

FACHBEITRAG

Wirksam Fördern Lernen

Ein Design-Based-Research-Ansatz für formatives Fördern in der sonderpädagogischen Lehrkräftebildung im sonderpädagogischen Schwerpunkt Lernen

Judith Zellner^{id}, Nikola Ebenbeck^{id}, Markus Gebhardt^{id}



Zusammenfassung

Pädagogische Diagnostik, Progress Monitoring und individualisierte Förderung zu vermitteln und zu erproben, sind Ziele der universitären Ausbildung in der Sonderpädagogik. Die vorliegende Studie evaluiert die in der Ausbildung durchgeführte Diagnostik und Förderung im studienbegleitenden Praktikum unter Einsatz eines Design-Based-Research-Ansatzes. Studierende ($N = 18$) förderten in kleinen Gruppen ($N = 43$) Schüler:innen im Lesen über 5 Messzeitpunkte anhand des begleitenden formativen Assessments von Levumi mit Prä- und Postphase im sonderpädagogischen Förderzentrum. Die eigene diagnostische Kompetenz wurde im Vorfeld und in Anschluss an die Fördereinheiten von den Studierenden eingeschätzt. Es zeigte sich eine Verbesserung über alle Dimensionen (fachdidaktische Diagnostikkompetenzen, Statusdiagnostik, Förderdiagnostik und Lernverlaufsdiagnostik) mit einem großen Effekt im Bereich der Lernverlaufsdiagnostik ($d = 1,21$). Zur Auswahl der geförderten Schüler:innen wurde verbindlich ein Screening durchgeführt, allerdings begründeten nur 34 % der Studierenden ihre Auswahl der Schüler:innen mithilfe der Testergebnisse. Die Förderung wurde in 89 % der Fälle außerhalb des regulären Unterrichts durchgeführt. Insgesamt verbesserten sich die Schüler:innen mit Effektstärken von $PND = 77\%$, $PEM = 83\%$ und $PAND = 84\%$ im Lesen. Diskutiert wird, inwieweit individuelle Förderung in Förderschulen systematisch umgesetzt und evidenzbasierte Verfahren als integraler Bestandteil des Unterrichts eingesetzt werden sollen, müssen und werden.

Keywords: Formatives Assessment, Lernverlaufsdiagnostik, Data-Based Decision-Making, Lehrkräftebildung, individuelle Förderung, Design-Based Research

1 DIAGNOSTIK UND FÖRDERUNG IM SONDERPÄDAGOGISCHEN SCHWERPUNKT LERNEN

Sonderpädagogik im Schwerpunkt Lernen ermöglicht Lernenden mit Lernschwierigkeiten in inklusiven Schulen und Förderschulen Lernerfolge und Teilhabe (Gebhardt, 2024). Zentrale Grundlage dieses Handelns ist sonderpädagogische Diagnostik. Nach der Empfehlung der KMK (2019) umfasst sie einerseits Feststellungsdiagnostik, etwa zur Entscheidung über zieldifferente Unterrichtsangebote,



This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.
© 2026 The Author(s). Breaking Barriers to Learning and Education published by Ludwigs-Maximilians-Universität München and Ludwigsburg University of Education in association with the German Hochschulkonferenz Sonderpädagogik des Lernens

und andererseits Förderdiagnostik, die der Planung individueller Bildungsangebote dient. Für sonderpädagogische Lehrkräfte ergibt sich daher ein doppelter Auftrag. Sie müssen sowohl diagnostische Entscheidungen über Förderbedarfe treffen, als auch Lernprozesse kontinuierlich beobachten, interpretieren und die für Gestaltungen von Förderung nutzbar machen.

Gerade im Schwerpunkt Lernen ist diese Verbindung von Diagnostik und Förderung besonders bedeutsam. Die Zielgruppe umfasst Lernende, die im allgemeinen Unterricht trotz unterrichtlicher Differenzierung und zusätzlicher Übung häufig nur geringe Lernfortschritte erzielen. Im englischsprachigen Raum werden diese Lernenden als *struggling students* bezeichnet (Vaughn et al., 2003). Für sie reicht eine bloße Intensivierung regulärer Unterrichtsmethoden häufig nicht aus. Vielmehr müssen Förderangebote eng an individuelle Lernvoraussetzungen, aktuelle Lernstände und beobachtbare Lernentwicklungen angepasst werden, womit die formative Funktion der Diagnostik in den Vordergrund rückt. Sonderpädagogische Diagnostik soll in diesem Kontext sichtbar machen, ob eine gewählte Förderung wirksam ist oder ob Ziele, Methoden oder Materialien angepasst werden müssen.

Diese Perspektive verschiebt den Fokus sonderpädagogischer Diagnostik. Neben der umfassenden Einzelfalldiagnostik, die Anamnese, standardisierte Testverfahren, Profilanalysen, Normvergleiche und Gutachtererstellung umfassen kann (Grünke, 2012; Heimlich et al., 2021), gewinnt eine klassenraumbezogene Diagnostik (*classroom-based assessment*) an Bedeutung. Classroom-based Assessment unterstützt Lehrkräfte dabei, Lernentwicklungen im Unterricht fortlaufend zu beobachten und Förderentscheidungen datenbasiert zu treffen (Brookhard & McMillan, 2020; Huber & Grosche, 2012; Jungjohann & Gebhardt, 2023). Im Zentrum stehen kurze, wiederholte und unterrichtsnah einsetzbare Messungen, mit deren Hilfe regelmäßig überprüft wird, ob die aktuellen Unterrichts- oder Fördersettings lernwirksam sind für das angestrebte Lernziel der Lernenden (Fuchs & Fuchs, 2006; Filderman et al., 2018).

Damit knüpft classroom-based Assessment an eine Entwicklung an, in der sonderpädagogische Diagnostik zunehmend als zyklischer Prozess verstanden wird. Frühere Konzepte beschrieben diagnostisches Handeln häufig in einer linearen Abfolge aus Diagnostik, Intervention und Evaluation (z. B. Heimlich, 2012; Bundschuh und Winkler, 2019). Neuere Ansätze betonen dagegen zyklische Modelle, in denen diagnostische Fragestellungen, Datenerhebung, Interpretation, Förderplanung und Evaluation fortlaufend und iterativ miteinander verbunden sind. Diese Entwicklung zeigt sich beispielsweise im Diagnostik-Förder-Kreislauf (Tönnissen et al., 2024), im GRID-Modell (Hagen & Hammes-Schmitz, 2025) oder in inklusiv fokussierten Modellen, die Statusdiagnostik, Lernverlaufsdagnostik und adaptive Förderung systematisch aufeinander beziehen. (z. B. Wilbert, 2014; Gebhardt & Jungjohann, 2020). Wie solche Ansätze praktisch umgesetzt werden können, zeigen Modellprojekte wie das Rügener Inklusionsmodell (Hartke, 2017; Voß et al., 2014) oder das Projekt *multimo* (Nitz et al., 2023). Eine breite Implementierung außerhalb solcher Modellregionen ist bislang nicht absehbar, aber aus sonderpädagogischer Sicht wünschenswert.

Zugleich verdeutlichen empirische Befunde, dass die förderbezogene Ausrichtung diagnostischen Handelns in der Praxis nicht automatisch und selbstverständlich gelingt. Zum einen sind sonderpädagogische Gutachten häufig nicht so differenziert und förderorientiert ausgestaltet, wie es für eine geeignete Förderpraxis erforderlich wäre (Moser Opitz et al., 2019). Das zeigt sich auch darin, dass die Abgrenzung zwischen den sonderpädagogischen Förderschwerpunkten oftmals nur eine geringe Trennschärfe aufweist und demnach nicht auf konkrete Förderung im jeweiligen Schwerpunkt fokussiert ist (Weber et al., 2025).

Zum anderen zeigen Lehrkräfte bei classroom-based Assessment eine eingeschränkte Urteilsgenauigkeit (Ebenbeck et al., 2026). Sie orientieren sich bei der Einteilung der Klasse in Fördergruppen intuitiv in den meisten Fällen nicht an standardisierten Testergebnissen, sondern stützen ihre Entscheidungen stattdessen auf subjektive Eindrücke und sozialen Vergleich (Ebenbeck et al., 2026). Diese Urteilsgenauigkeit ist zwar ausreichend für reguläre Differenzierung im Klassenzimmer (Südkamp et al., 2012), aber nicht für diagnostische Prozesse (Gebhardt, 2024).

Daraus ergibt sich ein Professionalisierungsbedarf, der über die Kenntnis einzelner Testverfahren hinausgeht. Lehrkräfte müssen lernen, diagnostische Informationen systematisch zu erheben, kritisch zu interpretieren und in adaptive Förderung zu überführen.

1.1 Formatives Fördern durch Lernverlaufsdagnostik

Eine Möglichkeit, diesen Anspruch umzusetzen, bietet Lernverlaufsdagnostik (Klauer, 2006). Bei Lernverlaufsdagnostik werden kurze und häufig skill-basierte Tests hochfrequent während mehrwöchiger Förderphasen eingesetzt, um die individuelle Lernentwicklung zu messen.

Lernverlaufsdiagnostik ist in Deutschland erst seit wenigen Jahren ein eigenständiges Feld und wird sukzessive in die universitäre Lehre integriert (Jungjohann et al., 2018). Ihre praktische Umsetzung setzt allerdings diagnostische Kompetenz und damit einhergehendes testtheoretisches Grundwissen und Data Literacy voraus (Jungjohann, 2019). Lehrkräfte müssen geeignete Messinstrumente in Abhängigkeit von der jeweiligen Fragestellung auswählen, Messwerte interpretieren, Lernentwicklungen einschätzen und auf dieser Grundlage Förderentscheidungen treffen können (Gebhardt, 2024; Espin et al., 2018; Jungjohann et al., 2018).

Digitale Lernverlaufsverfahren wie Levumi (www.levumi.de; Gebhardt et al., 2016) können diesen Prozess unterstützen, indem sie den diagnostischen Prozess und die Benutzerfreundlichkeit der Verfahren erleichtern, z. B. durch den Einsatz digitaler Endgeräte, Automatisierung, Datenmanagement und Algorithmik. Levumi setzt diesen Anspruch aufseiten der Lehrkraft z. B. so um, dass Lernende in digitalen Klassenbüchern organisiert sind und die Lehrkraft qualitative und quantitative Testergebnisse automatisiert in tabellarischer und grafisch aufbereiteter Form nach jeder Testdurchführung erhält (Jungjohann et al., 2018).

Die grafische Darstellung des Lernverlaufs eines Individuums unterstützt die schnelle Übersicht über Leistungstrends, Lernzuwächse oder stagnierende Entwicklungen. Allerdings zeigen empirische Befunde, dass die reine Verfügbarkeit solcher Daten nicht automatisch zu einer Verbesserung diagnostischer Entscheidungen führt (Espin et al., 2018). Entscheidend ist vielmehr, ob Lehrkräfte Daten theoriegeleitet deuten, mit konkreten Lernzielen verknüpfen und in begründete pädagogische Handlungen umsetzen können.

1.2 Datenbasierte Förderentscheidungen als universitäre Core Practice

Aufgrund ihrer hohen Praxisrelevanz und fachlichen Komplexität lassen sich datenbasierte Förderentscheidungen als Core Practice sonderpädagogischen Handelns verstehen. Core Practices (Grossman et al., 2009) bezeichnen in der universitären Lehre ein begrenztes Set an zentralen, hochwirksamen Handlungselementen, Lehrbausteinen und Kompetenzen, die in der Lehrkräftebildung gezielt angebahnt, geübt und reflektiert werden sollen. Für angehende Sonderpädagog:innen im sonderpädagogischen Schwerpunkt Lernen gehört dazu neben Klassenführung insbesondere in inklusiven Kontexten die Gestaltung und Verbindung von flexiblen Lernumgebungen und Unterstützungssystemen, insbesondere im Kontext Teamteaching und geöffneter inklusiver Unterricht, sowie das damit einhergehende diagnostische Erfassen von Lernprozessen, das Ableiten von Förderzielen, die Auswahl von Fördermaßnahmen und deren Evaluation (Gebhardt, 2024).

Für die universitäre Lehre folgt daraus, dass Studierende nicht nur Wissen über diagnostische Verfahren erwerben sollen. Sie müssen vielmehr im Studium und Praktikum lernen und üben, diagnostische Daten in konkreten Fällen zu interpretieren und daraus begründete Förderentscheidungen abzuleiten (Zellner et al., 2025a). Dies erfordert Lerngelegenheiten, in denen Studierende mit realistischen oder realen Lernverläufen arbeiten, Entscheidungsalternativen abwägen und die Passung von Fördermaßnahmen reflektieren. Fallbasierte, digitale und simulationsgestützte Formate bieten hierfür wichtige Zugänge, z. B. in Verbindung mit videobasierten Mikrointerventionen (Jungjohann et al., 2022) und simulationsgestützten Fallanalysen (Zellner & Gebhardt, 2025; Zellner et al., 2025b). Gleichzeitig bleibt der Transfer in reale schulische Kontexte eine besondere Herausforderung, weswegen das Praktikum ein zentraler Lernort für formatives Fördern ist.

1.3 Das sonderpädagogische Praktikum als Lernort formativer Förderdiagnostik

Die Forschung zu schulischen Praktika ist bislang vorwiegend in der Schulpädagogik und den Fachdidaktiken verortet. Diese konzentrieren sich auf die allgemeine Lehrkräftebildung und die Erfüllung von KMK-Standards (2004/2019) (Porsch & Gollub, 2021). Im Bereich der Sonderpädagogik gibt es zwar Studien zur aktuellen Lage und auch Modelle zur Vermittlung (Hartke, 2017), jedoch liegen bislang nur wenige empirisch fundierte Ansätze zur systematischen Umsetzung sonderpädagogischer Diagnostik und Förderung im Studium vor. Erste Arbeiten zeigen jedoch, dass sich entsprechende Konzepte im Rahmen von Praktika implementieren und über Einzelfallforschung empirisch begleiten lassen (Hövel, 2025).

In der bayerischen Lehrkräfteausbildung ermöglicht das studienbegleitende sonderpädagogische Praktikum (Bayerisches Staatsministerium für Unterricht und Kultus, 2008) Studierenden, eine Praktikumsklasse über ein gesamtes Schuljahr hinweg immer einen Tag pro Woche zu begleiten. Das Praktikum findet in sonderpädagogischen Förderzentren oder inklusiven Schulen statt und wird seitens der Universität jedes Semester von einem Seminar begleitet. Auf diese Weise ermöglicht es Studierenden über zwei Semester förderdiagnostisches Arbeiten (Teil I) und Erproben des eigenen Unterrichts (Teil II). Bisher lag vonseiten der Schulen der Fokus auf dem Erproben von möglichst vielen Unterrichtsstunden und dem Erstellen eines sonderpädagogischen Gutachtens. Nun liegt der Fokus neben dem Erproben von wenigen Unterrichtsstunden vorwiegend auf der Begleitung des langfristigen Lernprozesses einzelner Schüler:innen in einem Lernbereich (Capovilla et al., 2023). Ein Großteil der Praktikumslehrkräfte begleitete das Praktikum bereits vor dieser konzeptionellen Neuausrichtung, sodass das neue hochschulische Design in bestehende schulische Praxisstrukturen implementiert wurde, die weiterhin stark auf Unterrichtsversuche hin ausgerichtet waren. Ziel ist es, im Sinne des Response-to-Intervention-Ansatzes zu begleiten, zu fördern und formativ zu evaluieren und zu ermitteln, welche Fördermethoden und -materialien Lernen unterstützen.

Um dieses Ziel zu adressieren, wird der Ansatz der Design-Based Research (DBR) gewählt. DBR ist ein methodologischer Rahmen, der die theoriegeleitete Entwicklung, Erprobung und Weiterentwicklung von Interventionen unter realen Praxisbedingungen ermöglicht (Geßler, 2024). Charakteristisch für DBR ist ein iterativer Ablauf, der typischerweise mehrere Zyklen aus (1) theoretischer Fundierung und Design, (2) Implementation im Praxisfeld, (3) systematischer Analyse von Lern- und Entscheidungsprozessen sowie (4) Revision und Weiterentwicklung des Designs umfasst (Geßler, 2024). Ziel ist dabei nicht die Wirksamkeitsprüfung eines fertigen Produkts, sondern das vertiefte Verständnis von Lern- und Entwicklungsprozessen sowie die Generierung gestaltungsrelevanter Theorie.

Im deutschsprachigen Raum wurde DBR bislang insbesondere in der Mathematikdidaktik, etwa in der Arbeitsgruppe um Prediger (Abraham & Prediger, 2025), sowie im Sachunterricht (Schroeder et al., 2025) systematisch eingesetzt. Diese Forschungsarbeiten zeigen, dass DBR besonders geeignet ist, komplexe Lernprozesse in authentischen Settings zu analysieren und theoriegeleitet weiterzuentwickeln. International wird DBR zudem in den Naturwissenschaftsdidaktiken genutzt, um Entscheidungsprozesse von Lernenden systematisch zu erfassen und Interventionen entlang identifizierter Kompetenzdiskrepanzen zu optimieren (Frye et al., 2020; Obczovsky et al., 2023).

Für die Begleitung von Praktika bietet DBR mehrere Vorteile: Zum einen erlaubt der Ansatz die enge Verzahnung von universitärer Lehre und schulischer Praxis, indem Studierende und Lehrkräfte als aktive Beteiligte in den Entwicklungsprozess einbezogen werden. Zum anderen ermöglicht DBR die Analyse diagnostischer Entscheidungsprozesse über wiederholte Mess- und Reflexionsschleifen hinweg, was insbesondere für die Einübung sonderpädagogischer Core Practices relevant ist. Gleichzeitig ist DBR mit methodologischen Herausforderungen verbunden, etwa einer begrenzten Generalisierbarkeit einzelner Zyklen und einem erhöhten Aufwand in Planung, Datenerhebung und Auswertung (Geßler, 2024). Diese Einschränkungen werden im vorliegenden Beitrag bewusst in Kauf genommen, um die komplexen Bedingungen sonderpädagogischer Ausbildung angemessen abzubilden.

In Anlehnung an kontrollierte Einzelfallforschung werden im Rahmen des Praktikums diagnostische Verfahren eingesetzt, Fördermaßnahmen geplant und deren Wirksamkeit reflektiert. Die systematische Einübung und empirische Erfassung dieser Core Practices erfolgt dabei eingebettet in den DBR-Zyklus, um das Lehr-Lern-Arrangement theoriegeleitet weiterzuentwickeln und zugleich praxisnah zu gestalten.

2 FRAGESTELLUNG

Ziel der sonderpädagogischen Hochschullehre im Schwerpunkt Lernen im Bereich des formativen Assessments und der Förderung ist es, angehende Lehrkräfte zu befähigen, diagnostische Entscheidungen und Förderentscheidungen datenbasiert zu treffen. Obwohl die Bedeutung von Lernverlaufsdiagnostik für *struggling students* wissenschaftlich belegt ist (z. B. Fuchs et al., 2025), fehlt es an empirisch geprüften, praxisnahen Lehr-Lern-Formaten, die Studierende systematisch auf diese Form der Diagnostik vorbereiten. Hier setzt diese Studie an, um die Wirksamkeit des Ansatzes von Förderung und lernbegleitender Diagnostik in der Hochschullehre für Studierende der Sonderpädagogik zu betrachten.

Im Fokus des Projekts steht die Frage, wie Studierende zentrale Core Practices der formativen Diagnostik und Förderung im sonderpädagogischen Praktikum umsetzen. Die Studierenden erhalten hier-

für universitäre Vorgaben und Materialien, müssen diese jedoch unter den konkreten Bedingungen der Praktikumschule anwenden, anpassen und begründen. In Anlehnung an kontrollierte Einzelfallforschung setzen sie diagnostische Verfahren ein, planen Fördermaßnahmen und reflektieren deren Passung und Wirksamkeit anhand der erhobenen Daten. Der DBR-Zyklus bezieht sich dabei auf die systematische Entwicklung, Erprobung und Weiterentwicklung des hochschuldidaktischen Arrangements: Untersucht wird, wie die vorbereiteten Lehrbausteine in der Praxis umgesetzt werden, welche Schwierigkeiten bei der Implementation auftreten und wie das Format auf dieser Grundlage theoriegeleitet und praxisnah weiterentwickelt werden kann.

Die Fragestellungen lauten:

1. Wie verändert sich die Selbsteinschätzung der diagnostischen Kompetenzen der Studierenden zwischen Beginn und Ende des Praktikums, gemessen mit einem standardisierten Selbsteinschätzungsinstrument?
2. Wie wird die Förderung von Studierenden im Praktikum umgesetzt?
3. Welche Lernentwicklung zeigen die Ergebnisse der Lernverlaufsdiagnostik über den Praktikumszeitraum hinweg?

3 METHODE

3.1 Durchführung

Grundlage dieser Studie bildet die Evaluation der durchgeführten Förderdiagnostik von Studierenden im Rahmen ihres Praktikums an einer Hochschule in Bayern, Deutschland.

Die Studierenden besuchen im Rahmen ihres Studiums mehrere vorbereitende Veranstaltungen zu den Themen Diagnostik und Fördern, welche in Tabelle 1 gelistet sind. Die Lehre ist mit reflexiver Fallarbeit und konkreten Handlungsweisen verbunden, welche in einem Lehrkonzept verschriftlicht sind (Zellner et al., 2025a).

Die Studierenden haben die Vorgabe, im studienbegleitenden sonderpädagogischen Praktikum Teil I eine Lernumgebung für ein bis drei Schüler:innen über eine längere Zeit (min. 5 Interventionseinheiten) zu gestalten und den Kompetenzerwerb der Schüler:innen als langfristigen Prozess zu evaluieren. Daher sollte zunächst in den jeweiligen Praktikumsklassen sowohl in inklusiven Schulen als auch an sonder-

TABELLE 1: Übersicht über vorbereitende und praktikumsbegleitende Veranstaltungen im Studium

Semester	Universitäre Veranstaltungen	Praktikum
3	Vier Seminare zur Didaktik und Lernförderung in Deutsch und Mathematik	
4	Vorlesung zu Diagnostik und Testtheorie	
	Praxisseminar zu Status- und Lernverlaufsdiagnostik und Testtheorie	
5	Praktische Übung zur Förderplanung bei Lernschwierigkeiten	Praktikum Teil I – Fokus Diagnostik und Fördern
	Praktikumsbegleitseminar Teil I mit Fokus auf Diagnostik und Fördern	
6	Praktikumsbegleitseminar Teil II mit Fokus auf Didaktik und Unterrichten	Praktikum Teil II – Fokus Didaktik und Unterrichten

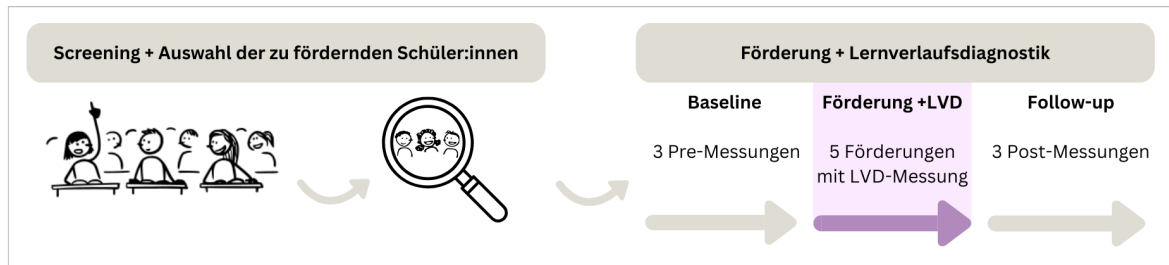


ABBILDUNG 1: Studiendesign

pädagogischen Förderzentren, ein Lesescreening durchgeführt und die Ergebnisse sollten gemeinsam mit der Klassenlehrkraft besprochen werden. Auf dieser Grundlage wählten die Studierenden in Absprache mit der Lehrkraft Schüler:innen mit Förderbedarf im Bereich Lesen aus. Für diese Schüler:innen planten sie eine individualisierte, evidenzbasierte Förderung im Bereich Lesen oder der phonologischen Bewusstheit und setzten diese über mindestens fünf aufeinanderfolgende Wochen um.

Die Förderung wurde durch regelmäßige Lernverlaufsdiagnostik auf der Plattform Levumi begleitet. Die Vorgabe war es, drei Baseline-Messungen, fünf Fördereinheiten in fünf aufeinanderfolgenden Wochen mit paralleler Testung sowie drei Follow-up-Messungen pro Schüler:in durchzuführen (Tabelle 1). Abweichungen von diesem Schema waren in Absprache möglich, wenn etwa Schüler:innen die Klasse vorzeitig verließen oder organisatorische Einschränkungen bestanden (z. B., wenn die Klasse für zwei Wochen im Praktikum war). Die Studierenden verfassten nach dem Praktikum einen Bericht, welcher die Grundlage für die folgenden Analysen bildet. Weitere Informationen zur Lehre, dem Material und Instrumenten können via Zenodo (<https://doi.org/10.5281/zenodo.23098048>) abgerufen werden.

Der vorliegende Beitrag bildet einen ersten Zyklus im Sinne des DBR ab. In der Designphase wurde das Lehrkonzept zur Verbindung von Diagnostik, Lernverlaufsdiagnostik und Förderung theoriegeleitet entwickelt und in universitäre Lehrveranstaltungen integriert. In der Implementationsphase setzten Studierende dieses Konzept im Rahmen ihres schulischen Praktikums um, indem sie diagnostische Verfahren durchführten, Fördermaßnahmen planten und Lernverläufe systematisch erhoben. Die Analysephase umfasste sowohl die Auswertung der Lernverlaufsdaten der Schüler:innen als auch die Selbsteinschätzungen der Studierenden sowie die in den Berichten dokumentierten Entscheidungsprozesse. Auf dieser Grundlage werden im Sinne der Revisionsphase zentrale Anpassungsbedarfe identifiziert, die in nachfolgenden DBR-Zyklen zur Weiterentwicklung des Lehr-Lern-Arrangements genutzt werden.

3.2 Stichprobe

In die Analyse der anonymisierten Berichte wurden nur Fälle einbezogen, bei welchen mindestens zwei Baseline-, vier Interventions- und zwei Follow-up-Messzeitpunkte vorliegen sowie das formative Assessment mit Tests von Levumi durchgeführt wurde. Daher wurden von insgesamt $N = 32$ Studierenden mit $N = 97$ geförderten Schüler:innen, nur

$N = 18$ Studierende (3 männlich, 15 weiblich; $M_{\text{Alter}} = 25,75$, $SD_{\text{Alter}} = 8,99$) verglichen, die an inklusiven Schulen oder an sonderpädagogischen Förderzentren ein Praktikum in den Jahrgangsstufen 1–8 absolvierten. Alle Teilnehmenden befanden sich im fünften Semester ihres Lehramtsstudiums. $N = 43$ Schüler:innen wurden von den Studierenden gefördert ($M = 2,93$, $SD = 0,26$ Schüler:innen pro Studierende:r), wovon 52,3 % männlich und 47,7 % weiblich waren. Das Alter der Schüler:innen lag im Mittel bei $M = 11,10$ Jahren ($SD = 2,40$, $Min = 7$, $Max = 16$).

3.3 Instrumente

Die Studierenden bearbeiteten vor und nach der Förderung einen Fragebogen zur Selbsteinschätzung der eigenen diagnostischen Kompetenz von Lehrkräften in der inklusiven Schule (DaKI, Version 0.2; Jungjohann & Gebhardt, 2023). DaKI misst die vier Dimensionen (1) fachdidaktische Diagnostikkompetenzen, (2) Schulleistungs- und Statusdiagnostik, (3) sonderpädagogische Feststell- und Förderdiagnostik und (4) Lernverlaufsdiagnostik anhand einer fünfstufigen Likert-Skala (stimme überhaupt nicht zu – stimme vollständig zu; Jungjohann & Gebhardt, 2023). Die individuellen Skalen weisen in der Validierungsstudie eine gute interne Kohärenz auf ($\alpha = .92, .82, .92, .93$).

3.4 Analyse

Zur Analyse der Veränderung in der Selbsteinschätzung der Studierenden wurde zunächst die Differenzverteilung auf Normalität geprüft (Shapiro-Wilk-Test) und ein t-Test für verbundene Stichproben durchgeführt. Für signifikante Ergebnisse wurde die Effektstärke nach Cohen's d berechnet; zusätzlich wurden Effektstärken auch bei nicht signifikanten Ergebnissen berichtet, um die Größe und potenzielle praktische Relevanz von Effekten unabhängig von der Stichprobengröße einschätzen zu können (Cano-Corres et al., 2012).

Die Umsetzung der formativen Förderung und Diagnostik wird deskriptiv dargestellt. Die Erhebung und Auswertung der Lernverläufe orientierte sich an zentralen Qualitätskriterien der WWC-Standards (What Works Clearinghouse) für Single-Case-Designs, insbesondere an der Phasenlogik sowie an wiederholten Messungen innerhalb der Baseline-, Interventions- und Follow-up-Phase (Kratochwill et al., 2010). Aufgrund der pragmatischen Umsetzung im Praktikumskontext und unvermeidbarer organisatorischer Abweichungen im Feld wurden jedoch nur Fälle in die Analyse einbezogen, die mindestens zwei Baseline-, vier Interventions- und zwei Follow-up-Messzeitpunkte aufwiesen; entsprechend werden die Einzelfalldaten nicht als streng experimenteller Wirksamkeitsnachweis, sondern als deskriptive und analytische Beschreibung von Lernverläufen im Rahmen datenbasierter Förderpraxis interpretiert. Auf Grundlage der Berichte wurde die Auswahl der geförderten Schüler:innen dichotom kodiert, abhängig davon, ob die Auswahl explizit auf die Ergebnisse des durchgeführten Screenings Bezug nahm oder primär auf nicht datenbasierten Begründungen beruhte. Die Einzelfalldaten wurden über alle Fälle hinweg aggregiert, da die Lernverläufe durchgängig mit demselben skillbasierten Lernverlaufsdiagnostik-Instrument im Lernbereich Lesen erhoben wurden und alle Studierenden auf Basis identischer Arbeitsanweisungen vorgingen, wodurch trotz unterschiedlicher Förderansätze eine vergleichbare Analyse der Lernentwicklungen möglich war. Zur Analyse der individuellen Lernverläufe der Schüler:innen wurden Differenzen der Prozent- und Rohwerte zwischen den Medianen der Baseline- und Follow-up-Messungen berechnet und deskriptiv ausgewertet. Ein Wilcoxon-Test überprüfte die Signifikanz der Differenzen.

Ergänzend wurden Nonoverlap-Indizes (PND, PEM, PAND) für den Vergleich der Phasen Baseline und Follow-up berechnet. PND (Percentage of Nonoverlapping Data) gibt den Anteil der Follow-up-Ergebnisse an, die jenen der Baseline-Messungen nicht überlappen. PND gibt somit an, wie viele Ergebnisse der Follow-up-Messungen über dem höchsten Wert der Baseline-Messungen liegen. PEM (Percentage Exceeding the Median) nutzt statt des Extremwerts den Median der Baseline und gibt an, wie viele Ergebnisse der Follow-up-Messungen darüber liegen, wodurch Ausreißer weniger starken Einfluss haben. PAND (Percentage of All Nonoverlapping Data) berücksichtigt den gesamten Datensatz und entfernt minimal viele Datenpunkte, um vollständige Trennung (keine Überlappung) zu erzielen. Der Anteil der verbleibenden Werte liefert den PAND-Wert. Je höher und näher an 100 % diese Werte sind, desto klarer trennen sich Follow-up- und Baseline-Phase. Das spricht für einen starken Interventions- oder Veränderungseffekt (Parker et al., 2011). Zur differenzierten Einzelfallanalyse wurde zudem der Tau-U-Test (Parker & Vannest, 2009) eingesetzt, um Effektstärke und Signifikanz unter Kontrolle von Trends abzuschätzen. Alle Analysen erfolgten mit dem Paket scan (Wilbert & Lüke, 2019) in R und der R-Studio-Umgebung (R Core Team, 2018).

Die eingesetzten Interventionen zur Förderung der Schüler:innen wurden entlang einer dreistufigen Skala kodiert: (1) gar nicht evidenzbasiert, (2) teilweise evidenzbasiert und (3) vollständig evidenzbasiert. Um die Effektivität von Maßnahmen beurteilen zu können, kategorisiert das Oxford Centre for Evidence-Based Medicine (2011) unterschiedliche Qualitätsstufen der Evidenz von deskriptiven Studien zu vollständig randomisiertem Kontrollgruppendesign oder Meta-Analysen. Casale et al. (2014) weisen allerdings darauf hin, dass etwa vollständig randomisierte Kontrollgruppendesigns in der pädagogischen Feldforschung nur eingeschränkt umgesetzt werden können. Dementsprechend wird in der vorliegenden Analyse nur überprüft, ob evidenzbasierte Maßnahmen eingesetzt wurden, und nicht, auf welcher Qualitätsstufe die Maßnahmen eingeordnet werden.

Die vorliegende Studie fokussiert somit im DBR-Zyklus nicht allein die Wirksamkeit einzelner Fördermaßnahmen, sondern die Implementation der universitären Lerngelegenheiten im Praktikum. Analysiert wird, inwiefern Studierende formative Diagnostik, Förderplanung und Evaluation als zusammenhängende Handlungspraxis realisieren und welche Bedingungen diese Umsetzung unterstützen oder erschweren.

4 ERGEBNISSE

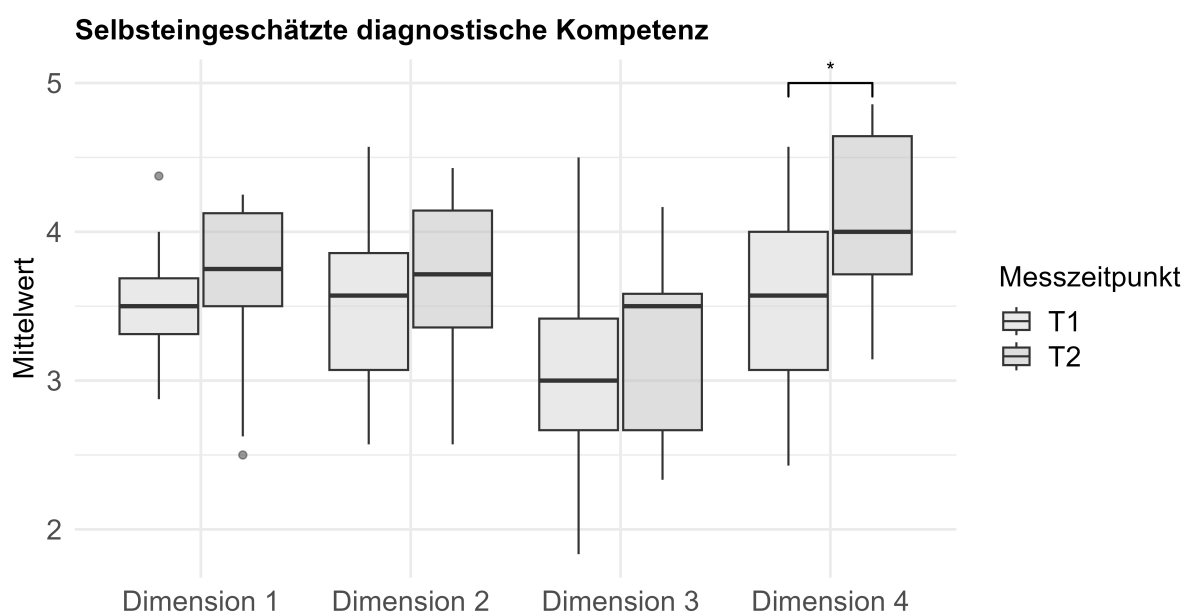
4.1 Selbsteinschätzung der Studierenden

Die Studierenden schätzten ihre diagnostischen Kompetenzen zu Semesterbeginn mit $M = 3,39$ (Range = 2,79–4,50) und zu Semesterende mit $M = 3,69$ (Range = 3,00–4,29) ein. Die Prüfung der Differenzen zeigte keine signifikante Abweichung von der Normalverteilung, $W = 0,88$, $p = ,13$, deshalb wurde ein t-Test angeschlossen. Der gepaarte t-Test zeigte eine signifikante Verbesserung der Gesamteinschätzung von Messzeitpunkt 1 zu Messzeitpunkt 2, $t(9) = -2,30$, $p < .05$, 95 %-CI $[-0,53, -0,00]$. Abbildung 2 zeigt die Verbesserung in allen vier Dimensionen des DaKI. Von den Verbesserungen war nur die Dimension Lernverlaufsdiagnostik mit einem großen Effekt signifikant ($t(9) = -3,84$, $p < .05$, 95 %-CI $[-0,95, -0,25]$). Die anderen Veränderungen waren nur klein bis mittel ($d_1 = 0,13$, $d_2 = 0,51$, $d_3 = 0,39$, $d_{\text{ges}} = 0,73$).

4.2 Umsetzung der Förderung durch die Studierenden

Für die Auswahl der zu fördernden Schüler:innen führten die Studierenden zunächst ein Screening in den Klassen durch. Hierbei kamen unterschiedliche Testverfahren zum Einsatz: Am häufigsten wurde die *ELFE II* (Ein Leseverständnistest für Erst- bis Siebtklässler – Version II; Lenhard & Schneider, 2017; $N = 24$) verwendet, gefolgt von digitalen *Levumi-Screenings* (Ebenbeck et al., 2023; $N = 10$), *WLLP* (Würzburger Leise Leseprobe; Küspert & Schneider, 1998; $N = 3$) und *LGVT 5–12+* (Lesegeschwindigkeits- und Verständnistest für die Klassen 5–12; Schneider et al., 2017; $N = 3$). Zwei Praktikumslehrkräfte äußerten gegenüber den Studierenden große Bedenken bezüglich der Sinnhaftigkeit eines Screenings in ihrer Klasse. Begründet wurden die Bedenken damit, dass die Schüler:innen der Klasse bereits im Schuljahr zuvor ein anderes Screening im Fach Deutsch durchgeführt hatten und dies als Belastung für die Schüler:innen der Klasse wahrgenommen wurde. Die Studierenden durften in gemeinsamer Absprache mit der Lehrkraft das Screening trotzdem durchführen.

Die Studierenden wählten die zu fördernden Schüler:innen zu 33,93 % auf Basis oder mit Bezug zu den Daten der Screening-Ergebnisse aus (z. B. „ähnliches Leistungsniveau, gute Kooperation, ge-



Note: * $p < .01$. Dimension 1: fachdidaktische Diagnostikkompetenzen, Dimension 2: Schulleistungs- und Statusdiagnostik, Dimension 3: sonderpädagogische Feststell- und Förderdiagnostik, Dimension 4: Lernverlaufsdiagnostik.

ABBILDUNG 2: Selbsteinschätzung der diagnostischen Kompetenz der Teilnehmenden zu Beginn (T1) und am Ende (T2) der Förderung

ringere Punktzahl im sinnkonstruierenden Satzlesen“ oder „schlechteste Schüler:innen der Klasse laut Screening und Wunsch der Eltern“). Die restlichen 66,07 % der Schüler:innen wurden auf Anraten der Praktikumslehrkraft als *struggling students* identifiziert und daher für die Förderung gewählt (z. B. „Die Praktikumslehrerin sieht bei diesen Schüler:innen großen Förderbedarf.“). Teilweise wurden auch Diskrepanzen zwischen Wort- und Satzebene, spezifische Auffälligkeiten im Bereich der phonologischen Bewusstheit oder der Sinnkonstruktion sowie die Kooperationsfähigkeit bei der Auswahl der Schüler:innen berücksichtigt.

Die Förderung selbst fand innerhalb eines Semesters statt. Die Studierenden berichteten, dass die Förderung zumeist wöchentlich im Rahmen von etwa 40-minütigen Einheiten durchgeführt wurde ($M = 40,79$ Minuten, $SD = 10,64$). Die Förderung wurde dabei zu 89,28 % außerhalb des Klassenzimmers durchgeführt (z. B. im Nebenraum), in Ausnahmen auch als Wochenplanarbeit (ein:e Studierende:r mit drei Schüler:innen) oder in ritualisierter Form zu Beginn des Klassenunterrichts (ein:e Studierende:r mit drei Schüler:innen).

Die Studierenden berichteten, dass 7,90 % der eingesetzten Interventionen vollständig evidenzbasierte Verfahren umfassten (z. B. das Programm *FILBY*, welches die Praktikumslehrkraft im regulären Klassenunterricht nutzt), zu 52,63 % Fördermaßnahmen, die zumindest teilweise evidenzbasiert sind (z. B. einzelne Leseübungen aus dem Fördermaterial von Levumi), und zu 39,47 % Fördermaßnahmen ohne Evidenzbasierung. Dies waren vorwiegend Lesestrategien, Übungen zur phonologischen Bewusstheit sowie unterschiedliche methodische Kombinationen wie Arbeitsblätter, Lesespurgeschichten, Lese-Dominos oder Textarbeit. Die Auswahl der Materialien orientierte sich an den jeweiligen Förderschwerpunkten der Schüler:innen und variierte zwischen einzelnen Klassen.

4.3 Umsetzung der Lernverlaufsdiagnostik durch die Studierenden

Anschließend wählten die Studierenden ein passendes Instrument für die Lernverlaufsdiagnostik. Am häufigsten kam der *Levumi-Test zum sinnkonstruierenden Satzlesen* (Jungjohann & Gebhardt, 2019; 67,44 %) zum Einsatz. Weitere eingesetzte Instrumente waren der *Levumi-Test Blitzlesen* (Zellner et al., 2024; 20,93 %), der *Levumi-Test zum Sichtwortschatz* (Jungjohann & Gebhardt, 2021; 6,98 %) und der *Levumi-Test zum Lesen von Pseudowörtern* (Jungjohann et al., 2021; 4,65 %).

Im Durchschnitt wurden von den Studierenden pro Schüler:in 2,7 Baseline-Messungen ($SD = 0,46$), 4,86 Förderungen mit paralleler Testung ($SD = 0,82$) sowie 2,58 Follow-up-Messungen ($SD = 0,64$) durchgeführt. Die Ergebnisse der Lernverlaufsdiagnostik wurden je nach Verfahren entweder als Rohwerte ($N = 38$) oder in Prozent ($N = 5$) angegeben.

Die Förderung durch die Studierenden wurde mittels Lernverlaufsdiagnostik evaluiert. Als Hinweis dafür, dass das Förderkonzept wirksam ist, wurden für die Lernentwicklung der Schüler:innen Steigungswerte der Ergebnisse in Prozent ($n = 5$) und Rohwerten ($n = 38$) zwischen der Baseline- und der Follow-up-Messung berechnet. Die Steigungswerte auf Prozentbasis lagen im Mittel bei $M = 24,49$ Prozentpunkten ($SD = 17,48$, $Mdn = 30,5$, $Min = 1,75$, $Max = 41,05$). Für die Steigungswerte auf Rohwertbasis ergab sich ein Mittelwert von $M = 6,96$ ($SD = 7,77$, $Mdn = 5,33$, $Min = -3,5$, $Max = 35$). Ein Wilcoxon-Test für verbundene Stichproben zeigte, dass sich die Medianwerte der Baseline- zur Follow-up-Messung signifikant verbesserten ($V = 48,5$, $p < .001$).

Die Overlap-Analyse ergab für den Vergleich der Baseline- mit der Follow-up-Phase (A–C) einen mittleren PND von 76,7 %, PEM von 83,4 % und PAND von 84,3 %. Werte zwischen 70 und 90 % des PND (Scruggs & Mastropieri, 1998), PEM (Ma, 2006) und PAND (Casale et al., 2017) weisen auf eine moderate Wirksamkeit der Intervention zwischen Baseline- und Follow-up-Messung hin. Insbesondere die Verfahren PEM und PAND, welche gegenüber dem PND robuster gegen Ausreißer sind (Parker et al., 2011), deuten auf eine Verbesserung der Lernentwicklung durch die Unterstützung der Studierenden hin.

Der Vergleich zwischen Phase A und Phase C mittels Tau-U-Test ergab in 35 % der Fälle einen signifikant positiven Effekt zugunsten der Phase C ($\tau = .75-.89$, $p < .05$). In weiteren Fällen lagen positive Tau-U-Werte vor, erreichten jedoch keine statistische Signifikanz (z. B. Schüler mit ID 19: $\tau = .62$, $p = .10$). Negative oder gegenläufige Effekte traten nicht signifikant auf. Insgesamt deuten die Ergebnisse auf wirksame Fördereffekte hin (Parker & Vannest, 2009).

Obwohl die aggregierten Overlap-Ergebnisse für den Vergleich zwischen Baseline und Follow-up insgesamt auf substanzielle Leistungszuwächse hinweisen, zeigten drei Einzelfälle keine oder sogar gegenläufige Entwicklungen. Insbesondere bei den Schüler:innen mit ID 22, 80 und 81 ergaben sich

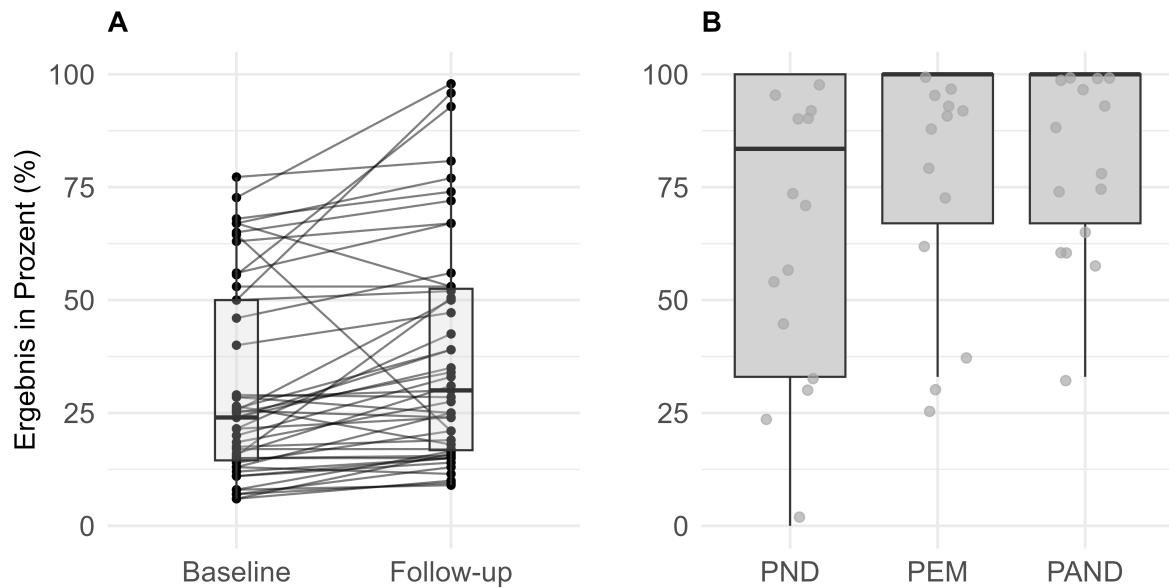


ABBILDUNG 3: (A) Lernentwicklung der Schüler:innen und (B) Effektstärke der Leistungssteigerung

negative Mittelwertsdifferenzen zwischen Baseline und Follow-up (Fall 22: $\Delta = -2,00$ Items; Fall 80: $\Delta = -2,33$ Items; Fall 81: $\Delta = 0,00$ Items). In diesen Fällen lagen auch die PND-Werte unter 35 % (Fall 22: 0 %; Fall 80: 33 %; Fall 81: 0 %), begleitet von niedrigen Effektstärken (Hedges' $g = -0,30, -0,16$ und $0,00$). Fall 80 fiel zudem durch einen negativen reskalierten NAP-Wert (-22 %) auf, was ebenfalls auf eine mögliche Verschlechterung hindeutet. Die Studierenden begründeten die schwankenden oder sinkenden Lernverläufe dieser drei Schüler:innen mit flüchtigem, oberflächlichem Arbeitsverhalten der Schüler:innen oder fehlenden sprachlichen Voraussetzungen.

Zusammenfassend konnte eine Wirksamkeit durch die Förderung der Studierenden festgestellt werden. Dennoch traten Schwierigkeiten auf, welche die Studierenden im Rahmen der Förderung und lernbegleitenden Diagnostik nicht lösen konnten.

5 DISKUSSION

Das vorgestellte Praxisprojekt zur formativen Diagnostik und Förderung durch Studierende unterscheidet sich von Ansätzen, welche aus schulpädagogischer Sicht eine hohe Anzahl von Unterrichtsversuchen in verschiedenen Fächern fokussieren. Stattdessen stand die Förderung einzelner Schüler:innen mit begleitender Lernverlaufsdagnostik im Mittelpunkt (Capovilla et al., 2023), da diese für die Sonderpädagogik als Core Practice (Grossman et al., 2009) in Förderzentren, aber insbesondere in der inklusiven Schule gesehen wird (Gebhardt, 2024). Umgesetzt wird dies an Förderzentren als Erweiterung der bestehenden Praxis. Diese Förderung dauerte im Schnitt rund 40 Minuten und fand zu fast 90 % außerhalb des regulären Unterrichts statt. Die Förderung blieb weitgehend additiv und war nicht in den alltäglichen Klassenunterricht integriert. Nur in wenigen Fällen kam es zu einer Integration der formativen Diagnostik und Förderung in den regulären Klassenunterricht. Aus inklusionspädagogischer Perspektive wäre an dieser Stelle eine engere Verzahnung wünschenswert, da Diagnostik und Förderung so Teil des gemeinsamen Lernprozesses würden (Brookhard & McMillan, 2020; Huber & Grosche, 2012; Jungjohann & Gebhardt, 2023). Ebenso bestanden viele Lehrkräfte auch weiterhin auf einer möglichst hohen Anzahl an verschiedenen Unterrichtsstunden im Praktikum, um eine möglichst hohe Übereinstimmung mit dem zukünftigen Referendariat zu ermöglichen. Aus Perspektive einer Klassenlehrkraft am sonderpädagogischen Förderzentrum ist dies nachvollziehbar. Die Autoren sehen dies dagegen eher kritisch, da der Dreiklang aus Diagnostik, Fördern und Evaluieren ansonsten nicht begleitet und geübt wird. Hinzu kommt, dass staatliche Vorgaben aktuell den Schwerpunkt eher auf summative Diagnostik und Förderplanung legen und Lernverlaufsdagnostik bislang nicht systematisch verankert ist (Jungjohann

& Gebhardt, 2018). Bei der Durchführung zeigten sich zudem Einschränkungen durch organisatorische Rahmenbedingungen. Förderungen fanden oft in der letzten Unterrichtsstunde statt, was den Lernerfolg der Schüler:innen einschränken kann, oder sie wurden zeitlich während der Kernfächer platziert, sodass Schüler:innen dafür im Regelunterricht fehlten. In der Gesamtauswertung konnten nur rund 56 % der teilnehmenden Schüler:innen berücksichtigt werden, da Förderstunden teils krankheitsbedingt ausfielen oder Studierende in ihren Praktikumsklassen mit dem Verfahren nicht zurechtkamen.

Ziel des Projekts ist es, dass Studierende lernen, Förderentscheidungen anhand von Testdaten zu treffen und die Wirksamkeit mithilfe von lernbegleitenden Tests zu evaluieren (Zellner et al., 2025a). Grundlage waren Screeningverfahren und Lernverlaufsdiagnostik, die von allen Studierenden umgesetzt wurden. Dennoch erfolgte die Auswahl der Schüler:innen zur Förderung lediglich zu etwa 33 % mit direktem Bezug zu diesen Daten. In rund 67 % der Fälle entschieden die Praktikumslehrkräfte, welche Kinder teilnehmen sollten, unabhängig von den erhobenen Ergebnissen. Damit wurden die vorgesehenen Daten in der Mehrzahl der Entscheidungen nicht ausschlaggebend genutzt. Hinzu kam, dass einzelne Lehrkräfte Bedenken hinsichtlich der Belastung der Schüler:innen durch die Diagnostik äußerten, wodurch die Umsetzungsmöglichkeiten für die Studierenden teilweise eingeschränkt waren. Ein möglicher Grund dafür könnte sein, dass der formative Förderansatz des Praktikumskonzepts sich von dem bestehenden Unterrichtskonzept der Klassen unterscheidet. Ebenso könnte die Sorge bestanden haben, dass der reguläre Unterricht zu stark beeinträchtigt würde, wenn einzelne Lernende wiederholt für die Förderung fehlen oder wenn dies zu Mehrarbeit führen würde (Moser Opitz et al., 2019). Dies sind jedoch Vermutungen, die sich anhand der vorliegenden Daten nicht klären lassen. Dennoch verdeutlichen die Überlegungen, dass die Implementierung formativer Verfahren ein längerer Prozess ist, der an bestehende Schulstrukturen angepasst werden muss.

Die Analyse der Fördermaßnahmen verdeutlicht, dass die Studierenden nur begrenzt evidenzbasierte Materialien einsetzten und sich stattdessen vermehrt für pädagogische Materialien entschieden. So wurde etwa das Leseförderprogramm FILBY, dessen Nutzbarkeit regelmäßig evaluiert wird (zuletzt Wild et al., 2022), von Studierenden ausgewählt, mit der Begründung, dass die Lehrkraft dies bereits im Klassenunterricht genutzt hatte und dies empfohlen hat. Eine verstärkte sonderpädagogische Expertise und wirksame Programme für die Sonderpädagogik werden daher in Zukunft angestrengt (Hillenbrand et al., 2015; 2018; Hagmann-von Arx et al., 2025). Die Förderungen zeigten insgesamt alle eine Verbesserung. Über alle Fälle hinweg ergaben sich ein mittlerer PND von 76,7 %, ein PEM von 83,4 % und ein PAND von 84,3 %, was auf eine moderate bis gute Lernentwicklung bei den geförderten Schüler:innen hinweist, insbesondere bei den gegen Ausreißer robusten Analysen PEM und PAND.

Die Studierenden selbst gaben an, dass sie nun über bessere diagnostische Kompetenzen verfügen. In der Selbsteinschätzung mit dem Fragebogen DaKI (Jungjohann & Gebhardt, 2023) zeigte sich eine signifikante Verbesserung in der Dimension Lernverlaufsdiagnostik ($p = .004$), verbunden mit einem großen Effekt ($d = 1.21$). Auch in den Gesamtskalen schätzten die Studierenden ihre diagnostischen Kompetenzen insgesamt am Ende des Seminars höher ein. Im Bericht gaben die Studierenden an, dass die Umsetzung von der Interaktion mit den Praktikumslehrkräften geprägt war. Je besser die Möglichkeiten der Studierenden zum selbstständigen Arbeiten und zur Verschränkung von eigener Förderung mit dem Fachunterricht gegeben waren, umso zufriedener beschrieben die Studierenden ihre durchgeführte Förderung und Evaluation. Diese Ergebnisse legen nahe, dass die Wirksamkeit des Designs wesentlich von den schulischen Rahmenbedingungen abhängt: Hohe Handlungsspielräume, eine konstruktive Zusammenarbeit mit den Praktikumslehrkräften und die systematische Verzahnung von individueller Förderung und Fachunterricht begünstigen sowohl die wahrgenommene Kompetenzentwicklung der Studierenden als auch die Qualität der durchgeführten Förderung.

Diagnostische Kompetenz und Förderplanung sind folglich nicht isoliert, sondern als wechselseitig aufeinander bezogene Kernpraktiken professionellen Handelns zu betrachten. Forschung zu data-based decision making und Lernverlaufsdiagnostik zeigt, dass eine fundierte Nutzung diagnostischer Daten die Grundlage für adaptive und wirksame Förderentscheidungen bildet (Filderman et al., 2018; Hövel, 2025). Insbesondere eine hohe Passung zwischen dem individuellen Lernbedarf der Schüler:innen und der geplanten Förderung erhöht die Wahrscheinlichkeit positiver Lernentwicklungen. Zwar lässt sich dieser Zusammenhang in komplexen schulischen Settings nur schwierig nachweisen (Horner et al., 2005), er ist jedoch theoretisch gut begründet und empirisch anschlussfähig, etwa über Befunde zur Bedeutung von Graph Literacy für förderbezogene Entscheidungen (Espin et al., 2018). Abweichende Lernentwicklungen können in diesem Zusammenhang nicht nur als Ergebnis kontextueller Rahmenbedingungen, sondern

auch als Anlass zur kritischen Reflexion diagnostischer Entscheidungen und der daraus abgeleiteten Förderplanung genutzt werden.

Vor diesem Hintergrund liegt ein Potenzial der Studie darin, die Entwicklung diagnostischer Kompetenzen von Studierenden und die lernbezogene Entwicklung geförderter Schüler:innen gemeinsam in den Blick zu nehmen und damit eine in der bisherigen Forschung häufig getrennt betrachtete Beziehung mithilfe des DBR-Ansatzes zu adressieren. Trotz begrenzter Implementationsbedingungen verbesserten sich die Lernleistungen der geförderten Schüler:innen moderat bis deutlich ($PND = 76,7\%$, $PEM = 83,4\%$, $PAND = 84,3\%$). Parallel dazu schätzten die Studierenden ihre Kompetenzen in der Lernverlaufsdiagnostik nach dem Praktikum signifikant höher ein als vorher ($p = .004$, $d = 1.21$). Diese Befunde sprechen dafür, dass das im Praktikum umgesetzte Design auch dann wirksam bleibt, wenn Förderung überwiegend additiv, zeitlich eingeschränkt und nur teilweise in den Regelunterricht integriert erfolgt, wie classroom-based Assessment im engeren Sinne vorsieht.

Die Ergebnisse machen somit zentrale Spannungsfelder sichtbar, die für die Weiterentwicklung des Designs relevant sind. Die Auswahl der Förderkinder erfolgte nur in etwa 33 % der Fälle datenbasiert durch die Studierenden, während in 67 % der Fälle Praktikumslehrkräfte unabhängig von den erhobenen Screeningdaten entschieden. Diagnostische Daten standen damit zwar zur Verfügung, beeinflussten die Förderentscheidungen jedoch nur begrenzt. Diese Diskrepanz verweist auf strukturelle und professionelle Aushandlungsprozesse, die die Umsetzung datenbasierter Diagnostik im Praktikum begrenzen. DBR ermöglicht es, diese Abweichungen zwischen intendiertem Design und tatsächlicher Nutzung systematisch zu dokumentieren und als Ansatzpunkte für Designanpassungen zu nutzen, anstatt sie als methodische Störfaktoren auszublenden.

5.1 Praktische Implikationen

Die Ergebnisse zeigen, dass Studierende im Praktikum in der Lage sind, diagnostische Verfahren systematisch umzusetzen und lernverlaufsdiagnostische Daten zur Evaluation von Fördermaßnahmen zu nutzen. Die signifikanten Kompetenzzuwächse in der Lernverlaufsdiagnostik der Studierenden sowie die nachgewiesenen Lernfortschritte der geförderten Schüler:innen bestätigen zum einen die Wirksamkeit von formativer Förderung und Diagnostik für das Lernen von *struggling students* (Fuchs et al., 2025) und verdeutlichen zugleich, dass Studierende diagnostische Core Practices praktisch anwenden können. Diese Befunde sprechen dafür, dass diagnostische Kompetenzen nicht erst im späteren Berufsverlauf, sondern bereits in frühen Praxisphasen gezielt aufgebaut und gefestigt werden können.

Für die Hochschullehre ergeben sich daraus konkrete Ansatzpunkte zur Stärkung von Transferfähigkeit und Skalierung, indem sie strukturierte Entscheidungs- und Reflexionshilfen bereitstellt, etwa durch verbindliche Dokumentationsformate, fallbasierte Reflexionsaufgaben oder den Vergleich eigener Entscheidungen mit alternativen datenbasierten Vorgehensweisen. Digitale Lernverlaufsdiagnostik bietet hierfür ein skalierbares Instrument. Auf diese Weise kann die Nutzung diagnostischer Daten systematisch gestärkt werden, ohne die notwendige Flexibilität im schulischen Alltag einzuschränken.

Für die Weiterentwicklung des DBR-Designs zeigen die Ergebnisse, dass praxisnahe Ausbildungsformate wirksam sind, wenn sie an bestehende schulische Routinen anschließen und gleichzeitig neue Handlungsoptionen eröffnen. Die positiven Effekte trotz organisatorischer Einschränkungen sprechen dafür, den eingeschlagenen Weg iterativ weiterzuverfolgen. Künftige DBR-Zyklen können darauf aufbauen, indem die Zusammenarbeit mit Praktikumslehrkräften frühzeitig einbezogen, diagnostische Prozesse transparenter gestaltet und erfolgreiche Umsetzungsbedingungen systematisch identifiziert werden.

5.2 Limitationen

Die Ergebnisse sind vor dem Hintergrund methodischer Limitationen zu interpretieren. Die Erfassung diagnostischer Kompetenzen basiert auf Selbsteinschätzungen der Studierenden, die durch soziale Erwünschtheit positiv verzerrt sein können. Zudem waren die Forschenden in die Konzeption und Durchführung des Seminars eingebunden, was sowohl die Auswahl der Instrumente als auch das Antwortverhalten der Studierenden beeinflusst haben könnte. Ein belastbarer Nachweis einer Intervention stand nicht im Vordergrund. Auch der Zusammenhang zwischen Kompetenzzuwachs der Studierenden und Lernzuwachs der Schüler:innen ist kritisch einzuordnen und erfordert ein stärker kontrolliertes Design. Aufgrund der Einbettung in den Studienbetrieb kann sich das vorliegende Design nur an zentralen

Kriterien der Einzelfallforschung, wie die klare Phasenstruktur (Baseline-Intervention-Follow-up) sowie wiederholte Messungen mit standardisierten Instrumenten orientieren, jedoch konnten wesentliche Anforderungen der WWC-Standards – wie stabile Baselines mit ausreichender Messdichte (typisch ≥ 3 –5, der vorliegende Beitrag hat lediglich $M = 2.7$ Messzeitpunkte in der Baseline und Follow-up-Phase), systematische intraindividuelle Replikation von Effekten oder die Kontrolle von Störvariablen durch experimentelle Designs (beispielsweise gestaffelte Interventionen) – unter den Bedingungen schulischer Praxis und des studienbegleitenden Praktikums nicht vollständig umgesetzt werden. Ebenso wurde aus Gründen der Vergleichbarkeit eine Aggregation der Einzelfalldaten über alle Fälle hinweg vorgenommen, wodurch die für Einzelfalldesigns zentrale fallbezogene Analyse und Replikationslogik eingeschränkt ist und differenzielle Wirkverläufe einzelner Interventionen nur begrenzt abgebildet werden konnten. Ferner ist die Studie in ein spezifisches landesspezifisches Schulsystem eingebettet, dessen rechtliche und strukturelle Ausgestaltung inklusiver Bildung die Übertragbarkeit der Ergebnisse einschränkt.

Die Umsetzung des Designs war durch eine hohe Variabilität in Umfang, Zeitpunkt und Integration der Förderung gekennzeichnet, sodass Effekte nicht eindeutig einzelnen Designelementen zugeordnet werden können. Der partizipative Charakter des DBR-Ansatzes zeigte sich primär in der praktischen Mitwirkung der Studierenden und Praktikumslehrkräfte. Aufgrund der geringen Stichprobengröße und der heterogenen Förderumsetzung war eine Differenzierung der Ergebnisse nach dem Grad der Evidenzbasierung nicht möglich und wird in einem weiteren Durchlauf im Sinne des DBR-Ansatzes durch ein geschärftes Design adressiert. Schließlich erlauben die Ergebnisse aufgrund des einmaligen Designzyklus keine Aussagen zur langfristigen Stabilität und Transferfähigkeit der berichteten Kompetenzgewinne über das Praktikum hinaus.

6 AUSBLICK

Aktuell besuchen Studierende eine Vielzahl von schulischen Praktika und arbeiten meist ab dem 2. Semester als Aushilfslehrkräfte. Dabei haben sie zwar eine Vielzahl von aktuellen schulischen Eindrücken, jedoch erlernen sie nur bedingt die Core Practices sonderpädagogischen Handelns systematisch. Insbesondere aus Sicht der Transformation zu einem inklusiven Bildungssystem gibt es pädagogische Herausforderungen und Notwendigkeiten, bestehende Handlungsweisen zu hinterfragen und zu verändern (Capovilla et al., 2023). Veränderung bedeutet, bestehende Strukturen vorsichtig aufzubrechen und weiterzuentwickeln. Zwar ist die bestehende Skepsis vor Veränderungen und insbesondere vor neueren digitalen Verfahren wie der Lernverlaufsdagnostik verständlich, aber kein Grund, mit der Entwicklung von innovativen Werkzeugen aufzuhören. So ist es unser Ansatz, durch gemeinsame Gespräche und anwendungsfreundliche Zugänge Akzeptanz zu erzeugen und die Nützlichkeit anhand von Projekten darzustellen.

LITERATUR

- Abraham, M., & Prediger, S. (2025). Scaffolding fifth graders' learning with a digital multi-representation applet: Design research on focusing multiplicative structures with dynamic dot arrays. *Digital Experiences in Mathematics Education*, 11(2), 219–246. <https://doi.org/10.1007/s40751-024-00163-7>
- Bayerisches Staatsministerium für Unterricht und Kultus. (2008). Bekanntmachung über die Organisation der Praktika für das Lehramt für Sonderpädagogik und für das Studium einer sonderpädagogischen Qualifikation im Rahmen der Lehramtsprüfungsordnung I vom 22. September 2008 (KWMBI. S. 373), geändert durch Bekanntmachung vom 6. April 2023 (BayMBI. Nr. 207). <https://www.gesetze-bayern.de/Content/Document/BayVwV154684>
- Brookhart, S. M., & McMillan, J. H. (2020). *Classroom assessment and educational measurement*. Taylor & Francis.
- Bundschuh, K., & Winkler, C. (2019). *Einführung in die sonderpädagogische Diagnostik*. utb.
- Cano-Corres, R., Sánchez-Álvarez, J., & Fuentes-Arderiu, X. (2012). The effect size: beyond statistical significance. *Ejifcc*, 23(1), 19–25.
- Capovilla, D., Falkenstörfer, S., & Gebhardt, M. (2023). Die Begleitung sonderpädagogischer Praktika aus akademischer Sicht. *Zeitschrift für Heilpädagogik*, (10), 446–449.
- Casale, G., Hennemann, T., & Hövel, D. (2014). Systematischer Überblick über deutschsprachige schulbasierte Maßnahmen zur Prävention von Verhaltensstörungen in der Sekundarstufe I. *Empirische Sonderpädagogik*, 4(1), 33–58.

- Casale, G., Husakovic, M., Hagen, T., Hövel, D. C., Krull, J., & Spilles, M. (2017). Effekte eines kognitiv-behavioralen Aufsatztrainings auf die Schreibleistung und das Lern- und Arbeitsverhalten bei Schülern mit ADHS in der Sekundarstufe einer Förderschule. *Empirische Sonderpädagogik*, 9(4), 341–364.
- Ebenbeck, N. (2023). *Computerized adaptive testing in inclusive education* [Doctoral dissertation, Universität Regensburg]. <https://doi.org/10.5283/epub.54551>
- Ebenbeck, N., Jungjohann, J., & Gebhardt, M. (2023). Testbeschreibung des digitalen Levumi-Lesescreenings LES-IN-DIG: Beschreibung der Testkonstruktion sowie der Items der digitalen Screeningtests „Phonologische Bewusstheit“, „Lexikalischer Abruf“, „Blitzlesen“ und „Sinnkonstruierendes Satzlesen“ in deutscher Sprache. <https://doi.org/10.5283/epub.53993>
- Ebenbeck, N., Zellner, J., Jungjohann, J., & Gebhardt, M. (2026, angenommen). Reference norms used in inclusive education – A closer look at teachers’ assessments of reading subskills. *Journal for Educational Research Online*
- Espin, C. A., Saab, N., Pat-El, R., Boender, P. D. M., & van der Veen, J. (2018). Curriculum-Based Measurement progress data: Effects of graph pattern on ease of interpretation. *Zeitschrift für Erziehungswissenschaft*, 21(4), 767–792. <https://doi.org/10.1007/s11618-018-0836-9>
- Filderman, M. J., Toste, J. R., Didion, L. A., Peng, P., & Clemens, N. H. (2018). Data-based decision making in reading interventions: A synthesis and meta-analysis of the effects for struggling readers. *The Journal of Special Education*, 52(3), 174–187. <https://doi.org/10.1177/0022466918790001>
- Frye, S., Terkowsky, C., Haertel, T., Franuszkiewicz, J., & Heix, S. (2020). Re-design eines Laborpraktikums im Lehramtsstudium – didaktische Optimierung mittels Design-Based Research. In C. Terkowsky, D. May, S. Frye, T. Haertel, T. R. Ortelt, S. Heix & K. Lensing (Hrsg.), *Labore in der Hochschullehre – Didaktik, Digitalisierung, Organisation* (S. 95–108). wbv. <https://doi.org/10.3278/6004804w095>
- Fuchs, A., Radkowsch, A., & Sommerhoff, D. (2025). Using learning progress monitoring to promote academic performance? A meta-analysis of the effectiveness. *Educational Research Review*, 46, 100648. <https://doi.org/10.1016/j.edurev.2024.100648>
- Gebhardt (2024). Pädagogische Diagnostik. Leistung, Kompetenz und Entwicklung messen, bewerten und für individuelle Förderung interpretieren. *Münchner Beiträge zur Inklusions- und Sonderpädagogik, Bd. 3. Version 0.3*. München: LMU München.
- Gebhardt, M., & Jungjohann, J. (2020). Analyse der Lernausgangslage und der Lernentwicklung – Prozesse der Förderdiagnostik. In U. Heimlich & F. Wember (Hrsg.), *Didaktik des Unterrichts bei Lernschwierigkeiten: Ein Handbuch für Studium und Praxis* (S. 367–380). Kohlhammer.
- Gebhardt, M., Diehl, K., & Mühling, A. (2016). Online Lernverlaufsmessung für alle SchülerInnen in inklusiven Klassen. *Zeitschrift für Heilpädagogik*, 67(10), 444–454.
- Grossman, P., Hammerness, K., & McDonald, M. (2009). Redefining teaching, re-imagining teacher education. *Teachers and Teaching*, 15(2), 273–289. <https://doi.org/10.1080/13540600902875340>
- Grünke, M. (2012). Auswertung von Daten aus kontrollierten Einzelfallstudien mit Hilfe von Randomisierungstests. *Empirische Sonderpädagogik*, 4(3/4), 247–264. <https://doi.org/10.25656/01:9301>
- Hagen, T. & Hammes-Schmitz, E. (2025). Mehrdimensionale Förderdiagnostik und -planung am Beispiel des GRID-Modells. In E. Hammes-Schmitz & T. Hagen (Hrsg.), *Mehrdimensionale Förderdiagnostik und -planung bei Schwierigkeiten im Lernen und Lernverhalten*. Verlag Modernes Lernen.
- Hagmann-von Arx, P., Savolainen, H., & Link, P. C. (2025). Evidenzbasierte sonderpädagogische Praxis als Baustein inklusiver Schulen: Eine Chance in Hinblick auf Prävention und Intervention im Bereich Verhalten. In M. Wicki & M. Törmänen (Hrsg.) *Bildung für alle stärken. Ein Handbuch für die evidenzbasierte Entwicklung inklusiver Schulen* (S. 129–141). Julius Klinkhardt.
- Hartke, B. (2017). *Handlungsmöglichkeiten Inklusion: Das Rügener Inklusionsmodell*. Kohlhammer.
- Heimlich, U. (2012). Schulische Organisationsformen sonderpädagogischer Förderung auf dem Weg zur Inklusion. In U. Heimlich & J. Kahlert (Hrsg.), *Inklusion in Schule und Unterricht. Wege zur Bildung für alle* (S. 80–116). Kohlhammer.
- Heimlich, U., Lutz, S., & Wilfert, K. (2021). *Ratgeber Förderplanung: Individuelle Lernförderung im Förderschwerpunkt Lernen (1. bis 9. Klasse)*. Persen Verlag.
- Heitzmann, N., Seidel, T., Opitz, A., Hetmanek, A., Wecker, C., Fischer, M. R., Ufer, S., Schmidmaier, R., Neuhäus, B., Siebeck, M., Stürmer, K., Obersteiner, A., Reiss, K., Girwidz, R., & Fischer, F. (2019). Facilitating diagnostic competences in simulations: A conceptual framework and a research agenda for medical and teacher education. *Frontline Learning Research*, 7(4), 1–24. <https://doi.org/10.14786/flr.v7i4.384>
- Hövel, D. C. (2025). Making behaviour visible: data-based behaviour support in schools. Evidence from 20 experimental single-case studies. *Emotional and Behavioural Difficulties*, 1–15. <https://doi.org/10.1080/13632752.2025.2594931>
- Huber, C., & Grosche, M. (2012). Das response-to-intervention-Modell als Grundlage für einen inklusiven Paradigmenwechsel in der Sonderpädagogik. *Zeitschrift für Heilpädagogik*, 63(8), 312–322.

- Jungjohann, J. (2019). *Theoriegeleitete Lernverlaufsdiagnostik zur Erfassung von basalen Lesekompetenzen in der inklusiven Grundschule: Entwicklung, Evaluation und Implementation von webbasierten Testverfahren* [Dissertation, Technische Universität Dortmund]. <https://doi.org/10.17877/DE290R-19805>
- Jungjohann, J., & Gebhardt, M. (2018). Lernverlaufsdiagnostik im inklusiven Anfangsunterricht Lesen – Verschränkung von Lernverlaufsdiagnostik, Förderplanung und Wochenplanarbeit. In F. Hellmich, G. Görel & M. F. Löper (Hrsg.). *Inklusive Schul- und Unterrichtsentwicklung* (S. 160–172). Kohlhammer.
- Jungjohann, J., & Gebhardt, M. (2019). *SinnL-Levumi-Tests zum sinnkonstruierenden Satzlesen als Lernverlaufsdiagnostik: „Sinnkonstruierendes Satzlesen“ der Onlineplattform www.levumi.de*
- Jungjohann, J., & Gebhardt, M. (2023). Dimensions of classroom-based assessments in inclusive education: A teachers' questionnaire for instructional decision-making, educational assessments, identification of special educational needs, and progress monitoring. *International Journal of Special Education*, 38(1), 131–144. <https://doi.org/10.52291/ijse.2023.38.12>
- Jungjohann, J., Diehl, K., & Gebhardt, M. (2021). *PseuWoL-Levumi: Tests der Leseflüssigkeit zur Lernverlaufsdiagnostik „Pseudowörter lesen“ der Onlineplattform www.levumi.de*
- Jungjohann, J., Diehl, K., Mühlhling, A., & Gebhardt, M. (2018). Graphen der Lernverlaufsdiagnostik interpretieren und anwenden – Leseförderung mit der Onlineverlaufsmessung Levumi. *Forschung Sprache*, 6(2), 84–91. <https://doi.org/10.17877/DE290R-19806>
- Jungjohann, J., Gebhardt, M., & Scheer, D. (2022). Understanding and improving teachers' graph literacy for data-based decision-making via video intervention. *Frontiers in Education*, 7, 919152. <https://doi.org/10.3389/educ.2022.919152>
- Kratochwill, T. R., Hitchcock, J., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M., & Shadish, W. R. (2010). Single-case designs technical documentation. *What Works Clearinghouse*. http://ies.ed.gov/ncee/wwc/pdf/wwc_scd.pdf
- Lai, M. K., McNaughton, S., Amituanai-Tolao, M., Turner, R., & Hsiao, S. (2009). Sustained acceleration of achievement in reading comprehension: The New Zealand experience. *Reading Research Quarterly*, 44(1), 30–56. <https://doi.org/10.1598/RRQ.44.1.2>
- Ma, H.-H. (2006). An Alternative Method for Quantitative Synthesis of Single-Subject Researches: Percentage of Data Points Exceeding the Median. *Behavior Modification*, 30(5), 598–617. <https://doi.org/10.1177/0145445504272974>
- Moser Opitz, E., Pool Maag, S., & Labhart, D. (2019). Förderpläne: Instrument zur Förderung oder “bürokratisches Mittel”? Eine empirische Untersuchung zum Einsatz von Förderplänen. *Empirische Sonderpädagogik*, 11(3), 210–224. <https://doi.org/10.25656/01:17780>
- Nitz, J., Niederelz, A., Hanisch, C., & Hennemann, T. (2024). Herausforderungen und Gelingensbedingungen erfolgreicher Implementation eines mehrstufigen Förderansatzes an Grundschulen. *Zeitschrift für Heilpädagogik*, 75(1), 4–17.
- Oxford Centre for Evidence-based Medicine (2011). *The Oxford 2011 Levels of Evidence*. <https://www.cebm.ox.ac.uk/resources/levels-of-evidence/ocebml-levels-of-evidence>
- Parker, R. I., & Vannest, K. J. (2009). An improved effect size for single-case research: Nonoverlap of all pairs. *Behavior Therapy*, 40, 357–367. <https://doi.org/10.1016/j.beth.2008.11.005>
- Parker, R. I., Vannest, K. J., & Davis, J. L. (2011). Effect size in single-case research: A review of nine nonoverlap techniques. *Behavior Modification*, 35(4), 303–322. <https://doi.org/10.1177/0145445511399147>
- Sailer, M., Bauer, E., Hofmann, R., Kiesewetter, J., Glas, J., Gurevych, I., & Fischer, F. (2023). Adaptive feedback from artificial neural networks facilitates pre-service teachers' diagnostic reasoning in simulation-based learning. *Learning and Instruction*, 83, 101620. <https://doi.org/10.1016/j.learninstruc.2023.101620>
- Schildkamp, K., & Ehren, M. (2013). From “intuition” to “data”-based decision making in Dutch secondary schools? In K. Schildkamp, M. K. Lai, & L. Earl (Eds.), *Data-based decision making in education: Challenges and opportunities* (pp. 49–67). Springer. https://doi.org/10.1007/978-94-007-4816-3_4
- Schneider, W., Schlagmüller, M., & Ennemoser, M. (2017). *LGVT 5-12+. Lesegeschwindigkeits- und -verständnis-test für die Klassen 5-13. Manual* (2., erweiterte und neu normierte Auflage). Hogrefe.
- Schroeder, R., Reh, A., Miller, S., Blumberg, E., Kottmann, B., Franzen, K., & Thevißen, M. (2025). Didaktisch-diagnostische Potenziale des inklusionsorientierten Sachunterrichts (DiPoSa) – Einblicke in ein designbasiertes Entwicklungsprojekt. *Förderbezogene Diagnostik in der inklusiven Bildung*, 1, 209–229.
- Scruggs, T. E., & Mastropieri, M. A. (1998). Summarizing single-subject research: Issues and applications. *Behavior Modification*, 22(3), 221–242. <https://doi.org/10.1177/014544559802200303>
- Staman, L., Timmermans, A., & Visscher, A. J. (2017). Effects of a data-based decision making intervention on student achievement. *Studies in Educational Evaluation*, 55, 58–67. <https://doi.org/10.1016/j.stueduc.2017.07.002>
- Südkamp, A., Kaiser, J., & Möller, J. (2012). Accuracy of teachers' judgments of students' academic achievement: A meta-analysis. *Journal of Educational Psychology*, 104(3), 743–762. <https://doi.org/10.1037/a0027627>

- Voß, S., Blumenthal, Y., Sikora, S., Mahlau, K., Diehl, K. & Hartke, B. (2014). Rügener Inklusionsmodell (RIM) – Effekte eines Beschulungsansatzes nach dem Response to Intervention-Ansatz auf die Rechen- und Leseleistungen von Grundschulkindern. *Empirische Sonderpädagogik*, 6(2), 114-132. <https://doi.org/10.25656/01:9248>
- Wilbert, J., & Lueke, T. (2019, August). *Package ‘scan’* (R package).
- Wild, J., Kraus, E., Steinert, M., Hilbert, S. & Schilcher, A. (2022). Ein Lesestrategietraining im Distanzunterricht? Wie sich das veränderte Lernumfeld während der Corona-Pandemie auf die Entwicklung des Leseverstehens von Schülerinnen und Schülern auswirkt. In M. Philipp & S. Jambor-Fahlen (Hrsg.), *Lesen: Prozess- und Produktperspektiven von der Wortebene bis zu multiplen Texten*. 77-99. Weinheim.
- Zellner, J. & Gebhardt, M. (2025). Scenario-Based Case Game to Promote Diagnostic Decision-Making in Inclusive Teacher Education. *Empirische Sonderpädagogik* 17(2). S. 124-141. <https://doi.org/10.2440/003-0044>
- Zellner, J., Ebenbeck, N., & Gebhardt, M. (2024). Designing a Digital Flash Reading Test for Data-Based Decisions in Inclusive Classrooms: Duration and Word Length as Difficulty-Generating-Item Characteristics. *Education Sciences*, 15(1), 5. <https://doi.org/10.3390/educsci15010005>
- Zellner, J., Giering, S., Ebenbeck, N., & Gebhardt, M. (2025a). *Von Statusdiagnostik zur prozessorientierten Förderdiagnostik. Pädagogisches Konzept für Seminare und Praktikum*. Ludwig-Maximilians-Universität: Münchener Beiträge zur Inklusion. https://epub.ub.uni-muenchen.de/127071/1/Seminarkonzept_und_best_practice_beispiele_final.pdf
- Zellner, J., Koch, J., Hagmann-von Arx, P., Tönnissen, L., Link, C.-P. & Gebhardt, M. (2025b, Juli). Simulationen zur Förderung der diagnostischen Kompetenzen in der heil- und sonderpädagogischen Hochschulbildung: Strukturiertes Klickspiel oder Einsatz GPTAvatar - ein Vergleich mit Noviz:innen und Expert:innen. [Posterpräsentation]. *Jahreskongress 2025 der Schweizerischen Gesellschaft für Bildungsforschung (SGBF) und der Schweizerischen Gesellschaft für Lehrerinnen- und Lehrerbildung (SGL)*. <https://doi.org/10.13140/RG.2.2.31321.22886>