criteria]. *Lernen und Lernstörungen*, *2*(2), 65-76. doi:10.1024/2235-0977/a000035
- Fischbach, A., Schuchardt, K., Mähler, C. & Hasselhorn, M. (2010). Zeigen Kinder mit schulischen Minderleistungen sozio-emotionale Auffälligkeiten? [Do children with low scholastic achievement show social-emotional problems?]. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, *42*(4), 201-210. doi:10.1026/0049-8637/a000025
- Fortes, I. S., Paula, C. S., Oliveira, M. C., Bordin, I. A., de Jesus Mari, J. & Rohde, L. A. (2016). A cross-sectional study to assess the prevalence of DSM-5 specific learning disorders in representative school samples from the second to sixth grade in Brazil. *European Child & Adolescent Psychiatry*, *25*(2), 195-207. doi:10.1007/s00787-015-0708-2
- Fuchs, L. S., Fuchs, D., Stuebing, K. K., Fletcher, J. M., Hamlett, C. L. & Lambert, W. (2008). Problem solving and computational skill: Are they shared or distinct aspects of mathematical cognition? *Journal of Educational Psychology*, *100*(1), 30-47. doi:10.1037/0022-0663.100.1.30
- Geary, D. C. (2004). Mathematics and learning disabilities. *Journal of Learning Disabilities*, *37*(1), 4-15. doi:10.1177/00222194040370010201
- Geary, D. C. (2010). Mathematical disabilities: Reflections on cognitive, neuropsychological, and genetic components. *Learning and Individual Differences*, *20*(2), 130-133. doi:10.1016/j.lindif.2009.10.008
- Geary, D. C. (2011). Consequences, characteristics, and causes of mathematical learning disabilities and persistent low achievement in mathematics. *Journal of Developmental and Behavioral Pediatrics*, *32*(3), 250-263. doi:10.1097/DBP.0b013e318209edef
- Geary, D. C., Hamson, C. O. & Hoard, M. K. (2000). Numerical and arithmetical cognition: A longitudinal study of process and concept deficits in children

- with learning disability. *Journal of Experimental Child Psychology*, 77(3), 236-263. doi:10.1006/jecp.2000.2561
- Geary, D. C., Hoard, M. K., Nugent, L. & Bailey, D. H. (2012). Mathematical cognition deficits in children with learning disabilities and persistent low achievement: A five-year prospective study. *Journal of Educational Psychology*, 104(1), 206-223. doi:10.1037/a0025398
- Geary, D. C., Hoard, M. K., Nugent, L. & Byrd-Craven, J. (2008). Development of number line representations in children with mathematical learning disability. *Developmental Neuropsychology*, 33(3), 277-299. doi:10.1080/87565640801982361
- Gersten, R., Jordan, N. C. & Flojo, J. R. (2005). Early identification and interventions for students with mathematics difficulties. *Journal of Learning Disabilities*, 38(4), 293-304. doi:10.1177/00222194050380040301
- Gleser, L. J. & Olkin, I. (2009). Stochastically dependent effect sizes. In H. Cooper, L. V. Hedges & J. C. Valentine (Hrsg.), *The handbook of research synthesis and meta-analysis* (2. Aufl., S. 357-376). New York, NY, US: Russell Sage Foundation.
- *Gold, A. B., Ewing-Cobbs, L., Cirino, P. T., Fuchs, L. S., Stuebing, K. K. & Fletcher, J. M. (2013). Cognitive and behavioral attention in children with math difficulties. *Child Neuropsychology*, 19(4), 420-437. doi:10.1080/09297049.2012.690371
- Grabner, R. H., Ansari, D., Koschutnig, K., Reishofer, G., Ebner, F. & Neuper, C. (2009). To retrieve or to calculate? Left angular gyrus mediates the retrieval of arithmetic facts during problem solving. *Neuropsychologia*, 47(2), 604-608. doi:10.1016/j.neuropsychologia.2008.10.013
- Gross-Tsur, V., Manor, O. & Shalev, R. S. (1996). Developmental dyscalculia: Prevalence and demographic features. *Developmental Medicine & Child Neurology*, 38(1), 25-33. doi:10.1111/j.1469-8749.1996.tb15029.x
- Haberstroh, S. & Schulte-Körne, G. (2019). Diagnostik und Behandlung der Rechenstörung. *Deutsches Ärzteblatt International*, 116(7), 107-114. doi:10.3238/arztebl.2019.0107
- Haffner, J., Baro, K., Parzer, P. & Resch, F. (2005). *Heidelberger Rechentest (HRT 1-4)*. Göttingen: Hogrefe.
- Hale, J., Alfonso, V., Berninger, V., Bracken, B., Christo, C., Clark, E., . . . Yalof, J. (2010). Critical issues in response-to-intervention, comprehensive evaluation, and specific learning disabilities identification and intervention: An expert white paper consensus. *Learning Disability Quarterly*, 33(3), 223-236. doi:10.1177/073194871003300310
- Hallgren, K. A. (2012). Computing inter-rater reliability for observational data: An overview and tutorial. *Tutorials in Quantitative Methods for Psychology*, 8(1), 23-34. doi:10.20982/tqmp.08.1.p023
- Hanich, L. B., Jordan, N. C., Kaplan, D. & Dick, J. (2001). Performance across different areas of mathematical cognition in children with learning difficulties. *Journal of Educational Psychology*, 93(3), 615-626. doi:10.1037/0022-0663.93.3.615
- Harper, P. (2005). *Labour force survey standard errors* (ABS Catalogue No. 6298.0). Verfügbar unter Australian Bureau of Statistics: [http://www.ausstats.abs.gov.au/ausstats/subscriber.nsf/0/49DE00F8584FDE17CA2570750075A1A5/\\$File/62980_2005.pdf](http://www.ausstats.abs.gov.au/ausstats/subscriber.nsf/0/49DE00F8584FDE17CA2570750075A1A5/$File/62980_2005.pdf) [Access 2005].

- Hecht, S. A., Torgesen, J. K., Wagner, R. K. & Rashotte, C. A. (2001). The relations between phonological processing abilities and emerging individual differences in mathematical computation skills: A longitudinal study from second to fifth grades. *Journal of Experimental Child Psychology*, 79(2), 192-227. doi:10.1006/jecp.2000.2586
- Hedges, L. V. (1981). Distribution theory for glass's estimator of effect size and related estimators. *Journal of Educational Statistics*, 6(2), 107-128. doi:10.2307/1164588
- Hedges, L. V., Tipton, E. & Johnson, M. C. (2010). Robust variance estimation in meta-regression with dependent effect size estimates. *Research Synthesis Methods*, 1(1), 39-65. doi:10.1002/jrsm.5
- Hegarty, M. & Kozhevnikov, M. (1999). Types of visual-spatial representations and mathematical problem solving. *Journal of Educational Psychology*, 91(4), 684-689. doi:10.1037/0022-0663.91.4.684
- Hegarty, M. & Waller, D. A. (2005). Individual differences in spatial abilities. In A. Miyake & P. Shah (Hrsg.), *The Cambridge handbook of visuospatial thinking* (S. 121-169). Cambridge, UK: Cambridge University Press.
- *Heine, A., Wissmann, J., Tamm, S., De Smedt, B., Schneider, M., Stern, E., . . . Jacobs, A. M. (2013). An electrophysiological investigation of non-symbolic magnitude processing: Numerical distance effects in children with and without mathematical learning disabilities. *Cortex*, 49(8), 2162-2177. doi:10.1016/j.cortex.2012.11.009
- Henik, A., Rubinsten, O. & Ashkenazi, S. (2011). The 'where' and 'what' in developmental dyscalculia. *The Clinical Neuropsychologist*, 25(6), 989-1008. doi:10.1080/13854046.2011.599820
- Higgins, J. P. T. & Green, S. (Hrsg.). (2011). *Cochrane handbook for systematic reviews of interventions version 5.1.0 [updated March 2011]*: The Cochrane Collaboration.
- Higgins, J. P. T. & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine*, 21(11), 1539-1558. doi:10.1002/sim.1186
- Higgins, J. P. T. & Thompson, S. G. (2004). Controlling the risk of spurious findings from meta-regression. *Statistics in Medicine*, 23(11), 1663-1682. doi:10.1002/sim.1752
- Hirsch, S., Lambert, K., Coppens, K. & Moeller, K. (2018). Basic numerical competences in large-scale assessment data: Structure and long-term relevance. *Journal of Experimental Child Psychology*, 167, 32-48. doi:10.1016/j.jecp.2017.09.015
- Hyde, J. S. (2016). Sex and cognition: Gender and cognitive functions. *Current Opinion in Neurobiology*, 38, 53-56. doi:10.1016/j.conb.2016.02.007
- Iglesias-Sarmiento, V. & Deaño, M. (2016). Arithmetical difficulties and low arithmetic achievement: Analysis of the underlying cognitive functioning. *The Spanish Journal of Psychology*, 19, 1-14. doi:10.1017/sjp.2016.40
- Jackson, D., Riley, R. & White, I. R. (2011). Multivariate meta-analysis: Potential and promise. *Statistics in Medicine*, 30(20), 2481-2498. doi:10.1002/sim.4172
- Jaekel, J. & Wolke, D. (2014). Preterm birth and dyscalculia. *Journal of Pediatrics*, 164(6), 1327-1332. doi:10.1016/j.jpeds.2014.01.069
- Johnson, E. S., Humphrey, M., Mellard, D. F., Woods, K. & Swanson, H. L. (2010). Cognitive processing deficits and students with specific learning

- disabilities: A selective meta-analysis of the literature. *Learning Disability Quarterly*, 33(1), 3-18. doi:10.1177/073194871003300101
- *Jolles, D., Ashkenazi, S., Kochalka, J., Evans, T., Richardson, J., Rosenberg-Lee, M., . . . Menon, V. (2016). Parietal hyper-connectivity, aberrant brain organization, and circuit-based biomarkers in children with mathematical disabilities. *Developmental Science*, 19(4), 613-631. doi:10.1111/desc.12399
- Jordan, N. C., Glutting, J. & Ramineni, C. (2010). The importance of number sense to mathematics achievement in first and third grades. *Learning and Individual Differences*, 20(2), 82-88. doi:10.1016/j.lindif.2009.07.004
- Jordan, N. C. & Hanich, L. B. (2003). Characteristics of children with moderate mathematics deficiencies: A longitudinal perspective. *Learning Disabilities Research & Practice*, 18(4), 213-221. doi:10.1111/1540-5826.00076
- Jordan, N. C., Kaplan, D., Ramineni, C. & Locuniak, M. N. (2009). Early math matters: Kindergarten number competence and later mathematics outcomes. *Developmental Psychology*, 45(3), 850-867. doi:10.1037/a0014939
- Jüni, P., Witschi, A., Bloch, R. & Egger, M. (1999). The hazards of scoring the quality of clinical trials for meta-analysis. *JAMA*, 282(11), 1054-1060. doi:10.1001/jama.282.11.1054
- Karagiannakis, G. N., Baccaglini-Frank, A. E. & Papadatos, Y. (2014). Mathematical learning difficulties subtypes classification. *Frontiers in Human Neuroscience*, 8, 1-5. doi:10.3389/fnhum.2014.00057
- Kassai, R., Futo, J., Demetrovics, Z. & Takacs, Z. K. (2019). A meta-analysis of the experimental evidence on the near- and far-transfer effects among children's executive function skills. *Psychological Bulletin*, 145(2), 165-188. doi:10.1037/bul0000180
- Kaufmann, L., Mazzocco, M. M., Dowker, A., von Aster, M., Göbel, S. M., Grabner, R. H., . . . Nuerk, H.-C. (2013). Dyscalculia from a developmental and differential perspective. *Frontiers in Psychology*, 4, 1-5. doi:10.3389/fpsyg.2013.00516
- Kaufmann, L. & von Aster, M. (2012). The diagnosis and management of dyscalculia. *Deutsches Ärzteblatt International*, 109(45), 767-778. doi:10.3238/arztebl.2012.0767
- Kaufmann, L., Wood, G., Rubinsten, O. & Henik, A. (2011). Meta-analyses of developmental fMRI studies investigating typical and atypical trajectories of number processing and calculation. *Developmental Neuropsychology*, 36(6), 763-787. doi:10.1080/87565641.2010.549884
- Kenneth, A. B. & Robert, W. J. (1985). Regression diagnostics: An expository treatment of outliers and influential cases. *Sociological Methods & Research*, 13(4), 510-542. doi:10.1177/0049124185013004004
- Keys, B. A. & White, D. A. (2000). Exploring the relationship between age, executive abilities, and psychomotor speed. *Journal of the International Neuropsychological Society*, 6(1), 76-82. doi:10.1017/S1355617700611098
- Klein, R. J., Proctor, S. E., Boudreault, M. A. & Turczyn, K. M. (2002). *Healthy people 2010: Criteria for data suppression* (Statistical Notes No. 24). Verfügbar unter Centers for Disease Control and Prevention: <https://www.cdc.gov/nchs/data/statnt/statnt24.pdf> [Access 2002].

- Knapp, G. & Hartung, J. (2003). Improved tests for a random effects meta-regression with a single covariate. *Statistics in Medicine*, 22(17), 2693-2710. doi:10.1002/sim.1482
- Kohn, J., Wyschkon, A. & Esser, G. (2013). Psychische Auffälligkeiten bei Umschriebenen Entwicklungsstörungen: Gibt es Unterschiede zwischen Lese-Rechtschreib- und Rechenstörungen? [Mental health problems in specific developmental disorders: Are there differences between dyslexia and dyscalculia?]. *Lernen und Lernstörungen*, 2(1), 7-20. doi:10.1024/2235-0977/a000027
- *Koontz, K. L. & Berch, D. B. (1996). Identifying simple numerical stimuli: Processing inefficiencies exhibited by arithmetic learning disabled children. *Mathematical Cognition*, 2(1), 1-23. doi:10.1080/135467996387525
- *Kucian, K. (2005). *Development of cerebral representations for numbers in normally achieving children and children with developmental dyscalculia*. (Doctoral dissertation). Verfügbar unter ProQuest Dissertations & Theses Global. (UMI No. 305368810)
- Kucian, K. (2016). Developmental dyscalculia and the brain. In D. B. Berch, D. C. Geary & K. Mann Koepke (Hrsg.), *Development of mathematical cognition: Neural substrates and genetic influences*. (Bd. 2, S. 165-193). San Diego, CA, US: Elsevier Academic Press.
- Kucian, K., Loenneker, T., Martin, E. & von Aster, M. (2011). Non-symbolic numerical distance effect in children with and without developmental dyscalculia: A parametric fMRI study. *Developmental Neuropsychology*, 36(6), 741-762. doi:10.1080/87565641.2010.549867
- Kuhn, J.-T. (2015). Developmental dyscalculia: Neurobiological, cognitive, and developmental perspectives. *Zeitschrift für Psychologie*, 223(2), 69-82. doi:10.1027/2151-2604/a000205
- Kuhn, J.-T., Ise, E., Raddatz, J., Schwenk, C. & Dobel, C. (2016). Basic numerical processing, calculation, and working memory in children with dyscalculia and/or ADHD symptoms. *Zeitschrift für Kinder- und Jugendpsychiatrie und Psychotherapie*, 44(5), 365-375. doi:10.1024/1422-4917/a000450
- Kuhn, J.-T., Schwenk, C., Raddatz, J., Dobel, C. & Holling, H. (2017). CODY-Mathetest: Mathematiktest für die 2.-4. Klasse (CODY-M 2-4). Düsseldorf: Kaasa health.
- Landerl, K. (2013). Development of numerical processing in children with typical and dyscalculic arithmetic skills - A longitudinal study. *Frontiers in Psychology*, 4, 1-14. doi:10.3389/fpsyg.2013.00459
- Landerl, K., Bevan, A. & Butterworth, B. (2004). Developmental dyscalculia and basic numerical capacities: A study of 8–9-year-old students. *Cognition*, 93(2), 99-125. doi:10.1016/j.cognition.2003.11.004
- Landerl, K., Fussenegger, B., Moll, K. & Willburger, E. (2009). Dyslexia and dyscalculia: Two learning disorders with different cognitive profiles. *Journal of Experimental Child Psychology*, 103(3), 309-324. doi:10.1016/j.jecp.2009.03.006
- Landerl, K. & Moll, K. (2010). Comorbidity of learning disorders: Prevalence and familial transmission. *Journal of Child Psychology and Psychiatry*, 51(3), 287-294. doi:10.1111/j.1469-7610.2009.02164.x
- Lee, K., Ng, S. F., Bull, R., Pe, M. L. & Ho, R. H. M. (2011). Are patterns important? An investigation of the relationships between proficiencies in

- patterns, computation, executive functioning, and algebraic word problems. *Journal of Educational Psychology*, 103(2), 269-281. doi:10.1037/a0023068
- Lewis, K. E. & Fisher, M. B. (2016). Taking stock of 40 years of research on mathematical learning disability: Methodological issues and future directions. *Journal for Research in Mathematics Education*, 47(4), 338-371. doi:10.5951/jresmetheduc.47.4.0338
- Lindquist, M., Philpot, R., Mullis, I. V. S. & Cotter, K. E. (2017). TIMSS 2019 mathematics framework In I. V. S. Mullis & M. O. Martin (Hrsg.), *TIMSS 2019 assessment frameworks* (S. 11-26). Boston, MA, US: International Association for the Evaluation of Educational Achievement.
- Lipsey, M. W. (2009). Identifying interesting variables and analysis opportunities. In H. Cooper, L. V. Hedges & J. C. Valentine (Hrsg.), *The handbook of research synthesis and meta-analysis* (2. Aufl., S. 147-148). New York, NY, US: Russell Sage Foundation.
- MacCallum, R. C. & Austin, J. T. (2000). Applications of structural equation modeling in psychological research. *Annual Review of Psychology*, 51, 201-226. doi:10.1146/annurev.psych.51.1.201
- *MacKinnon McQuarrie, M. A., Siegel, L. S., Perry, N. E. & Weinberg, J. (2014). Reactivity to stress and the cognitive components of math disability in grade 1 children. *Journal of Learning Disabilities*, 47(4), 349-365. doi:10.1177/0022219412463436
- *Mähler, C. & Schuchardt, K. (2016). Working memory in children with specific learning disorders and/or attention deficits. *Learning and Individual Differences*, 49, 341-347. doi:10.1016/j.lindif.2016.05.007
- Maki, K. E., Floyd, R. G. & Roberson, T. (2015). State learning disability eligibility criteria: A comprehensive review. *School Psychology Quarterly*, 30(4), 457-469. doi:10.1037/spq0000109
- Mammarella, I. C., Caviola, S., Cornoldi, C. & Lucangeli, D. (2013). Mental additions and verbal-domain interference in children with developmental dyscalculia. *Research in Developmental Disabilities*, 34(9), 2845-2855. doi:10.1016/j.ridd.2013.05.044
- *Mammarella, I. C., Caviola, S., Giofrè, D. & Szűcs, D. (2018). The underlying structure of visuospatial working memory in children with mathematical learning disability. *British Journal of Developmental Psychology*, 36(2), 220-235. doi:10.1111/bjdp.12202
- *Mammarella, I. C., Hill, F., Devine, A., Caviola, S. & Szűcs, D. (2015). Math anxiety and developmental dyscalculia: A study on working memory processes. *Journal of Clinical and Experimental Neuropsychology*, 37(8), 878-887. doi:10.1080/13803395.2015.1066759
- Mavridis, D. & Salanti, G. (2013). A practical introduction to multivariate meta-analysis. *Statistical Methods in Medical Research*, 22(2), 133-158. doi:10.1177/0962280211432219
- *McCall, C. A. (1999). *Math computation difficulties in grade 7 and 8 students*. (Doctoral dissertation). Verfügbar unter ProQuest Dissertations & Theses Global. (UMI No. 304576902)
- McCaskey, U., von Aster, M., Maurer, U., Martin, E., O'Gorman Tuura, R. & Kucian, K. (2018). Longitudinal brain development of numerical skills in typically developing children and children with developmental dyscalculia. *Frontiers in Human Neuroscience*, 11, 1-15. doi:10.3389/fnhum.2017.00629

- McCaskey, U., von Aster, M., O’Gorman Tuura, R. & Kucian, K. (2017). Adolescents with developmental dyscalculia do not have a generalized magnitude deficit – Processing of discrete and continuous magnitudes. *Frontiers in Human Neuroscience*, 11, 1-19. doi:10.3389/fnhum.2017.00102
- *McDonald, P. A. (2016). *Examining cognitive processing in children with an arithmetic disability*. (Master's thesis). Verfügbar unter ProQuest Dissertations & Theses Global. (UMI No. 1928901751)
- McDonald, P. A. & Berg, D. H. (2018). Identifying the nature of impairments in executive functioning and working memory of children with severe difficulties in arithmetic. *Child Neuropsychology*, 24(8), 1047-1062. doi:10.1080/09297049.2017.1377694
- McGill, R. J., Styck, K. M., Palomares, R. S. & Hass, M. R. (2016). Critical issues in specific learning disability identification: What we need to know about the PSW model. *Learning Disability Quarterly*, 39(3), 159-170. doi:10.1177/0731948715618504
- *Mejias, S., Mussolin, C., Rousselle, L., Gregoire, J. & Noël, M. P. (2012). Numerical and nonnumerical estimation in children with and without mathematical learning disabilities. *Child Neuropsychology*, 18(6), 550-575. doi:10.1080/09297049.2011.625355
- Menon, V. (2014). Arithmetic in the child and adult brain. In R. C. Kadosh & A. Dowker (Hrsg.), *The Oxford handbook of numerical cognition* Oxford, UK: Oxford University Press.
- Merdian, G., Merdian, F. & Schardt, K. (2015). *Bamberger Dyskalkuliediagnostik 1-4+ (R) (BADYS 1-4+)*. Bamberg: PaePsy.
- *Micallef, S. & Prior, M. (2004). Arithmetic learning difficulties in children. *Educational Psychology*, 24(2), 175-200. doi:10.1080/0144341032000160137
- Miciak, J., Fletcher, J. M., Stuebing, K. K., Vaughn, S. & Tolar, T. D. (2014). Patterns of cognitive strengths and weaknesses: Identification rates, agreement, and validity for learning disabilities identification. *School Psychology Quarterly*, 29(1), 21-37. doi:10.1037/spq0000037
- Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H. & Howerter, A. (2000). The unity and diversity of executive functions and their contributions to complex 'frontal lobe' tasks: A latent variable analysis. *Cognitive Psychology*, 41(1), 49-100. doi:10.1006/cogp.1999.0734
- Moeyaert, M., Ugille, M., Natasha Beretvas, S., Ferron, J., Bunuan, R. & van den Noortgate, W. (2017). Methods for dealing with multiple outcomes in meta-analysis: A comparison between averaging effect sizes, robust variance estimation and multilevel meta-analysis. *International Journal of Social Research Methodology*, 20(6), 559-572. doi:10.1080/13645579.2016.1252189
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G. & The PRISMA Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLOS Medicine*, 6(7), 1-6. doi:10.1371/journal.pmed.1000097
- *Moll, K., Göbel, S. M. & Snowling, M. J. (2015). Basic number processing in children with specific learning disorders: Comorbidity of reading and mathematics disorders. *Child Neuropsychology*, 21(3), 399-417. doi:10.1080/09297049.2014.899570

- Moll, K., Kunze, S., Neuhoff, N., Bruder, J. & Schulte-Körne, G. (2014). Specific learning disorder: Prevalence and gender differences. *Plos One*, 9(7), 1-8. doi:10.1371/journal.pone.0103537
- Möller, K., Fischer, U., Cress, U. & Nuerk, H.-C. (2012). Diagnostics and intervention in developmental dyscalculia: Current issues and novel perspectives. In Z. Breznitz, O. Rubinsten, V. J. Molfese & D. L. Molfese (Hrsg.), *Reading, writing, mathematics and the developing brain: Listening to many voices* (S. 233-275). New York, NY, US: Springer.
- Morgan, P. L., Farkas, G. & Wu, Q. (2009). Five-year growth trajectories of kindergarten children with learning difficulties in mathematics. *Journal of Learning Disabilities*, 42(4), 306-321. doi:10.1177/0022219408331037
- Moura, O., Simões, M. R. & Pereira, M. (2014). Executive functioning in children with developmental dyslexia. *The Clinical Neuropsychologist*, 28(S1), 1-22. doi:10.1080/13854046.2014.964326
- Moura, R., Lopes-Silva, J. B., Vieira, L. R., Paiva, G. M., de Almeida Prado, A. C., Wood, G. & Haase, V. G. (2015). From 'five' to 5 for 5 minutes: Arabic number transcoding as a short, specific, and sensitive screening tool for mathematics learning difficulties. *Archives of Clinical Neuropsychology*, 30(1), 88-98. doi:10.1093/arclin/acu071
- Murphy, M. M. (2009). A review of mathematical disabilities in children with fragile X syndrome. *Developmental Disabilities Research Reviews*, 15(1), 21-27. doi:10.1002/ddrr.49
- Murphy, M. M., Mazzocco, M. M. M., Hanich, L. B. & Early, M. C. (2007). Cognitive characteristics of children with mathematics learning disability (MLD) vary as a function of the cutoff criterion used to define MLD. *Journal of Learning Disabilities*, 40(5), 458-478. doi:10.1177/00222194070400050901
- Mussolin, C., Martin, R. & Schiltz, C. (2011). Relationships between number and space processing in adults with and without dyscalculia. *Acta Psychologica*, 138(1), 193-203. doi:10.1016/j.actpsy.2011.06.004
- *Mussolin, C., Mejias, S. & Noël, M. P. (2010). Symbolic and nonsymbolic number comparison in children with and without dyscalculia. *Cognition*, 115(1), 10-25. doi:10.1016/j.cognition.2009.10.006
- NCBI. (2016). MeSH. Verfügbar unter <https://www.ncbi.nlm.nih.gov/mesh> [19.02.2019].
- Niileksela, C. R. & Reynolds, M. R. (2014). Global, broad, or specific cognitive differences? Using a MIMIC model to examine differences in CHC abilities in children with learning disabilities. *Journal of Learning Disabilities*, 47(3), 224-236. doi:10.1177/0022219412453172
- Nolan, D. R., Hammeke, T. A. & Barkley, R. A. (1983). A comparison of the patterns of the neuropsychological performance in two groups of learning disabled children. *Journal of Clinical Child Psychology*, 12(1), 22-27. doi:10.1080/15374418309533108
- Olsson, L., Östergren, R. & Träff, U. (2016). Developmental dyscalculia: A deficit in the approximate number system or an access deficit? *Cognitive Development*, 39, 154-167. doi:10.1016/j.cogdev.2016.04.006
- Orraca-Castillo, M., Estevez-Perez, N. & Reigosa-Crespo, V. (2014). Neurocognitive profiles of learning disabled children with neurofibromatosis type 1. *Frontiers in Human Neuroscience*, 8, 1-9. doi:10.3389/fnhum.2014.00386

- Orwin, R. G. & Vevea, J. L. (2009). Evaluating coding decisions. In H. Cooper, L. V. Hedges & J. C. Valentine (Hrsg.), *The handbook of research synthesis and meta-analysis* (2. Aufl., S. 177-203). New York, NY, US: Russell Sage Foundation.
- Oxman, A. D. & Guyatt, G. H. (1992). A consumer's guide to subgroup analyses. *Annals of Internal Medicine*, *116*(1), 78-84. doi:10.7326/0003-4819-116-1-78
- *Peng, P., Congying, S., Beilei, L. & Sha, T. (2012). Phonological storage and executive function deficits in children with mathematics difficulties. *Journal of Experimental Child Psychology*, *112*(4), 452-466. doi:10.1016/j.jecp.2012.04.004
- Peng, P., Cuicui, W. & Jessica, N. (2018). Understanding the cognition related to mathematics difficulties: A meta-analysis on the cognitive deficit profiles and the bottleneck theory. *Review of Educational Research*, *88*(3), 434-476. doi:10.3102/0034654317753350
- Peng, P. & Fuchs, D. (2014). A meta-analysis of working memory deficits in children with learning difficulties: Is there a difference between verbal domain and numerical domain? *Journal of Learning Disabilities*, *49*(1), 3-20. doi:10.1177/0022219414521667
- Peng, P., Namkung, J., Barnes, M. & Sun, C. (2016). A meta-analysis of mathematics and working memory: Moderating effects of working memory domain, type of mathematics skill, and sample characteristics. *Journal of Educational Psychology*, *108*(4), 455-473. doi:10.1037/edu0000079
- Peng, P., Tao, S. & Li, B. L. (2013). The deficit profile of working memory, inhibition, and updating in Chinese children with reading difficulties. *Learning and Individual Differences*, *25*, 111-117. doi:10.1016/j.lindif.2013.01.012
- Peters, L. & de Smedt, B. (2018). Arithmetic in the developing brain: A review of brain imaging studies. *Developmental Cognitive Neuroscience*, *30*, 265-279. doi:10.1016/j.dcn.2017.05.002
- Peters, L., De Smedt, B. & Op de Beeck, H. (2015). The neural representation of Arabic digits in visual cortex. *Frontiers in Human Neuroscience*, *9*(517), 1-12. doi:10.3389/fnhum.2015.00517
- Pieters, S., Desoete, A., Roeyers, H., Vanderswalmen, R. & van Waelvelde, H. (2012). Behind mathematical learning disabilities: What about visual perception and motor skills? *Learning and Individual Differences*, *22*(4), 498-504. doi:10.1016/j.lindif.2012.03.014
- Pigott, T. D., Valentine, J. C., Polanin, J. R., Williams, R. T. & Canada, D. D. (2013). Outcome-reporting bias in education research. *Educational Researcher*, *42*(8), 424-432. doi:10.3102/0013189x13507104
- Poletti, M. (2016). WISC-IV intellectual profiles in Italian children with specific learning disorder and related impairments in reading, written expression, and mathematics. *Journal of Learning Disabilities*, *49*(3), 320-335. doi:10.1177/0022219414555416
- Pollack, C. & Ashby, N. C. (2018). Where arithmetic and phonology meet: The meta-analytic convergence of arithmetic and phonological processing in the brain. *Developmental Cognitive Neuroscience*, *30*, 251-264. doi:10.1016/j.dcn.2017.05.003
- *Powell, S. R., Fuchs, L. S., Fuchs, D., Cirino, P. T. & Fletcher, J. M. (2009). Do word-problem features differentially affect problem difficulty as a function

- of students' mathematics difficulty with and without reading difficulty? *Journal of Learning Disabilities*, 42(2), 99-110.
doi:10.1177/0022219408326211
- Proctor, B. (2012). Relationships between Cattell–Horn–Carroll (CHC) cognitive abilities and math achievement within a sample of college students with learning disabilities. *Journal of Learning Disabilities*, 45(3), 278-287.
doi:10.1177/0022219410392049
- R Core Team. (2017). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Verfügbar unter <https://www.R-project.org/>
- *Raddatz, J., Kuhn, J.-T., Holling, H., Moll, K. & Dobel, C. (2017). Comorbidity of arithmetic and reading disorder: Basic number processing and calculation in children with learning impairments. *Journal of Learning Disabilities*, 50(3), 298-308. doi:10.1177/0022219415620899
- Raghubar, K. P., Barnes, M. A. & Hecht, S. A. (2010). Working memory and mathematics: A review of developmental, individual difference, and cognitive approaches. *Learning and Individual Differences*, 20(2), 110-122. doi:10.1016/j.lindif.2009.10.005
- Raudenbush, S. W. (2009). Analyzing effect sizes: Random-effects models. In H. Cooper, L. V. Hedges & J. C. Valentine (Hrsg.), *The handbook of research synthesis and meta-analysis* (2. Aufl., S. 295-315). New York, NY, US: Russell Sage Foundation.
- Riley, R. D. (2009). Multivariate meta-analysis: The effect of ignoring within-study correlation. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 172(4), 789-811. doi:10.1111/j.1467-985X.2008.00593.x
- Ritchie, S. J. & Bates, T. C. (2013). Enduring links from childhood mathematics and reading achievement to adult socioeconomic status. *Psychological Science*, 24(7), 1301-1308. doi:10.1177/0956797612466268
- Roick, T., Göllitz, D. & Hasselhorn, M. (2018). *Deutscher Mathematiktest für dritte Klassen (DEMAT 3+)*. Göttingen: Hogrefe.
- *Rosenberg-Lee, M., Ashkenazi, S., Chen, T., Young, C. B., Geary, D. C. & Menon, V. (2014). Brain hyper-connectivity and operation-specific deficits during arithmetic problem solving in children with developmental dyscalculia. *Developmental Science*, 18(3), 351-372.
doi:10.1111/desc.12216
- Rosenberg-Lee, M., Ashkenazi, S., Chen, T., Young, C. B., Geary, D. C. & Menon, V. (2014). Brain hyper-connectivity and operation-specific deficits during arithmetic problem solving in children with developmental dyscalculia. *Developmental Science*, 18(3), 351-372.
doi:10.1111/desc.12216
- Rotzer, S., Loenneker, T., Kucian, K., Martin, E., Klaver, P. & von Aster, M. (2009). Dysfunctional neural network of spatial working memory contributes to developmental dyscalculia. *Neuropsychologia*, 47(13), 2859-2865. doi:10.1016/j.neuropsychologia.2009.06.009
- *Rousselle, L. & Noël, M. P. (2008). Mental arithmetic in children with mathematics learning disabilities: The adaptive use of approximate calculation in an addition verification task. *Journal of Learning Disabilities*, 41(6), 498-513. doi:10.1177/0022219408315638

- Rubinsten, O. & Henik, A. (2006). Double dissociation of functions in developmental dyslexia and dyscalculia. *Journal of Educational Psychology, 98*(4), 854-867. doi:10.1037/0022-0663.98.4.854
- Rücker, G., Carpenter, J. R. & Schwarzer, G. (2011). Detecting and adjusting for small-study effects in meta-analysis. *Biometrical Journal, 53*(2), 351-368. doi:10.1002/bimj.201000151
- Sala, G. & Gobet, F. (2017). Working memory training in typically developing children: A meta-analysis of the available evidence. *Developmental Psychology, 53*(4), 671-685. doi:10.1037/dev0000265
- Salthouse, T. A. (2000). Aging and measures of processing speed. *Biological Psychology, 54*(1), 35-54. doi:10.1016/S0301-0511(00)00052-1
- Sälzer, C. & Heine, J.-H. (2016). Students' skipping behavior on truancy items and (school) subjects and its relation to test performance in PISA 2012. *International Journal of Educational Development, 46*, 103-113. doi:10.1016/j.ijedudev.2015.10.009
- Sasanguie, D., Göbel, S. M., Moll, K., Smets, K. & Reynvoet, B. (2013). Approximate number sense, symbolic number processing, or number-space mappings: What underlies mathematics achievement? *Journal of Experimental Child Psychology, 114*(3), 418-431. doi:10.1016/j.jecp.2012.10.012
- Scammacca, N., Roberts, G. & Stuebing, K. K. (2014). Meta-analysis with complex research designs: Dealing with dependence from multiple measures and multiple group comparisons. *Review of Educational Research, 84*(3), 328-364. doi:10.3102/0034654313500826
- Schaupp, H., Lenart, F. & Holzer, N. (2010). *Eggenberger Rechentest 4+ (ERT 4+)*. Göttingen: Hogrefe.
- *Schleifer, P. & Landerl, K. (2011). Subitizing and counting in typical and atypical development. *Developmental Science, 14*(2), 280-291. doi:10.1111/j.1467-7687.2010.00976.x
- Schmidt, F. L., Le, H. & Oh, I.-S. (2009). Correcting for the distorting effects of study artifacts in meta-analysis. In H. Cooper, L. V. Hedges & J. C. Valentine (Hrsg.), *The handbook of research synthesis and meta-analysis* (2. Aufl., S. 317-333). New York, NY, US: Russell Sage Foundation.
- Schneider, M., Beeres, K., Coban, L., Merz, S., Susan Schmidt, S., Stricker, J. & de Smedt, B. (2017). Associations of non-symbolic and symbolic numerical magnitude processing with mathematical competence: A meta-analysis. *Developmental Science, 20*(3), 1-16. doi:10.1111/desc.12372
- Schneider, W. J. & McGrew, K. S. (2018). The Cattell-Horn-Carroll theory of cognitive abilities. In D. P. Flanagan & E. M. McDonough (Hrsg.), *Contemporary intellectual assessment: Theories, tests and issues* (4. Aufl., S. 73-163). New York, NY, US: The Guilford Press.
- Schuchardt, K., Fischbach, A., Balke-Melcher, C. & Mähler, C. (2015). Die Komorbidität von Lernschwierigkeiten mit ADHS-Symptomen im Grundschulalter. *Zeitschrift für Kinder- und Jugendpsychiatrie und Psychotherapie, 43*(3), 185-193. doi:10.1024/1422-4917/a000352
- *Schuchardt, K. & Mähler, C. (2010). Do subgroups of children with mathematical disabilities differ in their working memory, basic arithmetical knowledge, and numeric competences? *Zeitschrift für*

- Entwicklungspsychologie und Pädagogische Psychologie*, 42(4), 217-225. doi:10.1026/0049-8637/a000022
- *Schuchardt, K., Mähler, C. & Hasselhorn, M. (2008). Working memory deficits in children with specific learning disorders. *Journal of Learning Disabilities*, 41(6), 514-523. doi:10.1177/0022219408317856
- Schulte-Körne, G. (2016). Mental health problems in a school setting in children and adolescents. *Deutsches Ärzteblatt International*, 113(11), 183-190. doi:10.3238/arztebl.2016.0183
- Schwenk, C., Sasanguie, D., Kuhn, J.-T., Kempe, S., Doebler, P. & Holling, H. (2017). (Non-)symbolic magnitude processing in children with mathematical difficulties: A meta-analysis. *Research in Developmental Disabilities*, 64, 152-167. doi:10.1016/j.ridd.2017.03.003
- Shalev, R. S. (2007). Prevalence of developmental dyscalculia. In D. B. Berch & M. M. M. Mazzocco (Hrsg.), *Why is math so hard for some children? The nature and origins of mathematical learning difficulties and disabilities*. (S. 49-60). Baltimore, MD, US: Paul H Brookes Publishing.
- Shalev, R. S., Manor, O. & Gross-Tsur, V. (2005). Developmental dyscalculia: A prospective six-year follow-up. *Developmental Medicine & Child Neurology*, 47(2), 121-125. doi:10.1017/S0012162205000216
- Shin, M. & Bryant, D. P. (2015). A synthesis of mathematical and cognitive performances of students with mathematics learning disabilities. *Journal of Learning Disabilities*, 48(1), 96-112. doi:10.1177/0022219413508324
- *Silver, C. H., Elder, W. G. & DeBolt, A. J. (1999). Social skills self-appraisal of children with specific arithmetic disabilities. *Developmental Neuropsychology*, 16(1), 117-126. doi:10.1207/S15326942DN160107
- Skagerlund, K., Karlsson, T. & Träff, U. (2016). Magnitude processing in the brain: An fMRI study of time, space, and numerosity as a shared cortical system. *Frontiers in Human Neuroscience*, 10, 1-12. doi:10.3389/fnhum.2016.00500
- Skagerlund, K. & Träff, U. (2014). Development of magnitude processing in children with developmental dyscalculia: Space, time, and number. *Frontiers in Psychology*, 5, 1-15. doi:10.3389/fpsyg.2014.00675
- Skagerlund, K. & Träff, U. (2016). Number processing and heterogeneity of developmental dyscalculia: Subtypes with different cognitive profiles and deficits. *Journal of Learning Disabilities*, 49(1), 36-50. doi:10.1177/0022219414522707
- Stevens, J. P. (1984). Outliers and influential data points in regression-analysis. *Psychological Bulletin*, 95(2), 334-344. doi:10.1037/0033-2909.95.2.334
- Stock, P., Desoete, A. & Roeyers, H. (2010). Detecting children with arithmetic disabilities from kindergarten: Evidence from a 3-year longitudinal study on the role of preparatory arithmetic abilities. *Journal of Learning Disabilities*, 43(3), 250-268. doi:10.1177/0022219409345011
- Sutton, A. J. (2009). Publication bias. In H. Cooper, L. V. Hedges & J. C. Valentine (Hrsg.), *The handbook of research synthesis and meta-analysis* (2. Aufl., S. 435-452). New York, NY, US: Russell Sage Foundation.
- Swanson, H. L. (2004). Working memory and phonological processing as predictors of children's mathematical problem solving at different ages. *Memory & Cognition*, 32(4), 648-661. doi:10.3758/BF03195856

- *Swanson, H. L. (2012). Cognitive profile of adolescents with math disabilities: Are the profiles different from those with reading disabilities? *Child Neuropsychology*, 18(2), 125-143. doi:10.1080/09297049.2011.589377
- *Swanson, H. L. (2014). Does cognitive strategy training on word problems compensate for working memory capacity in children with math difficulties? *Journal of Educational Psychology*, 106(3), 831-848. doi:10.1037/a0035838
- Swanson, H. L. & Jerman, O. (2006). Math disabilities: A selective meta-analysis of the literature. *Review of Educational Research*, 76(2), 249-274. doi:10.3102/00346543076002249
- Swanson, H. L., Jerman, O. & Zheng, X. (2009). Math disabilities and reading disabilities: Can they be separated? *Journal of Psychoeducational Assessment*, 27(3), 175-196. doi:10.1177/0734282908330578
- *Swanson, H. L., Lussier, C. M. & Orosco, M. J. (2013a). Cognitive strategies, working memory, and growth in word problem solving in children with math difficulties. *Journal of Learning Disabilities*, 48(4), 339-358. doi:10.1177/0022219413498771
- *Swanson, H. L., Lussier, C. M. & Orosco, M. J. (2013b). Effects of cognitive strategy interventions and cognitive moderators on word problem solving in children at risk for problem solving difficulties. *Learning Disabilities Research & Practice*, 28(4), 170-183. doi:10.1111/ldrp.12019
- *Szanto, G. (1997). *Arithmetic disability of adults*. (Doctoral dissertation). Verfügbar unter ProQuest Dissertations & Theses Global. (UMI No. 304390867)
- *Szűcs, D., Devine, A., Soltesz, F., Nobes, A. & Gabriel, F. (2013). Developmental dyscalculia is related to visuo-spatial memory and inhibition impairment. *Cortex*, 49(10), 2674-2688. doi:10.1016/j.cortex.2013.06.007
- Tabachnick, B. G. & Fidell, L. S. (2007). *Using multivariate statistics* (5. Aufl.). Boston, MA: Allyn & Bacon/Pearson Education.
- Taub, G. E., Keith, T. Z., Floyd, R. G. & McGrew, K. S. (2008). Effects of general and broad cognitive abilities on mathematics achievement. *School Psychology Quarterly*, 23(2), 187-198. doi:10.1037/1045-3830.23.2.187
- Taylor, H. G., Espy, K. A. & Anderson, P. J. (2009). Mathematics deficiencies in children with very low birth weight or very preterm birth. *Developmental Disabilities Research Reviews*, 15(1), 52-59. doi:10.1002/ddrr.51
- Terrin, N., Schmid, C. H. & Lau, J. (2005). In an empirical evaluation of the funnel plot, researchers could not visually identify publication bias. *Journal of Clinical Epidemiology*, 58(9), 894-901. doi:10.1016/j.jclinepi.2005.01.006
- Tolar, T. D., Fuchs, L., Fletcher, J. M., Fuchs, D. & Hamlett, C. L. (2016). Cognitive profiles of mathematical problem solving learning disability for different definitions of disability. *Journal of Learning Disabilities*, 49(3), 240-256. doi:10.1177/0022219414538520
- *Torbeyns, J., Verschaffel, L. & Ghesquiere, P. (2004). Strategy development in children with mathematical disabilities: Insights from the choice/no-choice method and the chronological-age /ability-level-match design. *Journal of Learning Disabilities*, 37(2), 119-131. doi:10.1177/00222194040370020301

- Träff, U., Olsson, L., Östergren, R. & Skagerlund, K. (2017). Heterogeneity of developmental dyscalculia: Cases with different deficit profiles. *Frontiers in Psychology, 7*, 1-15. doi:10.3389/fpsyg.2016.02000
- Valentine, J. C., Pigott, T. D. & Rothstein, H. R. (2010). How many studies do you need? A primer on statistical power for meta-analysis. *Journal of Educational and Behavioral Statistics, 35*(2), 215-247. doi:10.3102/1076998609346961
- *van Daal, V., van der Leij, A. & Adèr, H. (2013). Specificity and overlap in skills underpinning reading and arithmetical fluency. *Reading and Writing, 26*(6), 1009-1030. doi:10.1007/s11145-012-9404-5
- van den Noortgate, W., Lopez-Lopez, J. A., Marin-Martinez, F. & Sanchez-Meca, J. (2013). Three-level meta-analysis of dependent effect sizes. *Behavior Research Methods, 45*(2), 576-594. doi:10.3758/s13428-012-0261-6
- van den Noortgate, W., Lopez-Lopez, J. A., Marin-Martinez, F. & Sanchez-Meca, J. (2014). Meta-analysis of multiple outcomes: A multilevel approach. *Behavior Research Methods, 47*(4), 1274-1294. doi:10.3758/s13428-014-0527-2
- *van der Sluis, S., de Jong, P. F. & van der Leij, A. (2004). Inhibition and shifting in children with learning deficits in arithmetic and reading. *Journal of Experimental Child Psychology, 87*(3), 239-266. doi:10.1016/j.jecp.2003.12.002
- *van der Sluis, S., de Jong, P. F. & van der Leij, A. (2005). Working memory in Dutch children with reading- and arithmetic-related LD. *Journal of Learning Disabilities, 38*(3), 207-221. doi:10.1177/00222194050380030301
- van Houwelingen, H. C., Arends, L. R. & Stijnen, T. (2002). Advanced methods in meta-analysis: Multivariate approach and meta-regression. *Statistics in Medicine, 21*(4), 589-624. doi:10.1002/sim.1040
- van Iterson, L., de Jong, P. F. & Zijlstra, B. J. H. (2015). Pediatric epilepsy and comorbid reading disorders, math disorders, or autism spectrum disorders: Impact of epilepsy on cognitive patterns. *Epilepsy & Behavior, 44*, 159-168. doi:10.1016/j.yebeh.2015.02.007
- van Rooijen, M., Verhoeven, L. & Steenbergen, B. (2015). From numeracy to arithmetic: Precursors of arithmetic performance in children with cerebral palsy from 6 till 8 years of age. *Research in Developmental Disabilities, 45-46*, 49-57. doi:10.1016/j.ridd.2015.07.001
- Vanbinst, K. & de Smedt, B. (2016). Individual differences in children's mathematics achievement: The roles of symbolic numerical magnitude processing and domain-general cognitive functions. *Progress In Brain Research, 227*, 105-130. doi:10.1016/bs.pbr.2016.04.001
- Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software, 36*(3), 1-48. doi:10.18637/jss.v036.i03
- Viechtbauer, W. & Cheung, M. W. L. (2010). Outlier and influence diagnostics for meta-analysis. *Research Synthesis Methods, 1*(2), 112-125. doi:10.1002/jrsm.11
- von Aster, M., Weinhold-Zulauf, M. & Horn, R. (2006). *Neuropsychologische Testbatterie für Zahlenverarbeitung und Rechnen bei Kindern (ZAREKI-R)*. Frankfurt am Main: Pearson Assessment & Information.

- von Elm, E., Poglia, G., Walder, B. & Tramer, M. R. (2004). Different patterns of duplicate publication - An analysis of articles used in systematic reviews. *JAMA*, 291(8), 974-980. doi:10.1001/jama.291.8.974
- Voyer, D. & Voyer, S. D. (2014). Gender differences in scholastic achievement: A meta-analysis. *Psychological Bulletin*, 140(4), 1174-1204. doi:10.1037/a0036620
- Wagner, R. K. & Torgesen, J. K. (1987). The nature of phonological processing and its causal role in the acquisition of reading skills. *Psychological Bulletin*, 101(2), 192-212. doi:10.1037/0033-2909.101.2.192
- Walker, E., Hernandez, A. V. & Kattan, M. W. (2008). Meta-analysis: Its strengths and limitations. *Cleveland Clinic Journal of Medicine*, 75(6), 431-439. doi:10.3949/ccjm.75.6.431
- *Wang, L.-C., Tasi, H.-J. & Yang, H.-M. (2012). Cognitive inhibition in students with and without dyslexia and dyscalculia. *Research in Developmental Disabilities*, 33(5), 1453-1461. doi:10.1016/j.ridd.2012.03.019
- Watson, S. M. R. & Gable, R. A. (2013). Unraveling the complex nature of mathematics learning disability: Implications for research and practice. *Learning Disability Quarterly*, 36(3), 178-187. doi:10.1177/0731948712461489
- *Willburger, E., Fussenegger, B., Moll, K., Wood, G. & Landerl, K. (2008). Naming speed in dyslexia and dyscalculia. *Learning and Individual Differences*, 18(2), 224-236. doi:10.1016/j.lindif.2008.01.003
- Willcutt, E. G., Petrill, S. A., Wu, S., Boada, R., DeFries, J. C., Olson, R. K. & Pennington, B. F. (2013). Comorbidity between reading disability and math disability: Concurrent psychopathology, functional impairment, and neuropsychological functioning. *Journal of Learning Disabilities*, 46(6), 500-516. doi:10.1177/0022219413477476
- *Wilson, A. J., Andrewes, S. G., Struthers, H., Rowe, V. M., Bogdanovic, R. & Waldie, K. E. (2015). Dyscalculia and dyslexia in adults: Cognitive bases of comorbidity. *Learning and Individual Differences*, 37, 118-132. doi:10.1016/j.lindif.2014.11.017
- Wilson, D. B. (2009). Systematic coding. In H. Cooper, L. V. Hedges & J. C. Valentine (Hrsg.), *The handbook of research synthesis and meta-analysis* (2. Aufl., S. 159-176). New York, NY, US: Russell Sage Foundation.
- *Winegar, K. L. (2013). *Inhibition performance in children with math disabilities*. (Doctoral dissertation). Verfügbar unter ProQuest Dissertations & Theses Global. (UMI No. 1430900289)
- Wood, W. & Eagly, A. H. (2009). Advantages of certainty and uncertainty. In H. Cooper, L. V. Hedges & J. C. Valentine (Hrsg.), *The handbook of research synthesis and meta-analysis* (2. Aufl., S. 455-472). New York, NY, US: Russell Sage Foundation.
- World Health Organization. (1978). *International classification of diseases (9th Revision)*. Geneva, CH: Author.
- World Health Organization. (1992). *International statistical classification of diseases and related health problems (10th Revision)*. Geneva, CH: Author.
- World Health Organization. (2018). *International statistical classification of diseases and related health problems (11th Revision)*. Verfügbar unter <https://icd.who.int/browse11/l-m/en> [21.02.2019].

- *Yaghoubzadeh, Z. (1995). *Arithmetic performance in learning disabled and normally achieving children: The relationship between error types and cognitive processes*. (Master's thesis). Verfügbar unter ProQuest Dissertations & Theses Global. (UMI No. 304270038)
- Zhang, X., Rasanen, P., Koponen, T., Aunola, K., Lerkkanen, M. K. & Nurmi, J. E. (2017). Knowing, applying, and reasoning about arithmetic: Roles of domain-general and numerical skills in multiple domains of arithmetic learning. *Developmental Psychology*, 53(12), 2304-2318.
doi:10.1037/dev0000432

Abbildungsverzeichnis

<i>Abbildung 1.</i> Neuronales Netzwerk beim Rechnen.	8
<i>Abbildung 2.</i> Schematische Darstellung des Kodiervorgehens.	38
<i>Abbildung 3.</i> Überblick über das Auswertungsprocedere.	50
<i>Abbildung 4.</i> Flussdiagramm zum Suchverlauf und Anwendung der Einschlusskriterien.	60
<i>Abbildung 5.</i> Ermittelte Effektstärken für domänenübergreifende Fähigkeiten.	74
<i>Abbildung 6.</i> Ermittelte Effektstärken für domänenübergreifende Fähigkeiten.	77

Tabellenverzeichnis

Tabelle 1 <i>Verwendete Suchparameter in Deutsch und Englisch</i>	34
Tabelle 2 <i>Kodierschema zu deskriptiven Studienangaben</i>	35
Tabelle 3 <i>Kodierschema zu Outcomes der Studienergebnisse und deren Operationalisierung</i>	40
Tabelle 4 <i>Deskriptive Studienangaben zu Studien- und Stichprobenmerkmalen</i>	62
Tabelle 5 <i>Deskriptive Studienangaben zu Diagnosemerkmalen</i>	66
Tabelle 6 <i>Ergebnisse der multivariaten Mehrebenen-Metaanalysen aller Datensätze</i>	70
Tabelle 7 <i>Modellfit mit fixierter Ebene 2 bzw. 3 und Anteile der Varianz je Ebene</i>	78
Tabelle 8 <i>Studienübergreifende Biasprüfung (II)</i>	80
Tabelle 9 <i>Kognitives Profil der Rechenstörung basierend auf dem RSE</i>	86
Tabelle 10 <i>Inhalte eines fiktiven Rechentests auf Basis des kognitiven Profils der Rechenstörung</i>	108

Danksagung

Für die durchweg hervorragende Betreuung möchte ich mich an dieser Stelle herzlichst bei meinem Doktorvater, Herrn Prof. Gerd Schulte-Körne, bedanken. Insbesondere möchte ich mich an dieser Stelle auch für seine allgemeine Unterstützung in vielerlei berufsbezogenen Fragen bedanken, die dieses Promotionsvorhaben begleitet haben.

Ein ebenso großer Dank gebührt meiner Familie für ihre große Unterstützung in den letzten Jahren und natürlich meiner kleinen Tochter für ihre emotionale Anteilnahme mitsamt den (wenigen) Nächten, in denen sie mich schlafen ließ.

Eidesstattliche Versicherung

Haberstroh, Stefan

Ich erkläre hiermit an Eides statt,

dass die vorliegende Dissertation mit dem Titel

Das kognitive Profil der Rechenstörung

selbständig verfasst, mich außer der angegebenen keiner weiteren Hilfsmittel bedient und alle Erkenntnisse, die aus dem Schrifttum ganz oder annähernd übernommen sind, als solche kenntlich gemacht und nach ihrer Herkunft unter Bezeichnung der Fundstelle einzeln nachgewiesen habe.

Ich erkläre des Weiteren, dass die hier vorgelegte Dissertation nicht in gleicher oder in ähnlicher Form bei einer anderen Stelle zur Erlangung eines akademischen Grades eingereicht wurde.

München, 19.03.2020

Ort, Datum

Stefan Haberstroh

Unterschrift Doktorand