

Selbstwirksamkeit und Einstellungen zur pädagogischen Diagnostik bei Lehrkräften und Lehramtsstudierenden

Robin Junker

Katholische Hochschule Nordrhein-Westfalen
r.junker@katho-nrw.de; <https://doi.org/10.17883/pa-ho-2026-01-05>

EINGEREICHT 5 DEZ 2025

ÜBERARBEITET 23 FEB 2026

ANGENOMMEN 4 MAR 2026

Die präzise Diagnose des Lernstands und der Förderbedarfe von Schüler:innen ist eine zentrale Kompetenz von Lehrkräften, die sowohl adaptiven Unterricht als auch Bildungsgerechtigkeit unterstützt. Neben fachlichem Wissen und methodischen Fertigkeiten spielen hierbei personale Faktoren wie die diagnostische Selbstwirksamkeit und die Einstellungen zur pädagogischen Diagnostik eine entscheidende Rolle. Die vorliegende Studie untersucht Ausprägungen und Korrelate dieser beiden Dimensionen bei Lehrkräften und Lehramtsstudierenden. Dafür wurden 107 Personen (85 Lehrkräfte, 22 Lehramtsstudierende) aus verschiedenen Schulformen befragt. Die Erhebung erfolgte über standardisierte Online-Fragebögen, darunter ein elaboriertes Instrument, das die Selbstwirksamkeit und Einstellung zu pädagogischer Diagnostik erstmals systematisch und differenziert nach kognitiven, emotionalen und verhaltensbezogenen Facetten erfasst. Die Daten wurden mittels deskriptiver Analysen, Korrelationsanalysen und Varianzanalysen (ANOVA) ausgewertet. Die Ergebnisse zeigen, dass Lehrkräfte eine höhere diagnostische Selbstwirksamkeit aufweisen als Studierende und dass diese mit der Berufserfahrung zunimmt. Zudem variieren Einstellungen zur Diagnostik je nach Schulform: Lehrkräfte in spezialisierten Schulformen zeigen positivere diagnostische Einstellungen und höhere Selbstwirksamkeit. Die Befunde unterstreichen die Notwendigkeit schulformbezogener Fortbildungen zur Förderung diagnostischer Kompetenzen.

Schlüsselwörter: Pädagogische Diagnostik, Selbstwirksamkeitsüberzeugungen, Einstellung, Lehrkräftebildung

1. Einleitung

Die Fähigkeit von Lehrkräften, Lernstände und Förderbedarfe ihrer Schüler:innen präzise zu diagnostizieren, ist eine zentrale Voraussetzung für adaptiven Unterricht und Bildungsgerechtigkeit (KMK, 2004; 2024). Sie gewinnt in Zeiten zunehmender Heterogenität, etwa durch Inklusionsbestrebungen, weiter an Bedeutung. Gleichzeitig zeigen Studien, dass die diagnostischen Fähigkeiten von Lehrkräften sehr unterschiedlich ausgeprägt sind (Helmke, 2017; Südkamp et al., 2012).

Initiativen wie „Leistung macht Schule“ (DLR, 2025), das „Startchancen-Programm“ (BMBF, 2025) und die stärkere Verankerung diagnostischer Kompetenzen (KMK, 2004; 2019) unterstreichen, dass die Förderung diagnostischer Fähigkeiten ein zentraler Baustein gelingenden Lernens in Gegenwart und Zukunft sein soll. Um diese Fähigkeiten auszubauen, spielen neben methodischem Wissen auch die diagnostische Selbstwirksamkeit und die Einstellung zur Diagnostik eine entscheidende Rolle (Baumert & Kunter, 2013; van Ophuysen & Behrmann, 2015).

Die vorliegende Studie untersucht deshalb, wie diagnostische Selbstwirksamkeitsüberzeugung und Einstellungen zur pädagogischen Diagnostik (im Folgenden SWÜPD und EPD) bei Lehrkräften und Lehramtsstudierenden ausgeprägt sind und welche Faktoren mit ihr zusammenhängen. Besonderes Augenmerk liegt auf Unterschieden je nach Berufserfahrung und Schulform. Nach einer theoretischen Einordnung werden dazu die Methode, die Ergebnisse und praxisrelevante Implikationen diskutiert.

2. Stand der Forschung

Pädagogische Diagnostik umfasst das systematische Erfassen und Analysieren von Informationen über Lernende, um fundierte pädagogische Entscheidungen zu treffen (van Ophuysen & Behrmann, 2015). Ingenkamp und Lissmann (2008) definieren sie als eine Reihe diagnostischer Tätigkeiten, die sich auf die Ermittlung von Lernvoraussetzungen, die Analyse von Lernprozessen und die Feststellung von Lernergebnissen beziehen, um Lernen zu optimieren. Sie ist zentral für Unterrichtsplanung, Notenvergabe und Bildungsentscheidungen wie Übergangsempfehlungen (van Ophuysen & Behrmann, 2015). Trotz dieser Bedeutung zeigt die Forschung, dass die diagnostische Kompetenz von Lehrkräften insgesamt nur mittelmäßig ausgeprägt ist (Anders et al., 2010; Südkamp et al., 2012). Während die Rangreihenbildung von Schülerleistungen meist gelingt, neigen Lehrkräfte dazu, Leistungen eher zu überschätzen, insbesondere bei leistungsschwächeren Schüler:innen (Westphal et al., 2017). Zudem variiert die Genauigkeit der Einschätzungen stark zwischen Lehrkräften, wobei kognitive Fähigkeiten meist präziser eingeschätzt werden als emotional-motivationale Merkmale (Rost & Hanses, 1998).

Herausforderungen bestehen in der Identifikation und Förderung leistungsstarker und leistungsschwächerer Schüler:innen (ebd.; Artelt et al., 2001; Westphal et al., 2017). Auch soziale Stereotype beeinflussen diagnostische Urteile: So werden beispielsweise Kinder aus bildungsnahen Elternhäusern positiver bewertet als vergleichbar leistungsstarke Kinder aus bildungsfernen Familien (Maaz & Nagy, 2010; Steinmayr, 2017). Zudem werden Potenziale in weniger stereotypbasierten Bereichen, wie hohe mathematische Fähigkeiten bei Mädchen, oft übersehen (Dicke et al., 2012; Lohman, 2005).

Eine fundierte diagnostische Kompetenz hat nachweislich positive Effekte auf den Unterricht und den Lernerfolg. Lehrkräfte mit ausgeprägten diagnostischen Fähigkeiten differenzieren ihren Unterricht stärker, setzen kognitiv aktivierende Materialien ein und wählen gezielt geeignete Fördermaßnahmen (Anders et al., 2010; van Ophuysen & Lintorf, 2013). Diagnostisch kompetente Lehrkräfte zeigen zudem eine größere Flexibilität in ihren Urteilen und sind eher bereit, Einschätzungen bei neuen Erkenntnissen anzupassen (Krolak-Schwerdt & Rummer, 2005). Langfristig trägt eine höhere diagnostische Kompetenz der Lehrkräfte somit zu einem erhöhten Lernerfolg der Schüler:innen bei, etwa im Bereich des Leseverständnisses (Förster & Souvignier, 2014).

Angesichts dieser Erkenntnisse betonen sowohl wissenschaftliche Studien als auch bildungspolitische Vorgaben (z. B. KMK, 2004) die Notwendigkeit, diagnostische Kompetenzen gezielt zu erheben und zu fördern.

2.1 Einstellung und Selbstwirksamkeitsüberzeugungen als Teil der diagnostischen Kompetenz

Die Qualität pädagogischer Diagnostik hängt nicht nur vom fachlichen, fachdidaktischen oder methodischen Wissen der Lehrkraft ab (Philipp & Gobeli-Egloff, 2022), sondern auch von ihren persönlichen Überzeugungen und Einstellungen (Müller & Wagner, 2024). Zwei zentrale personale Faktoren stehen dabei im Fokus: die SWÜPD, die das Vertrauen in die eigene Fähigkeit, diagnostische Aufgaben erfolgreich zu bewältigen, beschreibt und die EPD, welche die persönliche Werthaltung zur Relevanz und Nützlichkeit diagnostischer Tätigkeiten darstellt (Baumert & Kunter, 2013; Müller & Wagner, 2024).

Diagnostisches Handeln umfasst alle Aktivitäten von Lehrkräften, die systematisch Informationen über Lernstände, Fähigkeiten oder Unterstützungsbedarfe von Schüler:innen erheben, interpretieren und zur Gestaltung von Unterricht nutzen. Ob und wie dieses Handeln stattfindet, wird wesentlich durch SWÜPD und EPD beeinflusst: Lehrkräfte mit hoher Selbstwirksamkeit und positiver Einstellung setzen diagnostische Verfahren aktiver ein, während negative Haltungen zur Vermeidung führen können (van Ophuysen & Behrmann, 2015).

Die Rolle der SWÜPD und EPD wird erstens bereits durch das Kompetenzmodell professionellen Lehrerhandelns von Baumert & Kunter (2013) verdeutlicht, das neben professionellem Wissen auch motivationale und affektive Komponenten mit in die Betrachtung professionellen Handelns einbezieht. Zweitens werden die Selbstwirksamkeitsüberzeugung und Einstellung sowohl im DiaCoM-Rahmenmodell (Loibl et al., 2020) als auch im heuristischen Modell der Urteilsakurtheit nach Südkamp et al. (2012; 2017) ganz explizit als Personenmerkmale bei der Entstehung von diagnostischem Handeln betont. Drittens werden diese Merkmale bei dem Modell prozessdiagnostischer Kompetenzen von Reuker und

Künzell (2021) noch detaillierter als Grundlagen für die Entwicklung der kognitiven diagnostischen Prozesse wie Wahrnehmen, Interpretieren und Entscheiden (Blömeke et al., 2015) definiert. Trotz der zentralen Bedeutung, die SWÜPD und EPD bei der Entwicklung diagnostischen Handelns auf theoretischer Ebene zukommt, zeigt der Forschungsstand, dass motivationale Dispositionen wie SWÜPD und EPD bisher nur unzureichend untersucht sind (Wagner, 2024).

Bekannt ist weiterhin, dass die Wirkung dieser Dispositionen von Kontextfaktoren wie Berufserfahrung und Schulform abhängt: Berufserfahrung kann SWÜPD stärken, gleichzeitig können aber belastende Erfahrungen die EPD beeinträchtigen (Becker et al., 2020). Diagnostische Anforderungen variieren zudem zwischen Schulformen (Kopmann, 2021). Auch geschlechtsspezifische Unterschiede wurden nachgewiesen (Dicke et al., 2012).

2.2 Selbstwirksamkeitsüberzeugungen zur pädagogischen Diagnostik

Die Theorie der Selbstwirksamkeit geht auf Bandura (1997) zurück und beschreibt die subjektive Überzeugung einer Person, eine bestimmte Aufgabe erfolgreich zu bewältigen. Im Lehrberuf ist Selbstwirksamkeit stark mit der Qualität des Unterrichts, dem Umgang mit Herausforderungen und der Unterstützung der Schüler:innen verbunden (Caprara et al., 2006; Klassen & Tze, 2014; Zee & Koomen, 2016). Ross (1995) fand zudem heraus, dass Lehrkräfte mit höheren allgemeinen Selbstwirksamkeitsüberzeugungen sensibler auf Schüler:innen mit geringeren Leistungen eingehen und differenziertere Unterstützungsmaßnahmen ergreifen.

Spezifisch für die pädagogische Diagnostik bedeutet hohe Selbstwirksamkeit, dass Lehrkräfte sich zutrauen, valide diagnostische Urteile zu fällen und darauf basierend geeignete Fördermaßnahmen abzuleiten (van Ophuysen & Behrmann, 2015). Lehrpersonen mit ausgeprägter SWÜPD zeigen in der Regel eine höhere Bereitschaft, diagnostische Instrumente zu nutzen und die Ergebnisse aktiv in ihre Unterrichtsgestaltung zu integrieren (Müller & Wagner, 2024).

Empirische Studien belegen, dass SWÜPD nicht nur mit der Häufigkeit diagnostischer Tätigkeiten, sondern auch mit deren Qualität zusammenhängt: Ohle et al. (2015) fanden in einer Studie mit 121 Lehrkräften beispielsweise heraus, dass Lehrpersonen mit hoher SWÜPD deutlich mehr Zeit für diagnostische Tätigkeiten aufwendeten. SWÜPD wurde hier mit nur 5 Items gemessen, die auf einer Likert-Skala von 1 („stimme gar nicht zu“) bis 4 („stimme sehr zu“) eingeschätzt wurden. Diese Items bezogen sich aber nur auf die Verwendung von Text-Bild-Material im diagnostischen Prozess. Ähnliche Ergebnisse zeigen Westphal et al. (2018), die in einer Untersuchung mit 330 österreichischen Lehrkräften nachwiesen, dass SWÜPD mit einer effektiveren Nutzung diagnostischer Instrumente und einer besseren Ableitung individualisierter Fördermaßnahmen korreliert. Das Instrument von Westphal et al. (2018) stützte sich auf das Instrument von Ohle et al. (2015).

2.3 Einstellungen zu pädagogischer Diagnostik

EPD beeinflussen maßgeblich, ob und wie Lehrkräfte diagnostische Maßnahmen in ihren Unterricht integrieren (Kunter & Pohlmann, 2009). Eine positive EPD bedeutet, dass Lehrkräfte die Bedeutung diagnostischer Tätigkeiten erkennen und wertschätzen. Lehrkräfte betrachten Diagnostik nicht unbedingt als zusätzliche administrative Last, sondern oft als essenzielles Instrument zur Unterstützung der Lernprozesse ihrer Schüler:innen (Müller & Wagner, 2024). Solch positive Einstellungen können mit höherer Motivation, stärkerer Selbstwirksamkeit, stärkerer Selbstreflexion und einem höheren freiwilligen Zeitaufwand für Aufgaben der pädagogischen Diagnostik einhergehen (Ohle et al., 2015). EPD wurde in der Studie von Ohle et al. (2015) wie bei der SWÜPD mit 4-5 Items gemessen, die auf einer Likert-Skala von 1 („stimme gar nicht zu“) bis 4 („stimme sehr zu“) eingeschätzt wurden. Diese Items bezogen sich aber ebenfalls nur auf die Verwendung von Text-Bild-Material im diagnostischen Prozess. Westphal et al. (2018), deren Items auch hier auf Ohle et al. (2015) basierten, fanden heraus, dass Lehrkräfte mit positiver EPD häufiger diagnostische Instrumente einsetzen und ihre Unterrichtsmethoden stärker an den individuellen Lernvoraussetzungen der Schüler:innen ausrichten. Interessanterweise zeigte sich in dieser Studie, dass die EPD unabhängig von der tatsächlich investierten Zeit einen direkten Einfluss auf die wahrgenommene Weiterentwicklung des Unterrichts (beispielsweise im Sinne von Kompetenzorientierung und Differenzierung) hatte. Beide Untersuchungen (Ohle et al., 2015; Westphal et al., 2018) zeigen, dass Lehrkräfte mit positiven EPD stärker dazu neigen, diagnostische Daten systematisch zu nutzen und ihre pädagogischen Maßnahmen entsprechend anzupassen.

Eine negative EPD hingegen kann dazu führen, dass diagnostische Maßnahmen nur widerwillig durchgeführt oder deren Ergebnisse nicht konsequent genutzt werden (van Ophuysen & Behrmann, 2015). Diagnostisches Handeln kann aber ebenso von allgemeinen Einstellungen zum Fach abhängen (Larrain & Kaiser, 2022).

Empirische Studien zeigen, dass Lehramtsstudierende bereits während ihres Studiums grundsätzlich die Relevanz von Diagnostik erkennen, jedoch Schwierigkeiten haben, diagnostische Informationen professionell zu interpretieren (van Ophuysen & Behrmann, 2015).

Praktizierende Lehrkräfte berichten teils von einem hohen Zeitaufwand und empfundenen Belastungen durch diagnostische Verfahren, was auf ambivalente EPD hinweist (Müller & Wagner, 2024). Gleichzeitig belegen Studien, dass eine positive EPD eng mit der Qualität des diagnostischen Handelns zusammenhängt.

3. Ziele der Studie

Die vorliegende Studie untersucht, wie Lehrkräfte und Lehramtsstudierende ihre SWÜPD sowie ihre EPD einschätzen. Ziel ist es, Unterschiede zwischen verschiedenen Gruppen zu analysieren, zum Beispiel in Abhängigkeit von Schulform, Berufserfahrung oder Geschlecht. Die Studie soll damit einen Beitrag zum besseren Verständnis der Bedingungen leisten, unter denen diagnostische Kompetenzen im Lehrberuf entstehen und gestärkt werden können. Differenzierte oder inklusive Schulsysteme könnten eine höhere diagnostische Sensibilität erfordern als stärker homogene Bildungsgänge (Kopmann, 2021). Zudem stellt sich die Frage, ob männliche und weibliche Lehrkräfte Diagnostik unterschiedlich gewichten und ob Berufserfahrung eher Unsicherheiten abbaut oder durch schulische Belastungen den diagnostischen Fokus schmälert (Becker et al., 2020).

Zudem haben nur wenige Studien (Ohle et al., 2015; Westphal et al., 2018) SWÜPD bzw. EPD berücksichtigt, und wenn sie berücksichtigt wurden, waren die Operationalisierungen auf wenige Items beschränkt und stark kontextspezifisch. Mit der vorliegenden Untersuchung soll ein elaboriertes Messinstrument eingesetzt werden, das SWÜPD und EPD systematisch erfasst und dabei differenziert nach kognitiven, emotionalen und verhaltensbezogenen Facetten bzw. verschiedenen diagnostischen Bereichen abbildet.

4. Methode

In diesem Abschnitt werden die Stichprobe, die Instrumente, die Untersuchungsprozedur sowie das Auswertungsverfahren beschrieben.

4.1 Stichprobe

An der Untersuchung nahmen 107 Personen teil, darunter 85 Lehrkräfte und 22 Lehramtsstudierende (24 Männer, 83 Frauen). Entsprechend der ungleichen Geschlechterverteilung sind Ergebnisse, die auf geschlechtsspezifischen Vergleichen basieren, nur explorativ zu verstehen. Die Größe der Gruppe der Lehramtsstudierenden ($n = 22$) erklärt sich dadurch, dass lediglich eine Kohorte im Rahmen einer Lehrveranstaltung zur Verfügung stand. Die Gruppe der Lehrkräfte ($n = 85$) wurde im Kontext mehrerer Fortbildungsangebote rekrutiert, sodass sich die ungleichen Teilstichprobengrößen aus den organisatorischen Rahmenbedingungen ergeben. Auch wenn die Gruppengrößen variieren, erlauben sie einen explorativen Vergleich zwischen beiden Gruppen. Das Durchschnittsalter betrug 39,67 Jahre ($SD = 11,73$), bei Lehrkräften 43,98 Jahre ($SD = 9,09$). Die Lehramtsstudierenden befanden sich im Schnitt im 2,81. Semester ($SD = 1,19$). Die durchschnittliche Berufs-

erfahrung der Lehrkräfte lag bei 16,71 Jahren ($SD = 8,54$), über alle Teilnehmenden hinweg bei 13,28 Jahren ($SD = 10,18$).

Die Mehrheit war an Gymnasien tätig oder in Ausbildung (55,14 %, 59 Personen), gefolgt von Grundschulen (21,50 %, 23 Personen) und Haupt-, Real- und Gesamtschulen (HRSGe-Schulen) (15,89 %, 17 Personen). Zudem waren zwei Personen (1,87 %) an einer Fachschule, eine (0,93 %) an einer Förderschule und fünf (4,67 %) machten keine eindeutige Angabe.

Die Fächerverteilung bezog sich auf alle Lehrkräfte und Lehramtsstudierenden und zeigte insgesamt eine breite Streuung. Am häufigsten wurden Deutsch (50 %, 54 Personen), Mathematik (37,38 %, 40 Personen) und Geschichte (21,50 %, 23 Personen) unterrichtet.

Seltener vertreten waren Erdkunde (7,48 %, 8 Personen), Musik (6,54 %, 7 Personen) und Chemie (2,80 %, 3 Personen). Die geringste Verbreitung hatten Französisch, Informatik und Physik (je 3,74 %, 4 Personen) sowie Wirtschaft/Politik, Darstellendes Spiel/Kunst, Technik/Maschinenbau und sonstige Sprachen (je 2,80 %, 3 Personen).

4.2 Prozedur

Die Datenerhebung erfolgte online und unterschied sich leicht zwischen den Gruppen der Lehramtsstudierenden und der Lehrkräfte. Die Lehramtsstudierenden nahmen im Rahmen einer Lehrveranstaltung zum Thema pädagogische Diagnostik an der Befragung teil. Vor Beginn der Veranstaltung wurden sie gebeten, ihre EPD sowie ihre SWÜPD über „Learnweb“, das Learning Management System der Universität Münster, anzugeben. Die erhobenen Daten wurden anschließend genutzt, um innerhalb der Lehrveranstaltung gezielt auf individuelle und gruppenspezifische Tendenzen einzugehen. Die befragten Lehrkräfte wurden im Vorfeld von Veranstaltungen im Kontext des Projekts "Leistung macht Schule" (LemaS) und des ECHA-Diploms zur Hochbegabtenförderung zur Teilnahme eingeladen. Das ECHA-Diplom (European Council for High Ability) ist eine spezialisierte Weiterbildung für Lehrkräfte, die sich auf die Förderung von Hochbegabung konzentriert und vertiefte diagnostische sowie pädagogische Kompetenzen vermittelt. Die Lehrkräfte erhielten vor diesen Veranstaltungen einen Link zum Evaluationssystem „Evasys“ der Universität Münster und füllten den Fragebogen dort online aus. Wie bei den Studierenden wurden die erhobenen Informationen genutzt, um innerhalb der Fortbildungen gezielt auf SWÜPD und EPD der Teilnehmenden einzugehen.

Zu Beginn des Fragebogens wurde allen Teilnehmenden eine kurze Definition von pädagogischer Diagnostik präsentiert (vgl. Abschnitt 2.1), um ein einheitliches Verständnis des Begriffs sicherzustellen. Die Befragung war in beiden Grup-

pen anonym und freiwillig. Die Bearbeitung des Fragebogens dauerte im Durchschnitt etwa 10 Minuten.

4.3 Instrumente

Zur Erfassung der SWÜPD und EPD wurde ein neu entwickeltes Messinstrument verwendet, das auf dem Münsteraner Einstellungs- und Selbstwirksamkeitsinstrument zum Umgang mit Heterogenität (MESS-H) (Junker et al., 2020) basiert. Die Skala zur SWÜPD gliederte sich in drei zentrale Bereiche: Datenerhebung, Auswertung und Interpretation sowie Kommunikation und Implementation. Dabei erfasste sie, inwiefern Lehrkräfte sich in der Lage fühlten, relevante Informationen durch Beobachtung oder Tests zu erheben, diagnostische Ergebnisse korrekt zu analysieren und darauf aufbauend Fördermaßnahmen abzuleiten. Jeder dieser Bereiche wurde mit zwei Items, jeweils einer positiv und einer negativ formulierten Aussage, gemessen, sodass die Skala insgesamt sechs Items umfasste.

Die Erfassung der EPD erfolgte in drei Dimensionen: kognitive, emotionale und verhaltensbezogene Einstellungen. Während die kognitive Dimension sich auf die wahrgenommene Bedeutung von Diagnostik für die pädagogische Praxis bezog, erfasste die emotionale Dimension affektive Reaktionen wie Freude oder Stress im Umgang mit diagnostischen Tätigkeiten. Die verhaltensbezogene Einstellung spiegelte die Bereitschaft wider, diagnostische Erkenntnisse aktiv im Unterricht einzusetzen. Jede dieser drei Dimensionen wurde durch drei Items abgedeckt.

Alle Aussagen wurden auf einer fünfstufigen Likert-Skala von 1 („stimme überhaupt nicht zu“) bis 5 („stimme voll und ganz zu“) bewertet. Die interne Konsistenz der Skalen wurde mittels Cronbachs α berechnet. Die Skala zur SWÜPD zeigte mit einem α -Wert von 0,83 eine gute Reliabilität, wobei die mittlere Item-Korrelation bei 0,47 lag. Die Einstellungsskala erreichte eine interne Konsistenz von $\alpha = 0,76$ mit einer mittleren Item-Korrelation von 0,27, was eine zufriedenstellende Messgenauigkeit belegt. Die Skalen erwiesen sich insgesamt als reliabel und gut geeignet zur Erfassung von SWÜPD und EPD. Ergänzend wurden demografische und berufsbezogene Merkmale wie Schulform, Berufserfahrung und Geschlecht durch geschlossene Fragen erhoben.

4.4 Auswertung

Die Auswahl der statistischen Verfahren erfolgte hypothesen- und fragestellungsgeleitet. Die statistische Analyse erfolgte mit RStudio (ggplot2, dplyr, psych, glmnet). Deskriptive Analysen (Mittelwerte, Standardabweichungen) ermöglichten zunächst einen Überblick über die Verteilungen der zentralen Variablen. Korrelationen wurden eingesetzt, um Zusammenhänge zwischen SWÜPD, EPD und Kontextvariablen zu prüfen, da diese bivariaten Beziehungen für die theore-

tischen Annahmen zentral sind. Gruppenunterschiede (z. B. zwischen Lehrkräften und Studierenden, zwischen Schulformen) wurden mit *t*-Tests und ANOVAs untersucht, da diese Verfahren etablierte Methoden zum Vergleich von Mittelwerten zwischen Gruppen darstellen.

Die Reliabilitätsprüfungen mit Cronbachs α waren notwendig, da es sich um ein neu entwickeltes Instrument handelte und die interne Konsistenz der Skalen empirisch abgesichert werden musste.

Da es sich bei der vorliegenden Studie um eine Gelegenheitsstichprobe handelt, können die Ergebnisse nicht als repräsentativ im statistischen Sinn betrachtet werden. Entsprechend wurden die Analysen primär explorativ durchgeführt. Die Verteilungen der zentralen Variablen wurden mittels Shapiro-Wilk-Test und visueller Inspektion (Histogramme, Q-Q-Plots) auf Normalität geprüft. Für die beiden abhängigen Variablen (SWÜPD und EPD) ergaben sich in Bezug auf die betrachteten Gruppierungsvariablen (Geschlecht, Schulform, Berufsstand) keine signifikanten Unterschiede der Varianzen (alle $p > ,05$), sodass die Annahme der Varianzhomogenität als erfüllt betrachtet werden konnte.

Die erhobenen Daten wurden aus dem „Learnweb“ und dem Evasys-Evaluationssystem als .csv-Dateien exportiert und anschließend in einer einheitlichen Struktur aufbereitet. Dies umfasste die Prüfung auf fehlende Werte sowie die Bereinigung von Unregelmäßigkeiten in den Antworten. Die Datensätze wurden zunächst auf fehlende Werte geprüft. Einzelne nicht beantwortete Items (weniger als 2 % der Werte) wurden durch paarweisen Ausschluss behandelt. Vollständige Datensätze lagen bei allen zentralen Variablen (SWÜPD und EPD) vor, sodass keine systematischen Ausfälle bestanden.

5. Ergebnisse

In diesem Abschnitt werden die deskriptiven Analysen zu den SWÜPD und EPD, aber auch die Gruppenvergleiche in Bezug auf Berufsstatus, Geschlecht und Schulform dargelegt.

5.1 Deskriptive Ergebnisse

Die deskriptiven Analysen zeigen, dass die SWÜPD der Lehrkräfte im Durchschnitt bei $M = 3,64$ ($SD = 0,83$) liegen. Die EPD weist einen ähnlichen Mittelwert von $M = 3,68$ ($SD = 0,71$) auf.

Innerhalb der Einstellungsskala zeigen sich Unterschiede zwischen der kognitiven, emotionalen und verhaltensbezogenen Dimension. Die kognitive EPD ist mit $M = 3,99$ ($SD = 0,71$) am stärksten ausgeprägt. Die emotionale Einstellung fällt mit $M = 3,75$ ($SD = 0,72$) etwas niedriger aus. Die verhaltensbezogene EPD ist mit $M = 3,26$ ($SD = 0,97$) am geringsten ausgeprägt.

Insgesamt zeigen die Ergebnisse, dass Lehrkräfte der pädagogischen Diagnostik überwiegend positiv gegenüberstehen, jedoch insbesondere im Bereich der Umsetzung Unterschiede bestehen.

TAB. 1 Deskriptive Statistiken der Variablen SWÜPD und EPD bezüglich pädagogischer Diagnostik

	M	SD
SWÜPD	3,64	0,83
EPD	3,68	0,71
kognitiv	3,99	0,71
emotional	3,75	0,72
verhaltensbezogen	3,26	0,97

5.2 Korrelationen mit Berufserfahrung und Alter

Die Korrelationsanalyse zeigt, dass die SWÜPD signifikant positiv mit der EPD ($r = ,76$, $p < ,01$) zusammenhängen. Zudem besteht ein moderater positiver Zusammenhang zwischen den SWÜPD und der Berufserfahrung ($r = ,35$, $p < ,01$) sowie dem Alter ($r = ,32$, $p < ,01$). Die EPD weist hingegen keinen signifikanten Zusammenhang mit Berufserfahrung ($r = ,18$, $p = ,07$) oder Alter ($r = ,12$, $p = ,21$) auf.

Wie erwartet, korrelieren Berufserfahrung und Alter stark miteinander ($r = ,96$, $p < ,01$). Die sehr hohe Korrelation zwischen Alter und Berufserfahrung weist darauf hin, dass beide Variablen in dieser Stichprobe weitgehend dieselbe zeitliche Dimension abbilden.

TAB. 2 Mittelwerte, Standardabweichungen und Korrelationen mit Konfidenzintervallen hinsichtlich der Variablen SWÜPD, EPD, Berufserfahrung und Alter

Variable	M	SD	1	2	3	4
1. SWÜPD	3,64	0,83	—			
2. EPD	3,68	0,71	,76*	—		
			[,66; ,83]			
3. Berufs- erfahrung	13,24	10,19	,35*	,18	—	
			[,17; ,50]	[-,01; ,36]		
4. Alter	39,41	12,04	,32*	,12	,96*	—
			[,14; ,48]	[-,08; ,30]	[,94; ,97]	

Anmerkung. M = Mittelwert, SD = Standardabweichung. Die Werte in den eckigen Klammern stellen die 95%-Konfidenzintervalle dar. * $p < ,01$.

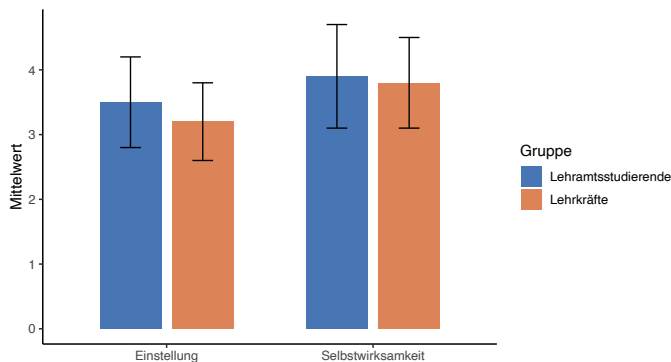
5.3 Unterschiede hinsichtlich des Berufsstands

Zur Untersuchung möglicher Unterschiede zwischen Lehramtsstudierenden und Lehrkräften in Bezug auf ihre SWÜPD und ihre EPD wurden unabhängige t -Tests durchgeführt. Zusätzlich wurde die Effektstärke mittels Cohens d berechnet.

Die Ergebnisse zeigen, dass Lehrkräfte eine signifikant höhere SWÜPD aufweisen als Lehramtsstudierende, $t(106) = 5,29$, $p < ,001$. Der Mittelwert für Lehrkräfte ($M = 3,77$, $SD = 0,83$) lag signifikant über dem Mittelwert der Studierenden ($M = 2,90$, $SD = 0,83$). Die berechnete Effektstärke beträgt $d = 1,26$ (95 %-KI: [0,76, 1,76]), was nach Cohen (1988) als großer Effekt interpretiert werden kann.

Auch hinsichtlich der EPD zeigte sich ein signifikanter Unterschied zwischen den Gruppen, allerdings mit einer geringeren Effektstärke. Der Mittelwert der Lehrkräfte ($M = 3,72$, $SD = 0,71$) war signifikant höher als der der Lehramtsstudierenden ($M = 3,32$, $SD = 0,71$), $t(106) = 2,69$, $p = ,008$. Die berechnete Effektstärke beträgt $d = 0,64$ (95 %-KI: [0,16, 1,12]), was einem mittleren Effekt entspricht.

ABB. 1 SWÜPD und EPD getrennt nach Berufsstand



5.4 Unterschiede hinsichtlich des Geschlechts

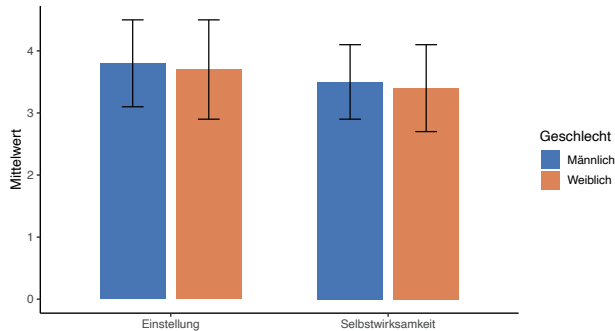
Bezüglich des Geschlechts wurde ein ähnliches Vorgehen wie beim Berufsstand gewählt.

Der t -Test zeigte jedoch keinen signifikanten Unterschied zwischen männlichen und weiblichen Teilnehmenden in Bezug auf ihre SWÜPD, $t(105) = -0,97$, $p = ,336$. Der Mittelwert für männliche Teilnehmende ($M = 3,44$, $SD = 0,68$) unterschied sich nicht signifikant vom Mittelwert der weiblichen Teilnehmenden ($M = 3,61$, $SD = 0,76$). Die berechnete Effektstärke beträgt $d = -0,22$ (95 %-KI: [-0,68, 0,23]), was auf einen kleinen Effekt hindeutet.

Auch hinsichtlich der EPD zeigte sich kein signifikanter Unterschied zwischen den Geschlechtern, $t(105) = -1,37$, $p = ,175$. Der Mittelwert für männliche Teilnehmende ($M = 3,47$, $SD = 0,57$) lag leicht unter dem der weiblichen Teilnehmenden

($M = 3,66$, $SD = 0,60$). Die berechnete Effektstärke beträgt $d = -0,32$ (95 %-KI: $[-0,77, 0,14]$), was einem kleinen bis mittleren Effekt entspricht.

ABB. 2 SWÜPD und EPD getrennt nach Geschlecht



5.5 Unterschiede nach Schulform

Zur Untersuchung möglicher Unterschiede zwischen verschiedenen Schulformen wurde eine einfaktorische Varianzanalyse (ANOVA) durchgeführt. Die Schulformen wurden in vier Gruppen unterteilt: Grundschule ($n = 17$), Gymnasium ($n = 61$), HRSGe (Haupt-, Real-, Sekundar- und Gesamtschulen) ($n = 20$) und Sonstige ($n = 12$). Zusätzlich wurden Post-hoc-Tests (Tukey HSD) durchgeführt, um signifikante Gruppenunterschiede genauer zu bestimmen.

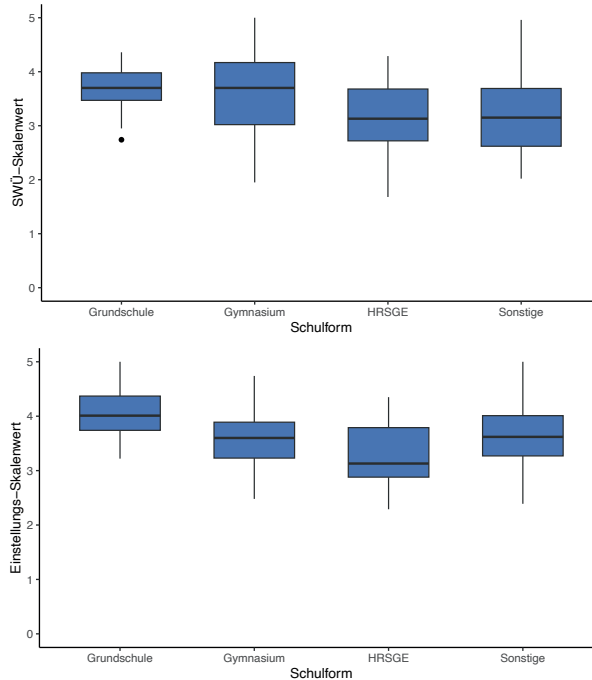
Die ANOVA ergab einen signifikanten Unterschied in der SWÜPD zwischen den Schulformen, $F(3, 106) = 2,74$, $p = ,047$, mit einer Effektstärke von $\eta^2 = 0,07$ (95 %-KI: $[0,00, 1,00]$), was auf einen kleinen Effekt hinweist.

Der Tukey-HSD-Post-hoc-Test zeigte jedoch, dass keine der paarweisen Vergleiche statistisch signifikant waren, mit Ausnahme des Vergleichs zwischen der Gruppe „Sonstige“ und „HRSGe“ ($p = ,049$). Hier zeigte sich, dass Teilnehmende aus der Gruppe „Sonstige“ eine höhere SWÜPD berichteten als Teilnehmende aus der HRSGe-Schulform (M -Differenz = $0,78$, 95 %-KI: $[0,002, 1,55]$). Für die anderen Gruppenvergleiche lagen die p -Werte über $,05$.

Auch für die EPD wurde eine ANOVA durchgeführt. Hier zeigte sich ein hochsignifikanter Unterschied zwischen den Schulformen, $F(3, 106) = 6,10$, $p < ,001$. Die Effektstärke betrug $\eta^2 = 0,15$ (95 %-KI: $[0,04, 1,00]$), was auf einen mittleren Effekt hindeutet.

Der Tukey-HSD-Post-hoc-Test ergab zwei signifikante Unterschiede: Erstens zeigte die Gruppe „Sonstige“ signifikant positivere EPD im Vergleich zur Gymnasialgruppe (M -Differenz = $0,75$, $p = ,003$, 95 %-KI: $[0,21, 1,30]$). Zweitens hatte die Gruppe „Sonstige“ auch eine signifikant positivere EPD als die HRSGe-Gruppe (M -Differenz = $0,81$, $p = ,006$, 95 %-KI: $[0,18, 1,44]$). Die anderen Gruppenvergleiche erreichten keine statistische Signifikanz.

ABB. 3 Boxplots der SWÜPD und EPD getrennt nach Schulform



6. Diskussion

Die vorliegende Studie untersuchte SWÜPD und EPD zur pädagogischen Diagnostik bei Lehrkräften und Lehramtsstudierenden. Dabei wurden Unterschiede in Abhängigkeit von Berufserfahrung, Schulform und Geschlecht analysiert. Während bisherige Arbeiten SWÜPD und EPD nur sehr eingeschränkt und kontextspezifisch operationalisiert haben (vgl. Ohle et al., 2015; Westphal et al., 2018), liegt hier erstmals ein vollständiges Instrument vor, das beide Konstrukte in mehreren Facetten differenziert erfasst. So werden differenzierte Aussagen zu Zusammenhängen zwischen diagnostischer Selbstwirksamkeit, Einstellungen und Kontextfaktoren möglich. Ein zentrales Ergebnis der Untersuchung ist, dass sich die SWÜPD mit zunehmender Berufserfahrung signifikant erhöht. Lehrkräfte wiesen im Vergleich zu Lehramtsstudierenden deutlich höhere SWÜPD-Werte auf. Diese Befunde stehen im Einklang mit bisherigen Studien, die zeigen, dass Lehrkräfte mit wachsender Erfahrung sicherer in der Anwendung diagnostischer Verfahren werden (van Ophuysen & Behrmann, 2015; Ohle et al., 2015). Die EPD veränderte sich mit der Berufserfahrung hingegen nur moderat, was darauf hindeutet, dass eine positive Haltung gegenüber Diagnostik weniger durch Erfahrung allein, sondern möglicherweise durch spezifische Schulungen oder persönliche Überzeugungen geprägt wird (Müller & Wagner, 2024).

Auch zwischen verschiedenen Schulformen zeigten sich Unterschiede, insbesondere in der EPD. Lehrkräfte aus der Gruppe „Sonstige“ (z. B. Förderschulen oder Fachschulen) wiesen im Vergleich zu Gymnasial- und HRSGe-Lehrkräften eine signifikant positivere EPD auf. Dies dürfte daran liegen, dass in spezialisierten Schulformen Diagnostik stärker zur Erstellung individueller Förderpläne dient (Westphal et al., 2018). In der SWÜPD hingegen zeigten sich nur kleinere Unterschiede zwischen den Schulformen. Dies könnte darauf hindeuten, dass Lehrkräfte über Schulformen hinweg ähnliche diagnostische Anforderungen erleben und sich daher keine stark divergierenden Überzeugungen ausbilden.

6.1 Methodische Überlegungen und Limitationen

Eine zentrale Limitation dieser Studie bezieht sich auf die Größe der Gesamtstichprobe. Eine erneute Durchführung der Untersuchung anhand einer weiteren, vielleicht größeren Stichprobe und mit Hilfe des hier entwickelten Instruments könnte die ersten Befunde stärken (Cohen, 1988).

Zudem basieren die Analysen ausschließlich auf Selbstberichten, was eine gewisse Verzerrung durch soziale Erwünschtheit oder Erinnerungsverzerrungen nicht ausschließt (Podsakoff et al., 2003). Eine Kombination von Fragebögen mit objektiveren Messungen, wie z. B. der Analyse realer diagnostischer Entscheidungen oder Beobachtungsdaten aus Unterrichtssituationen, könnte die Validität der Ergebnisse weiter erhöhen (He & Van de Vijver, 2016).

Zukünftige Studien könnten darüber hinaus durch ergänzende qualitative Daten (z. B. aus Interviews oder Fokusgruppen) profitieren (Creswell & Plano Clark, 2017), um besser zu verstehen, welche individuellen oder kontextuellen Faktoren die SWÜPD und EPD beeinflussen. Ebenso wären längsschnittliche Designs sinnvoll, um zu untersuchen, wie sich SWÜPD und EPD über die Zeit entwickeln und welche Faktoren langfristig zur Veränderung dieser Überzeugungen beitragen.

6.2 Praktische Implikationen

Da sich SWÜPD mit der Berufserfahrung signifikant erhöht, wäre es erstens sinnvoll, bereits in der universitären Lehramtsausbildung gezielt Maßnahmen zur Stärkung diagnostischer Kompetenzen zu implementieren. Lehramtsstudierende haben oft noch wenig Erfahrung mit diagnostischen Verfahren und fühlen sich in deren Anwendung unsicher (Kunter & Pohlmann, 2009; van Ophuysen & Behrmann, 2015). Studien zeigen, dass erfahrungsbasierte Lernformate, wie zum Beispiel simulationsbasierte Lehr-Lern-Ansätze, die Selbstwirksamkeit von Lehramtsstudierenden steigern können (Müller & Wagner, 2024; Ohle et al., 2015).

Die Ergebnisse zeigen zweitens, dass Diagnostik in spezialisierten Schulformen, wie Förderschulen oder Fachschulen, eine größere Rolle spielt als in ande-

ren Schulformen. Lehrkräfte in diesen Schulformen zeigten eine positivere EPD als Lehrkräfte an Gymnasien oder Gesamtschulen. Dies könnte darauf hindeuten, dass Diagnostik in stärker differenzierten oder inklusiven Bildungssettings als essenzielles Werkzeug zur individuellen Förderung wahrgenommen wird (van Ophuysen & Lintorf, 2013).

Um die diagnostische Kompetenz in allen Schulformen zu stärken, sollten Fortbildungen spezifisch an die jeweiligen schulischen Anforderungen angepasst werden. So könnten Gymnasiallehrkräfte beispielsweise stärker auf leistungsdifferenzierende Diagnostik und die Identifikation besonders begabter Schüler:innen geschult werden (Westphal et al., 2017).

Drittens müsste in Bezug auf die diagnostischen Anforderungen in verschiedenen Altersklassen und damit verbundenen Schulformen darüber nachgedacht werden, dass pädagogische Diagnostik zum Teil unterschiedliche Funktionen innehat. Während sie in der Grundschule Diagnostik häufig zur frühzeitigen Erkennung von Lernschwierigkeiten genutzt wird (Maaz & Nagy, 2010), liegt der Fokus in weiterführenden Schulen stärker auf der Bewertung von Leistungsentwicklungen und der Beratung zu Bildungswegen (Budde, 2008). Schulformspezifische Fortbildungen sollten daher gezielt auf diese unterschiedlichen Anforderungen eingehen.

Zusammenfassend zeigt sich, dass eine gezielte Förderung diagnostischer Kompetenzen entlang der Ausbildungs- und Berufslaufbahn der Lehrkräfte essenziell ist, um Diagnostik als integralen Bestandteil des professionellen Handelns zu etablieren. Maßnahmen, die sowohl die universitäre Ausbildung als auch die Fort- und Weiterbildung stärker auf SWÜPD ausrichten und die schulformspezifischen Besonderheiten berücksichtigen, könnten dazu beitragen, dass Lehrkräfte Diagnostik nicht nur als Mittel der Leistungsbewertung, sondern als wirksames Instrument zur individuellen Förderung begreifen. Letztlich wäre dies nicht nur für die Lehrkräfte selbst von Vorteil, sondern würde auch eine gezieltere Unterstützung der Schüler:innen ermöglichen und so die diagnostische Praxis langfristig verbessern.

Literatur

- Anders, Y., Kunter, M., Brunner, M., Krauss, S. & Baumert, J. (2010). Diagnostische Fähigkeiten von Mathematiklehrkräften und ihre Auswirkungen auf die Leistungen ihrer Schülerinnen und Schüler. *Psychologie in Erziehung und Unterricht*, 57(3), 175–193. <https://doi.org/10.2378/peu2010.art13d>
- Artelt, C., Schiefele, U. & Schneider, W. (2001). Predictors of reading literacy. *European Journal of Psychology of Education*, 16(3), 363–383. <https://doi.org/10.1007/BF03173188>.
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. Freeman.

- Baumert, J. & Kunter, M. (2013). The COACTIV model of teachers' professional competence. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss & M. Neubrand (Hrsg.), *Cognitive activation in the mathematics classroom and professional competence of teachers* (S. 25–48). Springer.
- Becker, S., Spinath, B., Ditzen, B. & Dörfler, T. (2020). Der Einfluss von Stress auf Prozesse beim diagnostischen Urteilen – eine Eye Tracking-Studie mit mathematischen Textaufgaben. *Unterrichtswissenschaft*, 48, 531–550. <https://doi.org/10.1007/s42010-020-00078-4>
- Blömeke, S., Gustafsson, J. E. & Shavelson, R. J. (2015). Beyond dichotomies competence viewed as a continuum. *Zeitschrift für Psychologie*, 223(1), 3–13. <https://doi.org/10.1027/2151-2604/a000194>
- BMBF (2025). *Startchancen-Programm*. https://www.bmbf.de/DE/Bildung/Schule/Startchancen-Programm/startchancen-programm_node.html
- Budde, J. (2008). Bildungs(miss)erfolge von Jungen und Berufswahlverhalten bei Jungen/männlichen Jugendlichen. In Bundesministerium für Bildung und Forschung (Hrsg.), *Bildungsforschung Band 23* (unveränderter Nachdruck 2010). Bonn, Berlin: BMBF.
- Caprara, G. V., Barbaranelli, C., Steca, P. & Malone, P. S. (2006). Teachers' self-efficacy beliefs as determinants of job satisfaction and students' academic achievement. *Journal of School Psychology*, 44(6), 473–490. <https://doi.org/10.1016/j.jsp.2006.09.001>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.
- Creswell, J. W. & Plano Clark, V. L. (2017). *Designing and conducting mixed methods research* (3rd ed.). SAGE Publications.
- Dicke, T., Parker, P. D., Marsh, H. W., Kunter, M., Schmeck, A. & Leutner, D. (2012). Self-efficacy in classroom management, classroom disturbances, and emotional exhaustion: A moderated mediation analysis of teacher candidates. *Journal of Educational Psychology*, 104(3), 774–786. <https://doi.org/10.1037/a0029356>
- DLR (2025). *Leistung macht Schule*. <https://www.leistung-macht-schule.de/>
- Förster, N. & Souvignier, E. (2014). Learning progress assessment and goal setting: Effects on reading achievement, reading motivation and reading self-concept. *Learning and Instruction*, 32, 91–100. <https://doi.org/10.1016/j.learninstruc.2014.01.002>
- Helmke, A. (2017). *Unterrichtsqualität und Lehrerprofessionalität* (7. Aufl.). Klett.
- Ingenkamp, K.-H. & Lissmann, U. (2008). *Lehrbuch der pädagogischen Diagnostik*. Beltz.
- Junker, R., Zeuch, N., Rott, D., Henke, I., Bartsch, C. & Kürten, R. (2020). Zur Veränderbarkeit von Heterogenitätseinstellungen und -selbstwirksamkeitsüberzeugungen von Lehramtsstudierenden durch diversitätssensible hochschuldidaktische Lehrmodule. *Empirische Sonderpädagogik*, 12(1), 45–63.

- Klassen, R.M. & Tze, V.M. C. (2014). Teachers' self-efficacy, personality, and teaching effectiveness: A meta-analysis. *Educational Research Review*, 12, 59–76. <https://doi.org/10.1016/j.edurev.2014.06.001>
- KMK (2004). *Standards für die Lehrerbildung: Bildungswissenschaften*. Kultusministerkonferenz. https://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2004/2004_12_16-Standards-Lehrerbildung.pdf
- KMK (2019). *Empfehlungen zur schulischen Bildung, Beratung und Unterstützung von Kindern und Jugendlichen im sonderpädagogischen Schwerpunkt LERNEN*. https://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2019/2019_03_14-FS-Lernen.pdf
- KMK (2024). *Empfehlungen zur schulischen Bildung, Beratung und Unterstützung von Kindern und Jugendlichen im sonderpädagogischen Schwerpunkt emotionale und soziale Entwicklung*. https://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2024/2024_12_13-Empfehlung-EmSoz-Entwicklung.pdf
- Kopmann, H. (2021). Zwischen Anspruch und Alltäglichkeit: Konzeptuelle und unterrichtspraktische Perspektiven auf eine inklusive Diagnostik. *Qualifizierung für Inklusion*, 3(2). <https://doi.org/10.25656/01:25442>
- Krolak-Schwerdt, S. & Rummer, R. (2005). Der Einfluss von Expertise auf den Prozess der schulischen Leistungsbeurteilung. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 37(4), 205–213.
- Kunter, M. & Pohlmann, B. (2009). Lehrer. In E. Wild & J. Möller (Hrsg.), *Pädagogische Psychologie* (S. 261–282). Springer.
- Larraín, M. & Kaiser, G. (2022). Interpretation of students' errors as part of the diagnostic competence of pre-service primary school teachers. *Journal für Mathematik-Didaktik*, 43(1), 39–66. <https://doi.org/10.1007/s13138-022-00198-7>
- Lohman, D. F. (2005). The role of nonverbal ability tests in identifying academically gifted students: An aptitude perspective. *Gifted Child Quarterly*, 49(2), 111–138. <https://doi.org/10.1177/001698620504900203>
- Loibl, K., Leuders, T. & Dörfler, T. (2020). A framework for explaining teachers' diagnostic judgements by cognitive modeling (DiacOM). *Teaching and Teacher Education*. <https://doi.org/10.1016/j.tate.2020.103059>
- Maaz, K. & Nagy, G. (2010). Der Übergang von der Grundschule in die weiterführenden Schulen des Sekundarschulsystems: Definition, Spezifikation und Quantifizierung primärer und sekundärer Herkunftseffekte. In J. Baumert, K. Maaz & U. Trautwein (Hrsg.), *Bildungsentscheidungen* (S. 153–182). VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-92216-4_7
- Müller, M.A. & Wagner G. (2024). Einblicke in die Vorstellungen von Lehrkräften betreffend Nutzen standardisierter Diagnoseverfahren. *Pädagogische Horizonte*, 8(1), 1–16. <https://doi.org/10.17883/pa-ho-2024-01-02>

- Ohle, A., McElvany, N., Horz, H. & Ullrich, M. (2015). Text-picture integration - Teachers' attitudes, motivation and self-related cognitions in diagnostics. *Journal for educational research online*, 7(2), 11–33. <https://doi.org/10.25656/01:11488>
- Philipp, K. & Gobeli-Egloff, I. (2022). Förderung diagnostischer Kompetenz im Rahmen der Ausbildung von Lehrkräften für die Primarschule: Eine Studie zum Erkennen von Stärken und Schwächen von Schülerinnen und Schülern am Beispiel von Größen. *Journal für Mathematik-Didaktik*, 43(2), 173–203. <https://doi.org/10.1007/s13138-022-00202-0>
- Podsakoff, P.M., MacKenzie, S.B., Lee, J.-Y. & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, 88(5), 879–903.
- Reuker, S. & Künzell, S. (2021). Learning diagnostic skills for adaptive teaching – A theoretical foundation. *Cogent Education*, 8(1), 1887432. <https://doi.org/10.1080/2331186X.2021.1887432>
- Ross, J.A. (1995). Strategies for enhancing teacher's beliefs in their effectiveness: Research on a school improvement hypothesis. *Teachers College record*, 97, 227–251. <https://doi.org/10.1177/016146819509700205>
- Rost, D.H. & Hanses, P. (1998). Wer nichts leistet, ist nicht begabt? Zur Identifikation hochbegabter Underachiever durch Lehrkräfte. *Labyrinth*, 21(57), 5–11.
- Schroeder, R., Blumberg, E., Kottmann, B., Miller, S. & Reh, A. (2021). Chancen des inklusionsorientierten Sachunterrichts für didaktisch-diagnostisches Handeln: Konzeptionelle und methodologisch-methodische Grundlagen eines forschungsbasierten Entwicklungsansatzes für die Lehrer*innenbildung. *Qualifizierung für Inklusion*, 3(2).
- Steinmayr, R. (Hrsg.) (2017). *FA(IR)BULOUS: FAIRe BeUrteilung des LeistungsPotenzials von Schülerinnen Und Schülern*. https://www.stiftung-mercator.de/content/uploads/2020/12/Fairbulous_Studie.pdf
- 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>
- Südkamp, A., Kaiser, J. & Möller, J. (2017). Ein heuristisches Modell der Akkuratheit diagnostischer Urteile von Lehrkräften. In A. Südkamp & A.-K. Praetorius (Hrsg.), *Diagnostische Kompetenz von Lehrkräften* (S. 33–38). Waxmann.
- He, J. & Van de Vijver, F.J.R. (2016). Response styles in factual items: Personal, contextual and cultural correlates. *International Journal of Psychology*, 51(6), 445–452. <https://doi.org/10.1002/ijop.12263>
- Van Ophuysen, S. & Behrmann, L. (2015). Die Qualität pädagogischer Diagnostik im Lehrerberuf - Anmerkungen zum Themenheft „Diagnostische Kompetenzen von Lehrkräften und ihre Handlungsrelevanz“. *Journal for educational research online*, 7(2), 82–98. <https://doi.org/10.25656/01:11491>

- van Ophuysen, S. & Lintorf, K. (2013). Pädagogische Diagnostik im Schulalltag. In S.-I. Beutel, W. Bos & Raphaela Porsch (Hrsg.), *Lernen in Vielfalt. Chance und Herausforderung für Schul- und Unterrichtsentwicklung* (S. 55–76). Waxmann.
- Wagner, L. (2024). Einflussfaktoren auf die Diagnosekompetenz (angehender) Lehrkräfte – ein systematisches Literaturreview. *Unterrichtswissenschaft*, 53, 99–128. <https://doi.org/10.1007/s42010-024-00215-3>
- Westphal, A., Vock, M. & Stubbe, T. (2017). Grade skipping from the perspective of teachers in Germany: The links between teachers' decisions, acceptance, and perceived knowledge. *Gifted Child Quarterly*, 61(1), 73–86. <https://doi.org/10.1177/0016986216670727>
- Westphal, A., Zuber, J. & Vock, M. (2018). Welche Rolle spielen Selbstwirksamkeit, Motivation und Einstellungen zu Diagnostik für die Nutzung datenbasierter Rückmeldungen? *Zeitschrift für Bildungsforschung*, 8, 289–307. <https://doi.org/10.1007/s35834-018-0223-x>
- Zee, M., & Koomen, H. M. Y. (2016). Teacher self-efficacy and its effects on classroom processes, student academic adjustment, and teacher well-being: A synthesis of 40 years of research. *Review of Educational Research*, 86(4), 981–1015. <https://doi.org/10.3102/0034654315626801>

