Taina Gabriel, Katja Köhlken, Ante Pavic, Claudia Mähler, Jürgen Wilbert & Moritz Börnert-Ringleb

Dynamisches Testen von Rechenfertigkeiten im Grundschulalter – zentrale Befunde des Verbundprojekts DYNAMIK

Zusammenfassung
Dynamisches Testen wird zunehmend als Alternative zur Statusdiagnostik diskutiert. Bisher liegt jedoch kein deutschsprachiges dynamisches Testverfahren zur Erfassung der Rechenfähigkeiten von Grundschüler:innen vor, sodass basierend auf dem Graduated-Prompt-Ansatz der DTR-ASM im Projekt DYNAMIK[1] entwickelt wurde. 37 Lehrkräfte wurden in drei verschiedenen Gruppen (dynamischer Test, Statustest, Kontrollgruppe) fortgebildet und führten eine entsprechende Diagnostik mit rechenschwachen Drittklässler:innen durch. Für jedes Kind wurde basierend auf den Ergebnissen ein Förderplan ($N = 99$) verfasst, welcher bezüglich seiner Qualität und seiner Inhalte untersucht wurde, um die Nützlichkeit des DTR-ASM für die Förderplanung zu untersuchen. Es zeigten sich keine Unterschiede zwischen den Gruppen bezüglich der Qualität der Förderpläne. Inhaltlich ließen sich Unterschiede finden, diese zeigten jedoch kein klares Muster. Die Implikationen der Ergebnisse werden diskutiert.
Schlüsselworte: Dynamisches Testen, Graduated Prompts, Rechenfähigkeiten, Förderpläne

Abstract
Dynamic testing is being increasingly discussed as an alternative to traditional tests. Up to now, there are no dynamic tests of arithmetic skills in Germany, so that the DTR-ASM was developed based on the graduated prompt approach in the project DYNAMIK. 37 teachers were educated in three different conditions (dynamic test, traditional test, control group) and asked to use the according approach with 3rd graders struggling in math. Based on the diagnostic results, teachers wrote educational plans for each child ($N = 99$) which were then analysed according to their quality and their content in order to investigate the usefulness of the DTR-ASM for educational planning. No differences between groups were found regarding the quality of the plans. In terms of content, multiple differences appeared but they did not follow a clear pattern. Implications of these results are being discussed.
Keywords: Dynamic testing, graduated prompts, arithmetic skills, educational plans

1. Einführung

Individuelle Förderung kann zur Teilhabe aller Kinder und Jugendlichen an Bildungsprozessen beitragen, setzt jedoch eine vorherige Beschreibung und Erfassung individueller Lernvoraussetzungen und Lernpotentiale voraus. In diesem Sinne stellen diagnostisches Handeln sowie geeignete diagnostische Ansätze und Methoden wesentliche Gelingensbedingungen individueller Förderung dar (Wilbert & Börnert,

1 Dieses Projekt wurde vom Bundesministerium für Bildung und Forschung (BMBF) unter den Förderkennzeichen 01NV2120A, 01NV2120B und 01NV2120C gefördert. Die Verantwortung für den Inhalt dieser Veröffentlichung liegt bei den Autor:innen.

2015). Mit Bezug zu pädagogischen Entscheidungsprozessen (wie z. B. der Förderplanung) wurden in den letzten Jahren insbesondere Ansätze des *data-based decision making* prominent diskutiert (Mandinach, 2012; Schildkamp, Lai & Earl, 2013). Diese argumentieren, dass es im pädagogischen Kontext nicht hinreichend sei, sich bei notwendigen Entscheidungen auf Überzeugungen und Bauchgefühle („gut feelings"; Mandinach, 2012, S.71) zu verlassen. Im Kontext schulischer Förderdiagnostik werden hierzu neben informellen diagnostischen Ansätzen insbesondere auch standardisierte Testverfahren eingesetzt (Bundschuh & Winkler, 2014), da diese eine objektive, reliable und valide Einschätzung der Fähigkeiten von Kindern und Jugendlichen zu ermöglichen scheinen. Zudem basieren sie häufig auf theoretischen Überlegungen zu den erfassten Konstrukten. In den letzten Jahren wurden jedoch Zweifel an der Eignung testdiagnostischer Zugänge für förderdiagnostische Entscheidungen geäußert. Diese Kritik lässt sich im Wesentlichen auf zwei Kernaussagen reduzieren: Zum einen wird argumentiert, dass testdiagnostische Informationen häufig keine Aussagen über die Ursachen der jeweiligen Testergebnisse und dementsprechend auch keine direkte Ableitung von notwendigen Impulsen und möglicher Unterstützung zur Entfaltung der abgebildeten Kompetenz ermöglichen (Bosma & Resing, 2012; Pameijer, 2006; Schlee, 2008). Zum anderen wird kritisiert, dass testdiagnostische Verfahren nicht geeignet seien, die tatsächlichen Kompetenzen der Kinder valide zu erfassen und Prognosen über Lernentwicklungen zu treffen (Hessels, 1997; Resing, Tunteler, De Jong & Bosma, 2009; Resing, Stevenson & Bosma, 2012; Tiekstra, Hessels & Minnaert, 2009). Diese Beobachtungen wurden für Kinder mit einer geistigen Beeinträchtigung (Resing et al., 2012; Tiekstra et al., 2009) und Lernende aus kulturellen Minderheiten (Hessels, 1997; Resing et al., 2009) beschrieben. Als Ursache für diese eingeschränkte Validität wird diskutiert, dass die Testverfahren unzureichend berücksichtigen, dass manche Kinder nicht über das notwendige Wissen über die jeweiligen Aufgabenformate und die metakognitiven Kompetenzen verfügen, um eine adäquate kognitive Repräsentation der Aufgabenstellung herzustellen (Tiekstra et al., 2009). Aus Perspektive einer förderorientierten Diagnostik erscheint insbesondere der Aspekt der fehlenden praktischen Relevanz bzw. der scheiternden Übertragung erfasster Testinformationen auf die konkrete Gestaltung der Förderpraxis problematisch. Es besteht daher ein Bedarf an alternativen diagnostischen Ansätzen, welche eine höhere Passung für entsprechende Zwecke aufweisen.

2. Dynamisches Testen im Rahmen der Förderplanung

Mit Blick auf die skizzierte Kritik und den resultierenden Bedarf stellt das Konzept des dynamischen Testens (DT; Guthke, 1992; Sternberg & Grigorenko, 2002) eine vielversprechende Alternative zu traditionellen Testverfahren dar. Dieses hat zum Ziel, die Lernpotentiale und benötigte Unterstützung zur Entfaltung dieser Potentiale von Kindern zu erfassen (Börnert & Wilbert, 2016; Guthke, Beckmann & Wiedl, 2003). Zu diesem Zweck wird der diagnostische Prozess erweitert und eine Intervention als wesentlicher Teil des diagnostischen Prozesses verstanden (Guthke et al.,

2003; Sternberg & Grigorenko, 2002). Dieses Vorgehen widerspricht dem grundlegenden Verständnis traditionellen Testens, bei welchem das Testergebnis nicht von der testleitenden Person beeinflusst werden sollte, da ansonsten das Gütekriterium der Objektivität gefährdet ist.

Im dynamischen Testen erhalten Kinder hingegen bei auftretenden Schwierigkeiten im Testprozess entweder eine aufgabenspezifische Förderung (Guthke et al., 2003) oder eine unmittelbare Unterstützung in Form eines Hinweises *(Prompts)* (Resing, 2013). Die Verbindung von Förderung und Testprozess kann dabei entweder strukturell getrennt (z. B. im Rahmen von Prä-Test, Training und Post-Test) oder eng verwoben (z. B. im Rahmen gestufter Hinweise) erfolgen. Erstere Verfahren nehmen dabei häufig mehrere Sitzungen in Anspruch und sind mit einem entsprechenden zeitlichen und ökonomischen Aufwand verbunden. Zudem erscheinen sie insbesondere für den Zweck der Identifikation von Lernpotentialen und der Feststellung des Ausmaßes an Veränderung relevant.

Für den Bereich der Förderplanung scheinen Varianten des dynamischen Testens geeignet, bei welchen Förderung und Unterstützung unmittelbar im Testprozess erfolgen. Der sogenannte *Graduated-Prompt-Ansatz* (Resing & Elliott, 2011) stellt hierbei eine prominente Variante dar. Bei diesem erhalten testleitende Personen ein Unterstützungsprotokoll, welches a priori definierte Hinweise beinhaltet. Die *Prompts* können dabei zur Aufgabenlösung notwendigen *kognitiven*, *metakognitiven* oder *strategischen* Prozessen entsprechen (Börnert-Ringleb & Wilbert, 2019). Sobald Kinder eine Aufgabe nicht lösen können bzw. falsch lösen, werden sie durch einen *Prompt* unterstützt. Das Ausmaß, in dem Kinder von den *Prompts* profitieren, kann das *Lernpotential* abbilden. Die Validität der erfassten Lernpotentiale konnte sowohl in zahlreichen querschnittlichen (Börnert-Ringleb & Wilbert, 2019; Fuchs, Compton, Fuchs, Bouton & Caffrey, 2011; Hessels, 1997; Stevenson, Heiser & Resing, 2016) als auch longitudinalen Untersuchungen (Caffrey, Fuchs & Fuchs, 2008; Cho, Compton, Fuchs, Fuchs & Boutin, 2014; Fuchs et al., 2011) nachgewiesen werden. Vielversprechend erscheint zudem die Möglichkeit der Ableitung von Förderempfehlungen aus dem dynamischen Testprozess, da über die Verarbeitung der *Prompts* durch die Kinder bereits konkrete Hinweise auf deren Entwicklungsstand sichtbar werden und die „Zone der nächsten Entwicklung" (Vygotskij & Cole, 1981) adressierbar wird. So wird argumentiert, dass sich aus der Responsivität von Individuen gegenüber einzelnen *Prompts* Schlüsse über relevante und zu adressierende Prozesse und Facetten ableiten lassen. Studien, welche die Eignung entsprechender Varianten des DT für diesen Zweck fokussieren, liegen bisher kaum vor. Bosma und Resing (2012) zeigten, dass sich interindividuelle Unterschiede in der Unterstützungsstruktur im dynamischen Testen beschreiben lassen und Lehrkräfte diese erfassten Testinformationen als wertvoll einschätzen. Überdies bevorzugten Lehrkräfte dynamische Testinformationen im Vergleich zu traditionellen oder verlaufsdiagnostischen Informationen für die Förderplanung (Bosma, Hessels & Resing, 2012). Die bestehenden Befunde beziehen sich insbesondere auf domänenübergreifende kognitive Prozesse. Mit Bezug auf schulisches Handeln erscheint dies insofern distal, als dass Unterricht in der Regel domänenspezifische Prozesse (z. B. die Rechenfähigkeit) adressiert. Entspre-

chende Ergebnisse sind aus diesem Grund nicht unmittelbar auf den Unterrichtsgegenstand übertragbar. Neben der Bestimmung eines domänenübergreifenden Lernpotentials erscheint daher die Anwendung des Konzeptes des dynamischen Testens in domänenspezifischen Kompetenzbereichen wie der Rechen- (Fuchs et al., 2008; Seethaler, Fuchs, Fuchs & Compton, 2012) oder Lesefähigkeit (Dörfler, Golke & Artelt, 2009; Sittner Bridges & Catts, 2011) sinnvoll. Zwar liegen bereits erste Vorschläge vor, diese adressieren jedoch insbesondere das Lesen und wurden in der Regel nicht hinsichtlich ihrer Eignung für förderdiagnostische Entscheidungen untersucht. Die vorliegenden Befunde in diesen Bereichen zeigen jedoch Vorteile dynamisch erfasster Lernpotentiale gegenüber Ergebnissen traditioneller Verfahren mit Bezug auf die Vorhersage der Lernentwicklung und des Lernerfolgs (Börnert-Ringleb & Wilbert, 2020).

2.1 Anwendungen des dynamischen Testens zur Erfassung der Rechenfähigkeit

Diagnostisches Handeln kann das frühe Erkennen von sowie eine zeitnahe Intervention bei sich entwickelnden Rechenschwierigkeiten unterstützen. Dies ist aus verschiedenen Gründen von hoher Relevanz (Balt, Fritz & Ehlert, 2020): Bereits zu Schuleintritt unterscheiden sich die mathematischen Fähigkeiten von Kindern (Bodovski & Farkas, 2007). Dies ist insbesondere relevant, da Rechenfertigkeiten mit der langfristigen Entwicklung von Kindern in Verbindung stehen (Ritchie & Bates, 2013). Lernerfolg in Mathematik konnte in verschiedenen Studien durch das frühe mathematische Wissen der Kinder vorhergesagt werden (Claessens & Engel, 2013; Mercader, Miranda, Presentación, Siegenthaler & Rosel, 2018). Schwierigkeiten im Rechnen führen dabei aber nicht nur zu einem negativeren Schulerfolg, sondern gehen darüber hinaus mit einer schlechteren psychosozialen Entwicklung einher (Parsons & Bynner, 2005).

In der schulischen Diagnostik des Rechnens werden insbesondere Verfahren eingesetzt, welche primär Rechenoperationen bzw. numerische Vorläuferfähigkeiten erfassen. Rechenfertigkeiten umfassen darüber hinaus jedoch unterschiedliche Komponenten, wie z. B. Gedächtnisstrategien, exekutive Prozesse und spezifisches mathematisches Wissen, die sich inter- und intraindividuell unterscheiden und zur Genese von Rechenschwierigkeiten beitragen können (Dowker, 2008; Kaufmann et al., 2013; von Aster, Schweiter & Weinhold Zulauf, 2007). Auch Defizite im Arbeitsgedächtnis haben sich zur Erklärung von Dyskalkulie als relevant erwiesen (Maehler & Schuchardt, 2016; Schuchardt & Mähler, 2010). Zur Diagnostik und Förderung bei Rechenschwierigkeiten bedarf es daher neben der Fokussierung konkreter Rechenoperationen und numerischer Kompetenzen insbesondere der Berücksichtigung übergeordneter kognitiver Prozesse wie exekutiver und Arbeitsgedächtnisfunktionen. Dynamisches Testen kann an dieser Stelle durch Beobachtung der Aufgabenbearbeitung und gezielte Unterstützung der Kinder mittels gestufter *Prompts* Einblick in für die Förderung relevante kognitive Prozesse geben

und daraus Schlussfolgerungen über zu berücksichtigende Komponenten ermöglichen.

3. Ziele und Aufbau des Forschungsprojekts DYNAMIK

Mit Bezug auf den dargestellten möglichen Nutzen des dynamischen Testens für die förderorientierte Diagnostik fokussierte das Forschungsprojekt DYNAMIK *(Dynamisches Testen als Perspektive für förderdiagnostische Entscheidungen in der Schule)* den Nutzen des dynamischen Testens für die Förderplanung im Bereich des Rechnens. Das Projekt umfasste eine Laufzeit von drei Jahren und wurde an den Standorten Hannover, Hildesheim und Potsdam durchgeführt. Das komplexe experimentelle Projektdesign gliederte sich in drei zentrale Teilschritte (Abbildung 1).

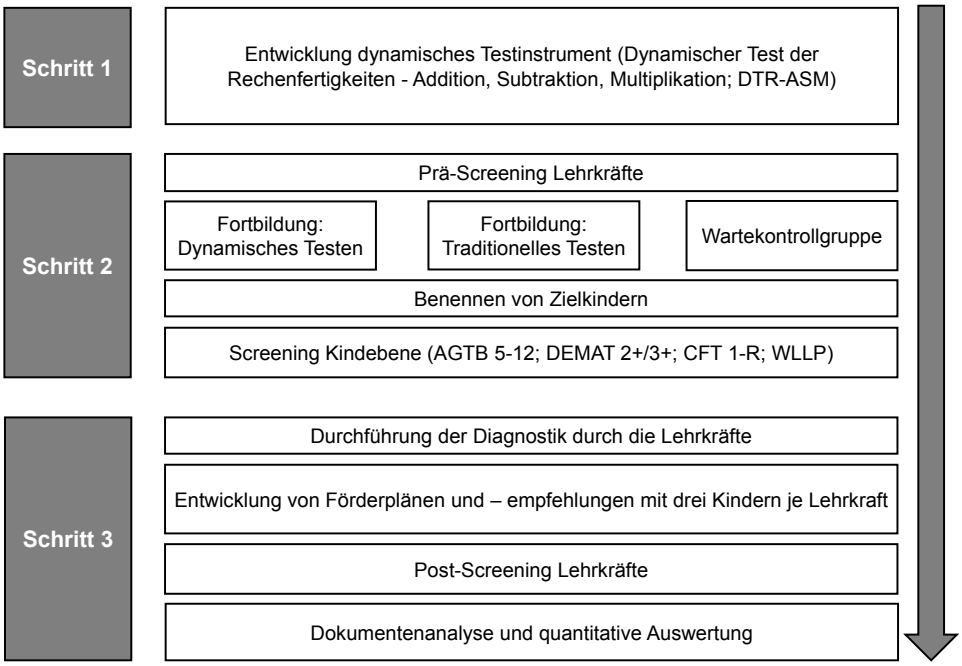

Abbildung 1: Design und Ablauf des Projekts DYNAMIK

In einem *ersten Schritt* wurde das Konzept des dynamischen Testens auf den Bereich der Erfassung des Rechnens von Grundschulkindern übertragen (Kapitel 4.1). Aufbauend wurden Fortbildungselemente konzipiert, welche für den folgenden Schritt des Projekts benötigt wurden. Im *zweiten Schritt* wurden die an der Studie teilnehmenden Lehrkräfte in verschiedenen Fortbildungsbedingungen qualifiziert. Hierbei nahmen alle Lehrkräfte zunächst an einer Fortbildung zur Entwicklung mathematischer Kompetenzen teil. Anschließend wurden die Lehrkräfte in der Anwendung von zwei unterschiedlichen testdiagnostischen Zugängen geschult. Ein Teil der Lehr-

kräfte (DT-G) nahm dabei an einer Fortbildung zur Anwendung des im Projekt entwickelten dynamischen Rechentests (Dynamischer Test der Rechenfertigkeiten – Addition, Subtraktion, Multiplikation, DTR-ASM) teil. Die zweite Gruppe (TT-G) lernte, einen bestehenden statusdiagnostischen Test (HRT 1–4; Haffner, Baro, Parzer & Resch, 2005) anzuwenden. Lehrkräfte in der Kontrollgruppe (KG) erhielten zunächst keine weitere Fortbildung, hatten aber nach Abschluss des Projekts die Möglichkeit, an einer Fortbildung zur Vorstellung des dynamischen Tests teilzunehmen. Nach der Teilnahme an den Fortbildungen wurden die Lehrkräfte gebeten, die erlernte Testdiagnostik mit jeweils drei Schüler:innen mit Rechenschwierigkeiten durchzuführen. Lehrkräfte der Kontrollgruppe führten keine zusätzliche Diagnostik durch, sondern sollten ihre üblichen diagnostischen Routinen (z. B. Alltagsbeobachtungen im Unterricht) durchführen.

In einem *dritten Schritt* verfassten alle Lehrkräfte abschließend Förderpläne für die ausgewählten und identifizierten Schüler:innen. Dabei wurden sie zunächst gebeten, den Lernstand des Kindes im Hinblick auf seine mathematischen Kompetenzen zu beschreiben. Anschließend sollten sie die Lernziele für das laufende Schuljahr (i. d. R. das zweite Schulhalbjahr der dritten Klasse) definieren und eine Förderung zur Erreichung der formulierten Lernziele planen. Die entstandenen Förderpläne wurden anschließend im Hinblick auf die Inhalte sowie die Qualität (Kohärenz) verglichen. Bei der Betrachtung der Förderpläne sollten folgende zentrale Forschungsfragen beantwortet werden:

1. Unterscheiden sich die Förderpläne hinsichtlich der Qualität zwischen den Untersuchungsbedingungen (DT-G, TT-G, KG)?
2. Unterscheiden sich die Förderpläne hinsichtlich der Inhalte (genannte Kompetenzen/Prozesse) zwischen den Untersuchungsbedingungen (DT-G, TT-G, KG)?

4. Zentrale Befunde und Produkte des Verbundprojekts

Im Folgenden sollen nun zentrale Befunde und Produkte des Projekts skizziert werden. Die Darstellung der Ergebnisse erfolgt dabei dem im Projekt vorgesehenen Ablauf.

4.1 Entwicklung des Dynamischen Tests der Rechenfertigkeiten (DTR-ASM)

Vor dem Hintergrund, dass mit Bezug auf das dynamische Testen noch kein deutschsprachiges Verfahren zur Erfassung der Rechenfertigkeit vorlag, war eines der zentralen Ziele des Projekts, einen entsprechenden Test zu entwickeln. Dieser sollte dabei den Entwicklungsstand in Bezug auf die arithmetischen Fertigkeiten von als rechenschwach wahrgenommenen Drittklässler:innen abbilden. Im Gegensatz zu statusdiagnostischen Verfahren, bei denen keinerlei Hilfestellung bei der Aufgabenlösung vorgesehen ist, beruht der entwickelte Test auf dem beschriebenen *Graduated-*

Prompt-Ansatz. Dabei werden dem Kind, sobald es an der Lösung einer Aufgabe scheitert, so lange Hinweise gegeben, bis die Aufgabe gelöst werden kann. Die Hinweise werden hierbei immer spezifischer, und häufig beinhaltet der finale Hinweis die Modellierung der Aufgabenlösung. Die Anzahl und Art der benötigten Hinweise können dann dazu genutzt werden, Auskunft über das Lernpotential der Kinder, die „Zone der nächsten Entwicklung" (Vygotskij & Cole, 1981) sowie die benötigten Hilfestellungen zu bekommen. Im Rahmen des Projekts sollten insbesondere *Prompts* berücksichtigt werden, die die große Bedeutung des Arbeitsgedächtnisses (Maehler & Schuchardt, 2016; Mann, Fischer & Nuerk, 2013) sowie emotionaler Prozesse (Roick, Gölitz & Hasselhorn, 2013) für die Mathematikleistung adressieren. Die *Prompts* sollten darüber hinaus motivierend sein und Erfolgserlebnisse ermöglichen (Schukajlow, Rakoczy & Pekrun, 2017). Um eine möglichst einfache Durchführung zu gewährleisten, sollten die Hinweise klar und einheitlich sein. Außerdem sollte das dynamische Testverfahren den zeitlichen Umfang eines etablierten standardisierten Testverfahrens nicht überschreiten.

4.1.1 Überblick über das Testinstrument DTR-ASM

Der konzipierte dynamische Test basiert auf den Modellen zur Entwicklung mathematischer Fertigkeiten nach Dehaene (1992) sowie Krajewski und Ennemoser (2013). Darauf aufbauend wurde ein hierarchisches *Promptschema* entwickelt, welches in den Test integriert ist und im Folgenden anhand einer Beispielaufgabe erläutert wird. Das *Promptschema* berücksichtigt hierbei verschiedene Prozesse wie das Arbeitsgedächtnis, die Metakognition sowie Rechen- und Zählstrategien. Zudem dürfen jederzeit motivational-emotionale *Prompts* zur Unterstützung, zum positiven Feedback, zur Stärkung des Selbstvertrauens und zur Beruhigung eingesetzt werden (z. B. „Das hast du gut gemacht!", „Ich helfe dir/unterstütze dich bei den Aufgaben!", „Es ist nicht schlimm, wenn du die Aufgabe nicht direkt löst.").

Auf Grundlage der Lehrpläne in den teilnehmenden Bundesländern (Niedersächsisches Kultusministerium, 2017; Bildungsserver Berlin-Brandenburg, 2015) sowie vergleichbaren Rechentests (z. B. DEMAT 3+; Roick, Gölitz & Hasselhorn, 2004) beinhaltet der Test jeweils sechs Additions-, Subtraktions- und Multiplikationsaufgaben zur Operationalisierung der arithmetischen Fertigkeiten. Die Division als vierte Grundrechenart ist zu diesem Zeitpunkt der Schullaufbahn noch nicht ausreichend gefestigt (Gliksman, Berebbi & Henik, 2022), weshalb sich besonders mit Berücksichtigung der rechenschwachen Zielgruppe gegen die Aufnahme von Divisionsaufgaben in den DTR-ASM entschieden wurde.

Bei den Additions- und Subtraktionsaufgaben handelt es sich um jeweils drei Aufgaben im Zahlenraum bis 20 und drei im Zahlenraum bis 100. Die sechs Multiplikationsaufgaben stammen aus dem kleinen Einmaleins und liegen damit ebenfalls im Zahlenraum bis 100. Die Aufgaben wurden dem HRT 1–4 (Haffner et al., 2005) entnommen und sollen mitsamt den *Prompts* in etwa dessen Durchführungsdauer entsprechen. Zur einfachen Testdurchführung wurde ein Testmanual verfasst,

das den Lehrkräften zur Verfügung gestellt wurde und genaue Anweisung zur Anwendung der *Prompts* enthielt. Die *Prompts* sind hierbei klar hierarchisch angeordnet und so einheitlich wie möglich gehalten, sodass nur minimale Unterschiede zwischen den *Prompts* für die verschiedenen Aufgabentypen bestehen. Während der Testdurchführung wurden die Lehrpersonen gebeten, einen Protokollbogen auszufüllen. In diesem vermerkten sie die angewandten *Prompts* für jede Aufgabe und die Häufigkeit motivational-emotionaler Hinweise für jede Aufgabenart (z. B. Addition im Zahlenraum bis 20).

4.1.2 Exemplarische Beschreibung einer Additionsaufgabe im Zahlenraum bis 20

Die Testdurchführung sowie das verwendete *Promptschema* soll nun an der Aufgabe „5 + 6“ exemplarisch dargestellt werden. Ein Überblick über das verwendete *Promptschema* findet sich zudem in Abbildung 2.

Nach der Präsentation der Aufgabe und der Aufforderung diese zu lösen, setzt das *Promptschema* bei einer fehlenden oder fehlerhaften Lösung mit der Aufforderung ein, sich zu überlegen, wie zur Lösung vorgegangen werden könnte *(Prompt 1a)*. Führt dies nicht zur korrekten Lösung, wird das Kind aufgefordert, sich Notizen zu Rechenschritten oder Teilergebnissen zu machen *(Prompt 1b)*, um zu sehen, ob eine Entlastung des Arbeitsgedächtnisses die Lösung der Aufgabe erleichtert, wie es Krajewski und Simanowski (2017) vermuten. Ist dies nicht der Fall, wird auf der Stufe des zweiten Prompts die Anwendung einer Rechenstrategie vorgeschlagen *(Prompt 2a)*.

Da für die Lösung von Additionsaufgaben verschiedene Rechenstrategien existieren (Padberg & Benz, 2021, S. 115–118), wurde für das Projekt eine konkrete Rechenstrategie vorgegeben. Die gewählte Strategie besteht darin, zunächst den ersten Summanden bis 10 aufzufüllen und anschließend den Rest des zweiten Summanden zu addieren. Dieser Ansatz bietet der Lehrkraft die Möglichkeit, zu beobachten, ob das Kind das Prinzip der Zahlzerlegung verstanden hat, was für die Entwicklung grundlegender arithmetischer Fähigkeiten bedeutsam ist (Wartha, Schulz & Benz, 2023). Wenn das Kind diese Rechenstrategie nicht erfolgreich anwendet, wird dem Kind die Möglichkeit gegeben, sich Notizen zu machen, um zu sehen, ob eine Entlastung des Arbeitsgedächtnisses in Kombination mit der Aufforderung zur Nutzung einer Rechenstrategie dem Kind die Lösung der Aufgabe ermöglicht *(Prompt 2b)*.

Erfolgt erneut keine korrekte Lösung, wird dem Kind vorgeschlagen, durch den Einsatz einer Zählstrategie zur richtigen Lösung der Aufgabenstellung zu gelangen *(Prompt 3a)*. Dabei gibt es zwei Möglichkeiten, Zählstrategien einzusetzen (Krajewski & Ennemoser, 2013), die auch beide im DTR-ASM angewandt werden dürfen. Die Lehrkraft kann das Kind entweder auffordern, die Strategie des Weiterzählens anzuwenden, bei der von einem der beiden Summanden bis zur korrekten Lösung weitergezählt wird, oder anbieten, die Aufgabe vollständig durch Zählen zu lösen. Diese Optionen sollen die Lehrkraft bei der Erkundung des Fähigkeitsniveaus des Kindes unterstützen. Ohne das Zählen mit den Fingern ist die Belastung des Ar-

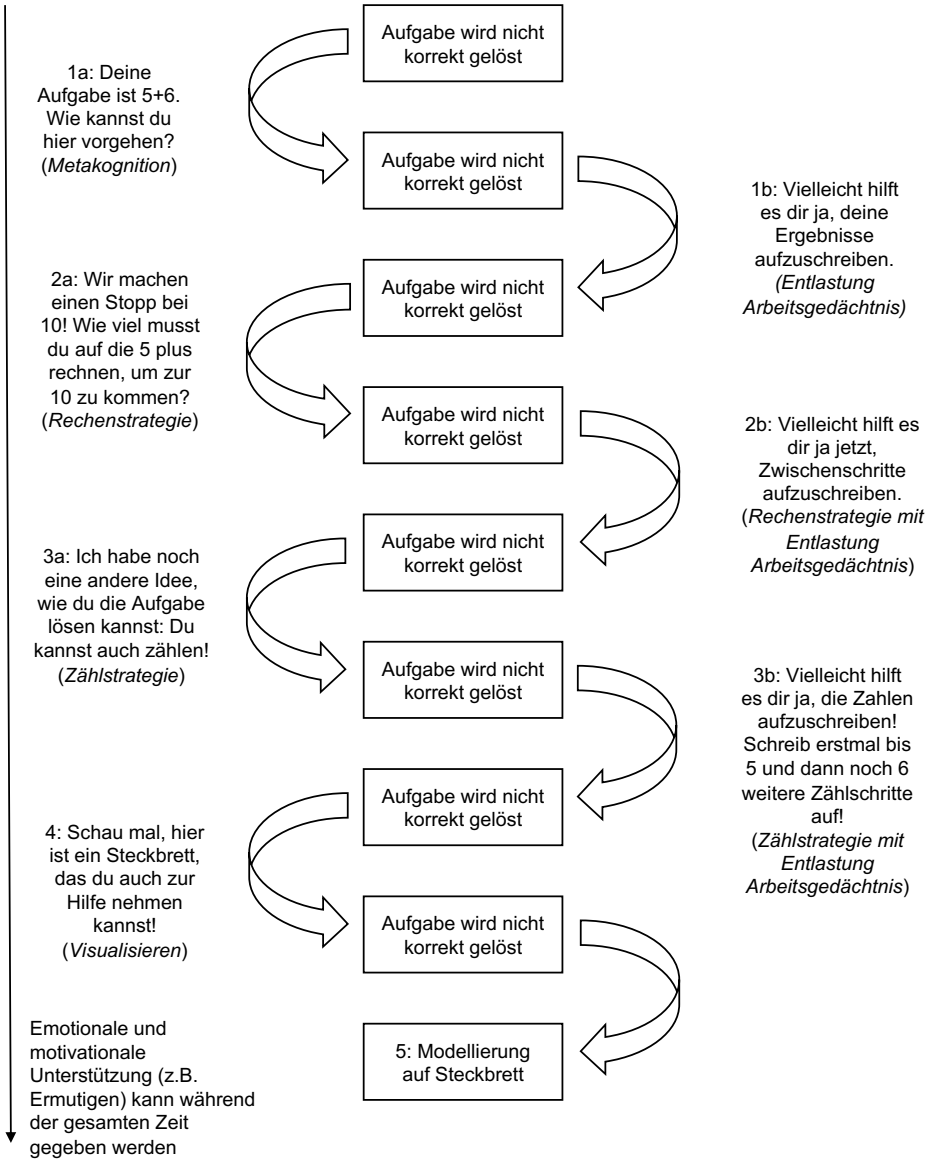

Abbildung 2: Protokoll der Unterstützung während der Testdurchführung; Hinweise wurden zu Abbildungszwecken angepasst.

beitsgedächtnisses beim Zählen und gleichzeitigen Behalten der Summanden weiterhin verhältnismäßig hoch. Deshalb wird bei fehlerhafter Lösung der nächste *Prompt* eingesetzt, um zu erkunden, ob das Kind an dieser Stelle vom Mitschreiben wie beispielweise der Notation der benötigten Zählschritte profitiert *(Prompt 3b)*.

Wenn die Aufgabe auch dann nicht korrekt gelöst ist, wird, ohne weitere Instruktion, ein Steckbrett zur Verfügung gestellt *(Prompt 4)*, um zu sehen, ob das Kind die

Visualisierung eigenständig zur Aufgabenlösung einsetzen kann. An dieser Stelle erhält die Lehrkraft Einblicke in die grundlegenden Fähigkeiten, wie das Abrufen der richtigen Zahlenfolge, das Abzählen und simultane Erfassen von Mengen sowie über das Verständnis grundlegender arithmetischer Prinzipien. Wenn das Kind die Aufgabe mit Hilfe des Steckbretts nicht selbst korrekt löst, wird die Aufgabe zusammen mit der Lehrkraft abgeschlossen *(Prompt 5)*. Dieser *Prompt* zielt vor allem darauf ab, Erfolgserlebnisse für das Kind zu schaffen und durch die Modellierung der Aufgabenlösung durch die Lehrkraft zu identifizieren, ob das Kind die vorgeführte Strategie bei der Lösung der darauffolgenden Aufgabe bereits anwenden kann, sodass die Festigung dieser Lösungsstrategie die Zone der nächsten Entwicklung des Kindes darstellen würde.

4.2 Entwicklung der Förderpläne und Förderplanratings

Um zu erfassen, inwieweit die durch die Lehrkräfte gesammelten diagnostischen Informationen in pädagogisches Handeln einfließen, wurden Förderpläne genutzt. Ein Förderplan ist „ein schriftlicher Plan zur gezielten Förderung von Schülerinnen und Schülern mit (sonder-)pädagogischem Förderbedarf oder Schülerinnen und Schülern, die von Schulversagen bedroht sind" (Melzer, 2014, S. 125). Für solche Förderpläne existieren eine Reihe von Qualitätskriterien, die eine effektive Förderplanung sicherstellen sollen. Darunter fallen neben der fachlichen und sachlichen Richtigkeit im Sinne einer wissenschaftlichen Fundierung und theoretischen Begründbarkeit des Vorgehens (Mutzeck, 2007) auch die Berücksichtigung des Entwicklungsstands des Kindes, die Festlegung von Förderzielen und konkreten dazugehörigen Fördermaßnahmen (Melzer, 2010).

Um zu untersuchen, ob die Durchführung des entwickelten dynamischen Tests einen Einfluss auf die Differenziertheit und Güte der Förderpläne sowie die Kohärenz der einzelnen Bestandteile des Förderplans zueinander hat, wurden die Pläne der drei Gruppen (DT-G, TT-G und KG) miteinander verglichen. Dabei wurde angenommen, dass sich die zusätzlichen diagnostischen Einblicke, die die dynamische Testdiagnostik gewähren kann, in einer höheren Qualität der Förderpläne widerspiegeln.

Zur Beurteilung der Güte der Förderpläne wurde vorab ein Ratingverfahren entwickelt. Obwohl zum Einsatz und den Effekten individueller Förderpläne in Schulen bereits Forschung vorliegt (insbesondere in Schweden; Andreasson, Asp-Onsjö & Isaksson, 2013; Isaksson, Lindqvist & Bergström, 2007; Rodger, 1995), existiert aktuell kein Ratingverfahren zur Erfassung der Qualität individueller Förderpläne. So untersuchten Isaksson et al. (2007) Förderpläne mittels einer Inhaltsanalyse hinsichtlich der dort dargestellten Probleme, führten jedoch keine Auswertung in Bezug auf die inhaltliche Qualität der Pläne durch. Eine entsprechende Inhaltsanalyse erfolgte ebenfalls im Rahmen des DYNAMIK-Projekts (Kap. 4.3).

4.2.1 Eingesetzte Ratingskala und Ablauf der Beurteilung der Förderpläne

Die entwickelte Ratingskala enthält neun Items mit Aussagen zur Differenziertheit der Darstellung des Entwicklungsstandes, der Lernziele sowie der Fördermaßnahmen. Die neun Items sind in Tabelle 1 dargestellt. Die Beurteilung erfolgt auf einer fünfstufigen Likert-Skala (1 = *trifft nicht zu*; 5 = *trifft voll zu*). Die Förderpläne von 99 Kindern wurden jeweils von sechs geschulten Rater:innen mittels der entwickelten Ratingskala beurteilt. Bei den Rater:innen handelt es sich um Studierende des Lehramts, der Lerntherapie und der Psychologie, die im Rahmen einer Fortbildung mit dem Aufbau der Pläne und der Bedeutung der Items der Ratingskala vertraut gemacht wurden und anhand eines exemplarischen fiktiven Förderplans ein erstes Rating durchführten. Durch das Rating eines fiktiven Plans und den gemeinsamen Austausch über dieses Rating sollte eine hohe Übereinstimmung der Ratings gewährleistet werden.

Tabelle 1: Deskriptive Statistik der Ratingskala zur Beurteilung der Güte der Förderpläne

Item	Mittelwert	*SD*	Median
Der Entwicklungsstand ist differenziert dargestellt.	3.3	1.1	3.0
Die Aspekte der Darstellung des Entwicklungsstands ergeben ein Gesamtbild.	3.1	1.1	3.0
Die Lern-/Förderziele sind differenziert dargestellt.	3.2	1.1	3.0
Die Lern-/Förderziele bauen aufeinander auf und ergänzen einander.	3.1	1.1	3.0
Die Fördermaßnahmen sind differenziert dargestellt.	3.0	1.1	3.0
Die Fördermaßnahmen bauen aufeinander auf und ergänzen einander.	3.0	1.2	3.0
Die Lern-/Förderziele und der Entwicklungsstand stehen in einem Zusammenhang.	3.2	1.0	3.0
Die Lern-/Förderziele und die Fördermaßnahmen stehen in einem Zusammenhang.	3.3	1.1	3.0
Die Fördermaßnahmen und der Entwicklungsstand stehen in einem Zusammenhang.	2.9	1.1	3.0

4.2.2 Ergebnisse des Förderplanratings

In einem ersten Analyseschritt erfolgte eine Betrachtung der Güte der entwickelten Ratingskala. Hierzu erfolgten Itemanalysen sowie eine explorative Faktorenanalyse. In den Ergebnissen zeigt sich eine ausgezeichnete Reliabilität der entwickelten Ratingskala (α = .94) sowie zufriedenstellende Werte in den Itemanalysen für Kennwerte wie Itemschwierigkeiten, Diskriminierung und Homogenität. Die Items der Skala laden zudem wie erwartet auf den Teilbereichen der Förderpläne (Entwicklungsstand, Lernziele, Förderung). Darüber hinaus erfolgte eine Betrachtung der In-

terraterreliabilität. Hier zeigten sich akzeptable bis zufriedenstellende Werte bei Betrachtung der *ICC*s, sodass von einer grundsätzlichen Eignung dieses Verfahrens zur Beurteilung der Qualität von Förderplänen auszugehen ist.

In einem zweiten Analyseschritt wurde nun die Güte der Förderpläne mittels der Ratings über die Bedingungen verglichen (Tabelle 2). Erste Analysen weisen darauf hin, dass sich die Verwendung der dynamischen Testdiagnostik nicht in einer erhöhten Güte der Förderpläne, sowohl verglichen zu statusdiagnostischer Diagnostik als auch zu informeller Diagnostik, äußert ($ß$ = -0.21). Die Ergebnisse deuten im Gegenteil vielmehr daraufhin, dass weitere Merkmale der Lehrkräfte entscheidend für den Großteil der Varianzaufklärung sind (*ICC* = .57). Darüber hinaus zeichnet sich ein signifikanter Effekt des Alters und des Geschlechts der Schüler:innen auf die Qualität ab: Für ältere ($ß$ = 0.05, p = .002) und weibliche ($ß$ = -0.07, p < .001) Schüler:innen zeigt sich eine höhere Qualität der Förderpläne.

Tabelle 2: Prädiktoren für die Güte der Förderpläne

Prädiktoren	Güte des Förderplans					
	Modell 1		Modell 2		Modell 3	
	ß	*p*	*ß*	*p*	*ß*	*p*
(Konstante)	0.16	.29	0.18	.24	0.15	.31
TT-G	-0.15	.32	-0.15	.31	-0.15	.29
DT-G	-0.15	.28	-0.22	.12	-0.21	.11
Hildesheim			-0.15	.09	-0.15	.09
Potsdam			0.13	.12	0.14	.08
Alter					0.05	.002**
Geschlecht m.					-0.07	<.001***
Modellparameter						
ICC	0.58		0.57		0.57	
Beobachtungen	569		569		563	
Marginal R^2 / Conditional R^2	0.02 / 0.58		0.05 / 0.59		0.08 / 0.60	

Anmerkungen. TT-G = statusdiagnostischer Test; DT-G = dynamischer Test; Geschlecht m. = männliches Geschlecht der Kinder.
** p < .01. *** p < .001.

4.3 Inhaltsanalytische Betrachtung der Förderpläne

Neben der Betrachtung der Qualität der Förderpläne lag ein wesentliches Erkenntnisinteresse des Projekts darin, ob sich die Förderpläne hinsichtlich der genannten Inhalte und Prozesse zwischen den Untersuchungsbedingungen unterscheiden. Hierzu wurde zunächst eine qualitative Inhaltsanalyse nach Mayring (2015) durchgeführt, um die Inhalte der Förderpläne zu kategorisieren. Dazu wurde in einem ersten Schritt ein Kodiermanual entwickelt, welches die inhaltlichen Kategorien beschreibt, die in den Förderplänen auftauchen *könnten*. Für jede Kategorie wurden zudem Ankerbeispiele aus den Förderplänen angegeben. In dem ersten Entwurf des Katego-

riensystems wurden die Kategorien auf theoretischen Annahmen basierend verfasst (z. B. in Anlehnung an die Entwicklung arithmetischer Fähigkeiten; Annahmen über diagnostische Informationen, die Lehrkräfte durch die Anwendung der Testverfahren erhalten; möglichen Einflussvariablen auf die Rechenfähigkeit). In einem zweiten Schritt wurden nach der Analyse von etwa 20 Prozent der Förderpläne weitere Kategorien induktiv ergänzt, die sich zwar in den Förderplänen finden ließen, sich aber nicht zwangsläufig aus theoretischen Vorannahmen ableiten lassen. Nach der Ausarbeitung des Kodiermanuals werteten drei Mitarbeiterinnen des Projekts zunächst gemeinsam eine Zufallsauswahl von Förderplänen aus. Dies diente der Klärung von Fragen sowie der Definition von induktiven Kategorien. Im Anschluss wurden zwanzig zufällig ausgewählte Förderpläne unabhängig voneinander ausgewertet. Die Ratings dienten als Grundlage für die Berechnung der Interraterreliabilität, welche in einem angemessenen Bereich lag. Abschließend wurden die verbleibenden Förderpläne aufgeteilt, sodass jeweils ein Drittel der Förderpläne von einer Raterin ausgewertet wurde.

4.3.1 Ergebnisse der Inhaltsanalyse

Nachfolgend sollen die Ergebnisse der Inhaltsanalyse überblicksartig dargestellt werden. Aufgrund der großen Anzahl an gefundenen Kategorien wird sich in diesem Text auf die a priori formulierten Kategorien (arithmetische Fähigkeiten, motivational-emotionale Prozesse, kognitive Prozesse, Unterstützungsbedarf) fokussiert, da diese den engsten inhaltlichen Zusammenhang zu den durchgeführten diagnostischen Verfahren aufwiesen. Darüber hinaus wird in diesem Kapitel auf eine Auswertung des dritten Teils der Förderpläne verzichtet. Da die Lehrkräfte gebeten wurden, anhand einer Tabelle Möglichkeiten zur Förderung der Kinder anzugeben, schrieben sie häufig nur einzelne Wörter in die Tabellenzellen, sodass in diesem Bereich generell wenig Informationen über die Kinder zu finden waren. Um einen Einblick in die Unterschiede zwischen den Gruppen zu gewinnen, wurde in einem ersten Schritt für alle Bereiche untersucht, wie viele unterschiedliche Kategorien in den Förderplänen genannt wurden. Dafür wurde für alle Subkategorien kodiert, ob diese im Förderplan genannt wurden, und anschließend ein Summenscore über die Anzahl der genannten Subkategorien gebildet. Anhand von Varianzanalysen wurde untersucht, ob sich diese Summenwerte zwischen den Gruppen unterscheiden. Im einem zweiten Analyseschritt wurde betrachtet, in wie vielen der Förderpläne der Gruppen die einzelnen Subkategorien auftraten, um einen detaillierten Überblick zu erhalten. Mithilfe von χ^2-Tests wurde untersucht, ob sich Unterschiede zwischen den drei Untersuchungsgruppen finden lassen. Insgesamt zeigen die Analysen dabei kein eindeutiges Ergebnismuster.

Arithmetische Fähigkeiten
In der Beschreibung des Lernstandes des Kindes unterschieden sich die Gruppen signifikant in der Anzahl an unterschiedlichen Kategorien aus dem Bereich der arith-

metischen Fähigkeiten. Post-hoc-Analysen zeigten hierbei, dass in der Kontrollgruppe mehr unterschiedliche Prozesse genannt wurden als in der Gruppe, die den HRT angewandt hat. Bei der Analyse der einzelnen Subkategorien zeigten sich Unterschiede in der Nennung des Stellenwertsystems sowie von Rechenstrategien. Obwohl sich die Anzahl an unterschiedlichen Kategorien bei der Definition der Lernziele insgesamt nicht unterschied, konnten auf der Ebene einzelner Kategorien Unterschiede gefunden werden. Die Gruppen unterschieden sich hierbei bezüglich des Anteils an Förderplänen, in welchen Lernziele zur Mengenerfassung, zum Stellenwertsystem sowie zu nicht näher definierten arithmetischen Fähigkeiten genannt wurden.

Motivational-emotionale Prozesse
Auch in der Anzahl an unterschiedlichen motivational-emotionalen Prozessen bei der Beschreibung des Entwicklungsstands des Kindes unterschieden sich die Gruppen signifikant voneinander. Post-hoc-Analysen zeigten hier, dass sich die DT-G sowie die TT-G signifikant voneinander unterschieden und in der DT-G weniger Prozesse genannt wurden. Auf der Ebene einzelner Subkategorien fanden sich Unterschiede in der Erwähnung des Selbstkonzepts sowie der Motivation. Anders als bei den arithmetischen Fähigkeiten wurden bei der Definition der Lernziele keinerlei Unterschiede zwischen den Gruppen gefunden.

Kognitive Prozesse
Bei der Untersuchung der Erwähnung kognitiver Prozesse konnten in der Beschreibung des Entwicklungsstands des Kindes keine Unterschiede zwischen den Gruppen gefunden werden. Auch bei der Definition der Lernziele fanden sich lediglich in Bezug auf die Erwähnung des logischen Denkens Unterschiede zwischen den Gruppen.

Unterstützungsbedarf
Bei der Beschreibung des Entwicklungsstands der Kinder beschrieben Lehrkräfte der DT-G mehr Unterstützungsbedarf auf Seiten des Kindes als Lehrkräfte der TT-G. Unterschiede fanden sich auf der Ebene der Subkategorien in Bezug auf motivational-emotionale sowie kognitive Unterstützung. Diese Unterschiede zeigten sich jedoch nicht mehr bei der Definition der Lernziele.

5. Diskussion und Einordnung der Ergebnisse

Das übergeordnete Ziel des DYNAMIK-Projekts war, das dynamische Testen auf den Bereich des Rechnens zu übertragen und zu untersuchen, inwiefern dieses für die Förderplanung hilfreich ist. Hierzu wurde im Projekt der erste deutschsprachige dynamische Test zur Erfassung von Rechenfertigkeiten entwickelt. Bei der Entwicklung des Testverfahrens gelang es, Prinzipien und Modelle (hier insbesondere den *Graduated-Prompt*-Ansatz) auf den Bereich der Arithmetik zu übertragen.

Darüber hinaus wurde die Nützlichkeit dieses Tests im Hinblick auf die Förderplanung von Lehrkräften untersucht. Dazu wurden Lehrkräfte in dessen Anwendung

des Verfahrens geschult und anschließend gebeten, eine Diagnostik mit rechenschwachen Kindern durchzuführen sowie für diese Kinder Förderpläne zu verfassen. Diese Pläne wurden dann mit Plänen von Lehrkräften, die einen statusdiagnostischen Test durchgeführt hatten, sowie den Plänen von Lehrkräften, die keine diagnostischen Testverfahren verwendet hatten, verglichen. Ergebnisse eines anschließenden Ratings der Förderpläne dokumentieren dabei keine Unterschiede in der Güte der Förderpläne in Abhängigkeit von der Fortbildungsbedingung. In den Analysen zum Inhalt der Förderpläne lassen sich Unterschiede zwischen den Gruppen finden, jedoch lässt sich aus den Ergebnissen insgesamt kein klares Bild bezüglich der Nützlichkeit des dynamischen Tests im Hinblick auf die Förderplanung ableiten. Die Interpretation der gefundenen Ergebnisse wird insbesondere durch mehrere Limitationen erschwert, welche im Folgenden kurz dargestellt werden sollen.

5.1 Limitationen

Die Durchführung des Projekts war von zahlreichen Herausforderungen geprägt, welche in dieser Form so nicht vorhersehbar waren. Eine der größten Limitationen der Studie, welche Auswirkungen auf die Auswertung sowie die Interpretation der Ergebnisse der Studie hatte, ist die hohe Drop-Out-Rate. Anfänglich gedacht war es, in jeder Bedingung dreißig Lehrkräfte bzw. je neunzig Förderpläne von dreißig Lehrkräften zu untersuchen. Trotz intensiver Bemühungen bei der Akquise konnte diese Zahl schon zu Beginn der Datenerhebung nicht erreicht werden, sodass bereits mit einer kleineren Stichprobe als ursprünglich geplant gearbeitet wurde. In allen Gruppen kam es während der Datenerhebung zu Drop-Outs, sodass am Ende nur mit 99 Förderplänen von insgesamt 37 Lehrkräften für alle Bedingungen gearbeitet wurde. Problematisch ist, dass sich die Drop-Outs ungleichmäßig auf die Untersuchungsgruppen verteilten, sodass sich die resultierenden Gruppengrößen stark unterschieden (DT-G: $n = 18$, TT-G: $n = 10$, KG: $n = 9$). Die hohen Drop-Out-Raten können darauf hindeuten, dass insbesondere die Lehrkräfte die Datenerhebung abschlossen, die besonders motiviert zur Teilnahme an der Studie waren, was zu Verzerrungen in den Ergebnissen beitragen könnte. Darüber hinaus verhinderte die geringe Anzahl an Lehrkräften, dass die genestete Datenstruktur in einigen der Analysen beachtet werden konnte. Dies ist besonders kritisch zu sehen, weil die Analysen zur Kohärenz der Pläne darauf hindeuten, dass die Qualität der Pläne zu einem Großteil durch die Lehrkraft determiniert wird.

Eine weitere Limitation ergibt sich aus der Tatsache, dass die Lehrkräfte die Kinder teilweise seit mehreren Jahren in Mathematik unterrichteten. Dadurch hatten sie wahrscheinlich weitgehende Einblicke in den Entwicklungsstand des Kindes, welche sie durch die alleinige Anwendung der Tests nicht bekommen hätten. Effekte der Testverfahren könnten demnach durch das Wissen der Lehrkraft über die Kinder überlagert worden sein, sodass sich kaum noch Unterschiede zwischen den Gruppen finden ließen. In zukünftigen Studien könnte dies umgangen werden, indem die

Wirkung des dynamischen Testverfahrens in Situationen untersucht wird, in denen sich die Lehrkraft und das Kind noch nicht kennen.

Die dritte Limitation betrifft die Operationalisierung der Nützlichkeit in den vorliegenden Analysen. Hierbei wurde sich auf die Qualität sowie die Inhalte der Förderpläne bezogen. Die Nützlichkeit für die Förderung von Kindern sollte jedoch auch ihre individuellen Bedürfnisse und die Passung zwischen den Bedürfnissen des Kindes und der Förderplanung berücksichtigen. Darüber hinaus wurden in den Förderplänen lediglich Absichten zur Förderung der Kinder durch die Lehrkräfte formuliert. Es blieb jedoch unklar, inwieweit sich die formulierten Ansätze auch tatsächlich im pädagogischen Handeln der Lehrkraft und damit der Förderung des Kindes widerspiegeln. Abschließend ist kritisch zu diskutieren, dass das verwendete dynamische Testverfahren im Rahmen des Projekts neu entwickelt wurde. Entsprechende tiefergehende Untersuchungen bezüglich der Validität sowie der Reliabilität der Testergebnisse stehen dementsprechend noch aus.

6. Fazit

Obwohl die Ergebnisse des Projekts vor dem Hintergrund verschiedener Limitationen betrachtet werden müssen, ergänzen sie den bisherigen Forschungsstand zum dynamischen Testen an mehreren Stellen. So lag bisher kein deutschsprachiger dynamischer Test zum Erfassen von Rechenfähigkeiten vor. Mit der Entwicklung des DTR-ASM wurde ein erster Versuch unternommen, ein solches Verfahren zu entwickeln. Weitere Studien zur Weiterentwicklung und Validierung des Tests werden benötigt und schließen sich an.

In der Vergangenheit wurde bereits untersucht, wie Lehrkräfte die Ergebnisse von dynamischen Testverfahren für die Förderplanung einschätzen. Dabei zeigte sich eine Präferenz für die Ergebnisse dynamischer Tests (Bosma & Resing, 2012; Bosma et al., 2012). In den bisherigen Studien erhielten die Lehrkräfte die Informationen jedoch häufig in Form von Testberichten bzw. Vignetten. Es war daher noch unklar, ob sie durch die eigenständige Anwendung eines dynamischen Testverfahrens auch selbstständig in der Lage sein würden, diese Informationen für ihre eigene Förderplanung zu nutzen. Mit dem durchgeführten Projekt wurde angestrebt, diese Forschungslücke zu schließen, sodass trotz der Limitationen neue Impulse für die weitere Forschung zum dynamischen Testen generiert werden konnten.

Literatur

Andreasson, I., Asp-Onsjö, L. & Isaksson, J. (2013). Lessons learned from research on individual educational plans in Sweden: Obstacles, opportunities and future challenges. *European Journal of Special Needs Education*, *28*(4), 413–426. https://doi.org/10.1080/08856257.2013.812405

Balt, M., Fritz, A., & Ehlert, A. (2020). Insights Into First Grade Students' Development of Conceptual Numerical Understanding as Drawn from Progression-Based Assessments. *Frontiers in Education*, *5*, 80. https://doi.org/10.3389/feduc.2020.00080

Bildungsserver Berlin-Brandenburg. (2015). *Rahmenlehrplan Mathematik 1–10*. https://bildungsserver.berlin-brandenburg.de/rlp-online/c-faecher/mathematik/materialien

Bodovski, K. & Farkas, G. (2007). Mathematics Growth in Early Elementary School: The Roles of Beginning Knowledge, Student Engagement, and Instruction. *The Elementary School Journal*, *108*(2), 115–130. https://doi.org/10.1086/525550

Börnert, M. & Wilbert, J. (2016). Dynamisches Testen als neue Perspektive in der sonder-pädagogischen Diagnostik – Theorie, Evidenzen, Impulse für Forschung und Praxis. *Zeitschrift für Heilpädagogik*, *67*, 156–167.

Börnert-Ringleb, M. & Wilbert, J. (2019). Beitrag eines dynamischen gegenüber einem statischen Test kognitiver Fähigkeiten zur Vorhersage der Schulleistung. *Empirische Sonderpädagogik*, *11*(3), 175–190. http://www.psychologie-aktuell.com/index.php?id=empirische-sonderpaedagogik

Börnert-Ringleb, M. & Wilbert, J. (2020). Die Vorhersage von Mathe- und Leseleistungen durch dynamisches Testen. *Lernen und Lernstörungen*, *10*(2), 102–113. https://doi.org/10.1024/2235-0977/a000331

Bosma, T. & Resing, W. C. M. (2012). Need for instruction: Dynamic testing in special education. *European Journal of Special Needs Education*, *27*(1), 1–19. https://doi.org/10.1080/08856257.2011.613599

Bosma, T., Hessels, M. G. P. & Resing, W. C. M. (2012). Teachers' preferences for educational planning: Dynamic testing, teaching' experience and teachers' sense of efficacy. *Teaching and Teacher Education*, *28*(4), 560–567. https://doi.org/10.1016/j.tate.2012.01.007

Bundschuh, K. & Winkler, C. (2014). *Einführung in die sonderpädagogische Diagnostik* (8., überarbeitete Auflage). Ernst Reinhardt Verlag.

Caffrey, E., Fuchs, D. & Fuchs, L. S. (2008). The Predictive Validity of Dynamic Assessment: A Review. *The Journal of Special Education*, *41*(4), 254–270. https://doi.org/10.1177/0022466907310366

Cho, E., Compton, D. L., Fuchs, D., Fuchs, L. S. & Bouton, B. (2014). Examining the Predictive Validity of a Dynamic Assessment of Decoding to Forecast Response to Tier 2 Intervention. *Journal of Learning Disabilities*, *47*(5), 409–423. https://doi.org/10.1177/0022219412466703

Claessens, A. & Engel, M. (2013). How Important is Where you Start? Early Mathematics Knowledge and Later School Success. *Teachers College Record, 115(6),* 1–29. https://doi.org/10.1177/016146811311500603.

Dehaene, S. (1992). Varieties of numerical abilities. *Cognition*, *44*(1–2), 1–42. https://doi.org/10.1016/0010-0277(92)90049-N

Dörfler, T., Golke, S. & Artelt, C. (2009). Dynamic assessment and its potential for the assessment of reading competence. *Studies in Educational Evaluation*, *35*(2–3), 77–82. https://doi.org/10.1016/j.stueduc.2009.10.005

Dowker, A. (2008). Individual differences in numerical abilities in preschoolers. *Developmental Science*, *11*(5), 650–654. https://doi.org/10.1111/j.1467-7687.2008.00713.x

Fuchs, D., Compton, D. L., Fuchs, L. S., Bouton, B. & Caffrey, E. (2011). The Construct and Predictive Validity of a Dynamic Assessment of Young Children Learning to Read: Implications for RTI Frameworks. *Journal of Learning Disabilities*, *44*(4), 339–347. https://doi.org/10.1177/0022219411407864

Fuchs, L. S., Compton, D. L., Fuchs, D., Hollenbeck, K. N., Craddock, C. F. & Hamlett, C. L. (2008). Dynamic assessment of algebraic learning in predicting third graders' development of mathematical problem solving. *Journal of Educational Psychology*, *100*(4), 829–850. https://doi.org/10.1037/a0012657

Haffner, J., Baro, K., Parzer, P. & Resch, F. (2005). *HRT 1–4 – Heidelberger Rechentest*. Hogrefe.

Hessels, M. G. P. (1997). Low IQ but high learning potential: why Zeyneb and Moussa do not belong in special education. *Educational and Child Psychology*, *14*, 121–136.

Gliksman, Y., Berebbi, S. & Henik, A. (2022). Math Fluency during Primary School. *Brain Sciences*, *12*(3), 371. https://doi.org/10.3390/brainsci12030371

Guthke, J. (1992). Learning tests – the concept, main research findings, problems and trends. *Learning and Individual Differences*, *4*(2), 137–151. https://doi.org/10.1016/1041-6080(92)90010-C

Guthke, J., Beckmann, J. F. & Wiedl, K. H. (2003). Dynamik im dynamischen Testen. *Psychologische Rundschau*, *54*(4), 225–232. https://doi.org/10.1026//0033-3042.54.4.225

Isaksson, J., Lindqvist, R. & Bergström, E. (2007). School problems or individual shortcomings? A study of individual educational plans in Sweden. *European Journal of Special Needs Education*, *22*(1), 75–91. https://doi.org/10.1080/08856250601082323

Kaufmann, L., Mazzocco, M. M., Dowker, A., Von Aster, M., Göbel, S. M., Grabner, R. H., Henik, A., Jordan, N. C., Karmiloff-Smith, A. D., Kucian, K., Rubinsten, O., Szucs, D., Shalev, R. & Nuerk, H.-C. (2013). Dyscalculia from a developmental and differential perspective. *Frontiers in Psychology*, *4*. https://doi.org/10.3389/fpsyg.2013.00516

Krajewski, K. & Ennemoser, M. (2013). Entwicklung und Diagnostik der Zahl-Größen-Verknüpfung zwischen 3 und 8 Jahren. In M. Hasselhorn, A. Heinze, W. Schneider & U. Trautwein (Hrsg.), *Test und Trends: Bd. 11. Diagnostik mathematischer Kompetenzen* (S. 41–66). Hogrefe.

Krajewski, K. & Simanowski, S. (2017). Qualitätskriterien für Förderansätze zur Prävention von Rechenschwäche. *Frühförderung Interdisziplinär*, *36*(2), 93. https://doi.org/10.2378/fi2017.art08d

Maehler, C., & Schuchardt, K. (2016). Working memory in children with specific learning disorders and/or attention deficits. *Learning and Individual Differences*, *49*, 341–347. https://doi.org/10.1016/j.lindif.2016.05.007

Mann, A., Fischer, U. & Nuerk, H.-C. (2013). TEDI-MATH – Test zur Erfassung numerisch-rechnerischer Fertigkeiten vom Kindergarten bis zur 3. Klasse. In M. Hasselhorn, A. Heinze, W. Schneider & U. Trautwein (Hrsg.), *Tests und Trends: Bd. 11. Diagnostik mathematischer Kompetenzen* (S. 97–112). Hogrefe.

Mandinach, E. B. (2012). A Perfect Time for Data Use: Using Data-Driven Decision Making to Inform Practice. *Educational Psychologist*, *47*(2), 71–85. https://doi.org/10.1080/00461520.2012.667064

Mayring, P. (2015). Qualitative Content Analysis: Theoretical Background and Procedures. In A. Bikner-Ahsbahs, C. Knipping & N. Presmeg (Hrsg.), *Approaches to Qualitative Research in Mathematics Education* (S. 365–380). Springer Netherlands. https://doi.org/10.1007/978-94-017-9181-6_13

Melzer, C. (2010). Wie können Förderpläne effektiv sein und eine professionelle Förderung unterstützen? *Zeitschrift Für Heilpädagogik*, *61*(6), 212–220.

Melzer, C. (2014). Förderplanung. In U. Heimlich, R. Stein & F. Wember (Hrsg.), *Handlexikon Lernschwierigkeiten und Verhaltensstörungen* (S. 125–128). Kohlhammer.

Mercader, J., Miranda, A., Presentación, M. J., Siegenthaler, R. & Rosel, J. F. (2018). Contributions of Motivation, Early Numeracy Skills, and Executive Functioning to Mathematical Performance. A Longitudinal Study. *Frontiers in Psychology*, *8*, 2375. https://doi.org/10.3389/fpsyg.2017.02375

Mutzeck, W. (2007). Förderplanung. In K. Bundschuh, U. Heimlich & R. Krawitz (Hrsg.), *Wörterbuch Heilpädagogik* (3. Aufl., S.79–84). Klinkhardt.

Niedersächsisches Kultusministerium (Hrsg.). (2017). Kerncurriculum für die Grundschule Schuljahrgänge 1–4: Mathematik. In *Niedersächsisches Landesinstitut für Schulische Qualitätsentwicklung*. https://cuvo.nibis.de/cuvo.php?p=search&k0_2=Schulbereich&v0_2=Primarbereich&k0_5=Schulform&v0_5=Grundschule&k0_6=Fach&v0_6=Mathematik

Pameijer, N. (2006). Towards needs-based assessment: Bridging the gap between assessment and practice. *Educational and Child Psychology*, *23*(3), 12–24. https://doi.org/10.53841/bpsecp.2006.23.3.12

Padberg, F. & Benz, C. (2021). *Didaktik der Arithmetik: Fundiert, vielseitig, praxisnah* (5., überarbeitete Auflage). Springer Spektrum.

Parsons, S. & Bynner, J. (2005). *Does Numeracy matter more?* Institute of Education University of London.

Resing, W. C. M. (2013). Dynamic Testing and Individualized Instruction: Helpful in Cognitive Education? *Journal of Cognitive Education and Psychology*, *12*(1), 81–95. https://doi.org/10.1891/1945-8959.12.1.81

Resing, W. C. M., Stevenson, C. E. & Bosma, T. (2012). Dynamic Testing: Measuring Inductive Reasoning in Children with Developmental Disabilities and Mild Cognitive Impairments. *Journal of Cognitive Education and Psychology*, *11*(2), 159–178. https://doi.org/10.1891/1945-8959.11.2.159

Resing, W. C. M., Tunteler, E., De Jong, F. M. & Bosma, T. (2009). Dynamic testing in indigenous and ethnic minority children. *Learning and Individual Differences*, *19*(4), 445–450. https://doi.org/10.1016/j.lindif.2009.03.006

Resing, W. C. M. & Elliott, J. G. (2011). Dynamic testing with tangible electronics: Measuring children's change in strategy use with a series completion task. *British Journal of Educational Psychology, 81*(4), 579–605. https://doi.org/10.1348/2044-8279.002006

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. https://doi.org/10.1177/0956797612466268

Rodger, S. (1995). Individual Education Plans Revisited: A Review of the Literature. *International Journal of Disability, Development and Education*, *42*(3), 221–239. https://doi.org/10.1080/0156655950420305

Roick, T., Gölitz, D. & Hasselhorn, M. (2004). *DEMAT 3+ Deutscher Mathematiktest für dritte Klassen*. Hogrefe.

Roick, T., Gölitz, D. & Hasselhorn, M. (2013). Affektive Komponenten der Mathematikkompetenz. Die Mathematikangst-Ratingskala für vierte bis sechste Klassen (MARS 4–6). In M. Hasselhorn, A. Heinze, W. Schneider & U. Trautwein (Hrsg.), *Tests und Trends: Bd. 11. Diagnostik mathematischer Kompetenzen* (S. 205–221). Hogrefe.

Schildkamp, K., Lai, M. K. & Earl, L. (Hrsg.). (2013). *Data-based Decision Making in Education*. Springer Netherlands. https://doi.org/10.1007/978-94-007-4816-3

Schlee, J. (2008). 30 Jahre „Förderdiagnostik" – eine kritische Bilanz. *Zeitschrift für Heilpädagogik*, *4*, 122–133.

Schuchardt, K. & Mähler, C. (2010). Unterscheiden sich Subgruppen rechengestörter Kinder in ihrer Arbeitsgedächtniskapazität, im basalen arithmetischen Faktenwissen und in den numerischen Basiskompetenzen? *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, *42*(4), 217–225. https://doi.org/10.1026/0049-8637/a000022

Schukajlow, S., Rakoczy, K. & Pekrun, R. (2017). Emotions and motivation in mathematics education: Theoretical considerations and empirical contributions. *ZDM*, *49*(3), 307–322. https://doi.org/10.1007/s11858-017-0864-6

Seethaler, P. M., Fuchs, L. S., Fuchs, D. & Compton, D. L. (2012). Predicting first graders' development of calculation versus word-problem performance: The role of dynamic assessment. *Journal of Educational Psychology*, *104*(1), 224–234. https://doi.org/10.1037/a0024968

Sittner Bridges, M. & Catts, H. W. (2011). The use of a dynamic screening of phonological awareness to predict risk for reading disabilities in kindergarten children. *Journal of Learning Disabilities*, *44*(4), 330–338. https://doi.org/10.1177/0022219411407863

Sternberg, R. J. & Grigorenko, E. L. (2002). *Dynamic Testing: The Nature and Measurement of Learning Potential*. Cambridge University Press.

Stevenson, C. E., Heiser, W. J. & Resing, W. C. M. (2016). Dynamic testing: Assessing cognitive potential of children with culturally diverse backgrounds. *Learning and Individual Differences*, *47*, 27–36. https://doi.org/10.1016/j.lindif.2015.12.025

Tiekstra, M., Hessels, M. G. P. & Minnaert, A. E. M. G. (2009). Learning Capacity in Adolescents with Mild Intellectual Disabilities. *Psychological Reports*, *105*(3), 804–814. https://doi.org/10.2466/PR0.105.3.804-814

von Aster, M., Schweiter, M. & Weinhold Zulauf, M. (2007). Rechenstörungen bei Kindern. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, *39*(2), 85–96. https://doi.org/10.1026/0049-8637.39.2.85

Vygotskij, L. S. & Cole, M. (1981). *Mind in society: The development of higher psychological processes* (Nachdr.). Harvard University Press.

Wartha, S., Schulz, A. & Benz, C. (2023). Zusammenhänge zwischen Zahlzerlegungen, Addition und Subtraktion im Zahlenraum bis 10: Empirische Untersuchung der Kompetenzen von Lernenden im zweiten und dritten Schuljahr. *Lernen und Lernstörungen*, *12*(3), 155–167. https://doi.org/10.1024/2235-0977/a000415

Wilbert, J. & Börnert, M. (2015). Unterricht. In I. Hedderich, G. Biewer, J. Hollenweger & R. Markowetz (Hrsg.), *Handbuch Inklusion und Sonderpädagogik* (1. Aufl, S. 346–353). UTB.