

Die dimensionale Struktur diagnostischer Urteilkriterien

Zusammenfassung: Diagnostische Urteile sind ein wichtiger Bestandteil der täglichen Praxis von Lehrkräften und mitentscheidend für die schulische Entwicklung von Schülerinnen und Schülern. Bisherige Untersuchungen geben Einblicke in Urteilsprozesse und deuten auf unterschiedliche inhaltsbezogene und generische Kriterien hin, auf deren Basis diagnostische Urteile getroffen werden. Die Frage, anhand welcher Kriterien angehende Lehrkräfte Lösungen von Schülerinnen und Schülern im Fach Mathematik in unterschiedlichen Phasen des Lehramtsstudiums beurteilen und wie diese Urteile strukturiert sind, ist bisher ungeklärt. Dahingehende Befunde können zum Verständnis diagnostischer Urteile beitragen und perspektivisch Implikationen für die Lehramtsausbildung liefern. Vor diesem Hintergrund wurden Begründungen zur Beurteilung von Schülerlösungen ($n = 110$ angehende Lehrkräfte im Primarstufenstudium) erhoben und inhaltsanalytisch ausgewertet. Auf Basis der identifizierten Urteilkriterien wurde in mehreren Teilstudien mit weiteren Stichproben angehender Primarstufenlehrkräfte ($n_1 = 168$, $n_2 = 209$, $n_3 = 209$) die dimensionale Struktur diagnostischer Urteile angehender Lehrkräfte mittels konfirmatorischer Faktorenanalysen untersucht. Entsprechend der Ergebnisse weisen diagnostische Urteile von angehenden Lehrkräften die folgenden vier Dimensionen bei der Beurteilung mathematischer Schülerlösungen auf: Verstehen, Qualität der Lösung, Präsentation der Lösung und Motivation. Zusätzlich zu den Erkenntnissen über die Struktur diagnostischer Urteile angehender Lehrkräfte wurde ein Erhebungsinstrument entwickelt, das in weiteren Studien für die Untersuchung, wie sich diagnostische Urteile durch entsprechende Lerngelegenheiten entwickeln, genutzt werden kann.

Schlüsselwörter: *Urteilkriterien, Diagnostik, dimensionale Struktur, Lehrkräfteausbildung*

Abstract: Diagnostic judgements are part of teachers' daily practice and crucial for students' academic development. Previous studies provide insights into judgement processes and point to different content-related and generic criteria diagnostic judgements are drawn on. Up to know it has not been clarified what criteria pre-service teachers use to judge students' solutions in mathematics in different phases of teacher education and how these judgements are structured. Findings in this regard can contribute to the understanding of diagnostic judgements and can provide implications for teacher education. Against this background, justifications for the judgements of student solutions ($n = 110$ pre-service teachers in primary education) were collected and investigated by content analysis. Based on the identified judgement criteria, the dimensional structure of diagnostic judgements of pre-service teachers was examined in several sub-studies with further samples of ($n_1 = 168$, $n_2 = 209$, $n_3 = 209$) using confirmatory factor analyses. According to the results, diagnostic judgements of prospective teachers show the following four dimensions in the judgements of mathematical student solutions: Understanding, Quality of solution, Presentation of Procedure, and Motivation. In addition to the findings on the structure of pre-service teachers' diagnostic judgements, a survey instrument was developed that can be used in further studies to investigate how diagnostic judgements develop through appropriate learning opportunities.

Keywords: *judgement criteria, diagnosing, dimensional structure, teacher education*

1. Einleitung

Lehrkräfte stehen regelmäßig vor der Herausforderung, Schülerinnen- und Schülerlösungen mit Blick auf ihren Lernstand, insbesondere in Bezug auf den fachspezifischen Kompetenz- und Wissenserwerb, zu beurteilen. In solchen Anforderungssituationen kommen diagnostische Urteile zum Tragen. Diese entscheiden nicht nur darüber, was den Schülerinnen und Schülern zurückgemeldet wird, sondern auch welche weiteren Lernangebote in welcher Form gemacht werden und haben daher Konsequenzen für die

schulische Entwicklung von Schülerinnen und Schülern (Zhu et al., 2018). Ausgehend von den für die unterschiedlichen Schulfächer formulierten Lernzielen, bilden der Aufbau und die Verknüpfung von Wissens-elementen eine wesentliche Basis für die Kompetenzentwicklung von Schülerinnen und Schülern. In Anlehnung an die Theorien der Kognitionspsychologie zur Konzeptualisierung verschiedener Wissens-elemente (Anderson, 2005) nimmt das prozedurale und konzeptuelle Wissen fach- und gegenstandsübergreifend eine Schlüsselfunktion ein (Scheffler, 1965). Beim Lernen von Mathematik kommt beiden Wissens-elementen eine besonders wichtige Rolle zu (Goldin, 2018). Daraus ergibt sich die Anforderung, zu diagnostizieren, inwieweit prozedurales und konzeptuelles Wissen erworben wurde. Dies stellt insbesondere im Fach Mathematik eine besondere Herausforderung für Lehrkräfte dar, da mathematische Schülerinnen- und Schülerlösungen vielfältig sein können und Aufgaben häufig mehrere Lösungswege zulassen (Durking et al., 2017).

Bisher gibt es nur wenige und keine einheitlichen Befunde dazu, welche Kriterien angehende Lehrkräfte ihren diagnostischen Urteilen zugrunde legen, wenn diese Schülerinnen- und Schülerprodukte in Mathematik beurteilen und inwieweit sich diese auf konzeptuelles oder prozedurales Wissen konzentrieren. Bisherige Befunde deuten darauf hin, dass angehende Lehrkräfte bei der Beurteilung von Schülerinnen- und Schülerlösungen eine fachliche und fachdidaktische Perspektive einnehmen, sich dabei aber häufig auf die oberflächliche Beschreibung des Vorgehens der Schülerinnen und Schüler beschränken, daraus aber keine Rückschlüsse auf die Denkwege ziehen, die dem Vorgehen zugrunde liegen könnten und zudem keine tieferen Analysen dazu vornehmen, ob Schülerinnen und Schüler ein konzeptuelles Verständnis erlangt haben (Crespo, 2000; Talanquer et al., 2015; Baldinger, 2020). Bislang liegen keine umfassenden Erkenntnisse dazu vor, welche Kriterien angehende Lehrkräfte bei der diagnostischen Beurteilung von Schülerinnen- und Schülerlösungen zugrunde legen und wie diese strukturiert sind.

Ausgehend von den Ergebnissen einer qualitativen Vorstudie, in deren Rahmen Kriterien bei der Beurteilung einer mathematischen Schülerlösung offen erhoben wurden, wurden Items entwickelt und anschließend in drei Teilstudien bei unterschiedlichen Gruppen von Lehramtsstudierenden eingesetzt, um die dimensionale Struktur diagnostischer Urteile angehender Lehrkräfte zu untersuchen. In Teilstudie I wurden die entwickelten Items zur Erfassung der Urteilkriterien mit Ratingskalen kombiniert, um eine Schülerlösung einer Aufgabe aus dem Bereich Wahrscheinlichkeiten von angehenden Lehrkräften von zwei Universitäten während ihres Masterstudiums beurteilen zu lassen. In Teilstudie II wurden die Urteilkriterien kombiniert mit Ratingskalen zur Beurteilung einer Schülerlösung aus dem Bereich Arithmetik und angehenden Lehrkräften in unterschiedlichen Phasen ihres

Lehramtsstudiums vorgelegt. Ein auf Basis der Ergebnisse von Teilstudie I und II revidiertes Instrument wurde dann in Teilstudie III erneut bei Lehramtsstudierenden am Ende ihres Bachelorstudiums eingesetzt. Die in den Teilstudien verwendeten Schülerlösungen wurden jeweils so gewählt, dass Studierende zu Beginn, im fortgeschrittenen und am Ende des Bachelorstudiums einbezogen werden konnten. Folglich stammen die Lösungen aus den Inhaltsbereichen, die in den jeweiligen Phasen der Lehramtsstudiums gelehrt wurden, um die fachlichen Voraussetzungen für die diagnostische Beurteilung sicherzustellen.

2. Theorierahmen

Professionelle Kompetenzen bilden eine wichtige Voraussetzung für diagnostische Urteile und umfassen neben Wissen und affektiven Facetten auch situationsspezifische Fertigkeiten (Blömeke et al., 2015). Dazu gehören das Wahrnehmen, Interpretieren und das Treffen von Handlungsentscheidungen, die als Subprozesse des *Noticing* konzeptualisiert werden können (van Es & Sherin, 2021; Jacobs et al., 2011) und insbesondere bei der diagnostischen Beurteilung von Schülerinnen- und Schülerlösungen zum Tragen kommen. In einem Rahmenmodell, das als Heuristik für die empirische Analyse diagnostischer Urteilsbildung dienen soll, integrieren Loibl et al. (2020) diese Subprozesse als Kernstück diagnostischer Urteilsbildung. Das Modell postuliert, dass die diagnostische Urteilsbildung sowohl von Situationscharakteristika als auch von Personencharakteristika, wie den Kompetenzen der urteilenden Person, abhängt. Der Logik des Modells folgend ließe sich für angehende Lehrkräfte annehmen, dass sich in unterschiedlichen Phasen des Lehramtsstudiums diagnostische Urteile unterscheiden, basierend auf den dafür notwendigen Kompetenzen, die sich im Laufe des Studiums entwickeln und in diagnostische Urteile einfließen. Diese Annahme legen auch Modelle zur professionellen Handlungskompetenz von Lehrkräften (Blömeke et al., 2015) und diagnostischer Kompetenz (Herppich et al., 2017) nahe, die Kompetenz als erlernbar, domänenspezifisch und abgrenzbar zu allgemeinen kognitiven Fähigkeiten verstehen (Koeppen et al., 2008). Dementsprechend könnten diagnostische Urteile zu Beginn und am Ende des Studiums unterschiedlich strukturiert sein, sofern im Studium dafür relevante Kompetenzen erworben werden, z. B. dahingehend, welche Kriterien bei der diagnostischen Beurteilung von Schülerinnen- und Schülerprodukten angewendet werden sollten. Empirische Befunde, nach denen z. B. die Verfügbarkeit von Urteilskriterien zu akkurateren Urteilen führt (Quinn, 2020), stützen diese Annahme. Welche Urteilskriterien, auf die sich diagnostische Urteile von (angehenden) Lehrkräften stützen können, in den einzelnen Fächern angemessen sind, klärt sich mit Blick auf die Lernziele. Wie in vielen weiteren Fächern stellt der Erwerb von Wissen zu

Konzepten und Prozeduren ein wichtiges Lernziel in Mathematik dar. Neben prozeduralem Wissen, somit Wissen dazu, wie Verfahren, Algorithmen oder Methoden umgesetzt werden, ist konzeptuelles Wissen im Sinne eines inhaltlichen Verständnisses wesentlicher Begriffe und Vorgehensweisen sowie ihrer Wechselbeziehungen (Rittle-Johnson & Schneider, 2015) ausschlaggebend. Lehrkräfte sollten dementsprechend im Fach Mathematik bei der Beurteilung von Schülerinnen- und Schülerprodukten diagnostizieren, inwieweit Prozeduren angemessen und korrekt auf die Aufgabe angewendet wurde und Schlüsse auf das Schülerinnen- und Schülerdenken und deren konzeptuelles Verständnis der Inhalte ziehen. Urteilkriterien sollten insofern fachbezogene prozedurale und konzeptuelle Aspekte adressieren.

Die wichtige Rolle dieser Kriterien für diagnostische Urteile wird mit Blick auf das heuristische Modell von Loibl et al. deutlich. Urteilkriterien lassen sich den Personencharakteristika zuordnen, die neben den Situationscharakteristika Einfluss auf das diagnostische Denken und schließlich auf das diagnostische Verhalten nehmen. Urteilkriterien kommen in diagnostischen Situationen zum Tragen, in denen (angehende) Lehrkräfte Schülerinnen- und Schülerlösungen beurteilen, in der Regel mit dem Ziel, Schlussfolgerungen über den Lernstand und die erworbenen Kompetenzen zu treffen (Herppich et al., 2017). Was konkret (angehende) Lehrkräfte an einer Schülerinnen- und Schülerlösung in solchen Situationen wahrnehmen, kann in simulierten Situationen mittels verschiedener Methoden erhoben werden, z. B. indem (angehende) Lehrkräfte ihre Überlegungen bei der Beurteilung von Schülerinnen- und Schülerprodukten im Sinne lauten Denkens dokumentieren (Loibl et al., 2020). Die wahrgenommenen Aspekte können dann Hinweise auf mögliche Urteilkriterien liefern.

Wie angehende Lehrkräfte diagnostizieren, wurde bisher im Rahmen von Studien zum *Noticing* von mathematischen Schülerinnen- und Schülerprodukten untersucht, wobei insbesondere die Subfacetten der Wahrnehmung und Interpretation in den Blick genommen wurden. Da die Studien nicht zum Ziel hatten, Kriterien, die angehende Lehrkräfte für die Beurteilung von Schülerinnen- und Schülerprodukten zugrunde legen, zu erheben, verweisen die Ergebnisse nur indirekt auf Urteilkriterien und verschiedene Dimensionen diagnostischer Urteile.

Den empirischen Befunden nach verfolgen angehende Lehrkräfte drei Strategien bei der Beurteilung von Schülerinnen- und Schülerprodukten: Sie argumentieren mathematisch, mathematikdidaktisch und mit Blick auf ihre eigene Lösung (Baldinger, 2020). Laut weiteren Studien beschränken sich die Urteile oft auf eine oberflächliche Beschreibung des Vorgehens in der Schülerinnen- und Schülerlösungen. Inwieweit deutlich wird, ob die Schülerinnen und Schüler ein konzeptuelles Verständnis erlangt hat, wird häufig nicht fokussiert (Talanquer et al., 2015). Weiterhin sind die diagnostischen Urteile

angehender Lehrkräfte eher evaluativer als interpretativer Natur und es werden eher schnelle und umfassende Schlüsse gezogen, anstatt sorgfältig und überlegt zu interpretieren und Schlussfolgerungen zu den Denkwegen und zum Verständnis zu ziehen (Crespo, 2000). Weitere Ergebnisse zeigen, dass Fehler, die auf ein mangelndes konzeptuelles Verständnis der Schülerinnen und Schüler zurückzuführen sind, von angehenden Lehrkräften als Ergebnis eines fehlenden Verständnisses davon, wie ein Prozedur anzuwenden ist bzw. wie genau vorzugehen ist, interpretiert werden anstatt auf mangelndes Konzeptverständnis (Son, 2013) und angehende Lehrkräfte dazu neigen, direkt darauf zu reagieren oder zu korrigieren, anstatt den Schülerinnen und Schülern Fragen zu ihren Lösungen zu stellen, um deren Denken zu ergründen (Cai et al., 2021). Die Ergebnisse von Interventionsstudien deuten jedoch darauf hin, dass Lernangebote diagnostische Urteile von angehenden Lehrkräften dahingehend stärken können, dass diese eine genauere Untersuchung des Denkens der Schülerinnen und Schüler vornehmen (Monson et al., 2018). Die bisherigen Studien zeigen, dass auch wenn sich angehende Lehrkräfte zwar auf fachbezogene Aspekte wie konzeptuelles Verständnis und Prozeduren fokussieren, sie sich aber häufig eher auf Oberflächenmerkmale konzentrieren, um die Schülerinnen- und Schülerlösungen zu beschreiben. Insgesamt geben die Befunde Einblicke in die Vielfalt der Beurteilungskriterien und deuten auf eine multidimensionale Struktur des Konstrukts hin.

3. Ziel und Forschungsfragen

Um zu untersuchen, wie diagnostische Urteile strukturiert sind, wurden drei Teilstudien durchgeführt.

Forschungsfrage in Studie I: Wie sind diagnostische Urteile angehender Lehrkräfte am Ende des Studiums strukturiert, wenn diese eine Schülerlösung¹ aus dem Bereich Wahrscheinlichkeiten beurteilen?

Es wird ausgehend von dem heuristischen Rahmenmodell von Loibl et al. (2020) angenommen, dass der Prozess des *Noticing* einer mathematischen Schülerlösung, der entscheidend für das diagnostische Urteil ist, davon abhängt, welche Kriterien der Beurteilung zugrunde gelegt werden. Es wird weiterhin angenommen, dass die Aspekte, auf die angehende Lehrkräfte fokussieren, wenn sie eine Schülerlösung beurteilen, Rückschlüsse auf ihre Urteilkriterien zulassen. Die einzelnen Urteilkriterien repräsentieren die Dimensionen diagnostischer Urteile. Empirische Befunde legen nahe, dass diagnostische Urteile angehender Lehrkräfte unterschiedliche fachbezogene Kriterien beinhalten (Baldinger, 2020), die sowohl den Lösungsprozess adressierende

¹ Die Lösungen wurden mit Namen von männlichen Schülern kombiniert, daher wird nur die männliche Form benannt.

Oberflächenmerkmale als auch das Verständnis adressierende Tiefenstrukturmerkmale beinhalten (Talanquer et al., 2015, Crespo).

Forschungsfrage in Studie II: Sind diagnostische Urteile angehender Lehrkräfte über die Phasen des Studiums hinweg in gleicher Weise strukturiert, wenn diese eine Schülerlösung aus dem Bereich Arithmetik beurteilen?

Bisherige Befunde geben keine Hinweise darauf, dass für Aufgaben aus unterschiedlichen mathematischen Bereichen unterschiedliche diagnostische Urteilkriterien angewandt werden. Insofern wird eine identische dimensionale Struktur erwartet.

Forschungsfrage in Studie III: Durch welche dimensionale Struktur sind die diagnostischen Urteile angehender Lehrkräfte am Ende ihres Bachelorstudiums am besten repräsentiert?

Analog zu den vorherigen Annahmen wird davon ausgegangen, dass das diagnostische Urteil den Prozess adressierende Oberflächen- und das Konzeptverständnis adressierende Tiefenstrukturmerkmale umfasst.

4. Methodik

Um die Vielfalt der Beurteilungskriterien zu erheben wurden in einer Vorstudie mit einer Gelegenheitsstichprobe von $n = 110$ angehenden Lehrkräften im Rahmen einer Vorlesung zu Grundlagen der Stochastik für die Grundschule eine Schülerlösung zu einer Wahrscheinlichkeitsaufgabe (siehe Abb. 1) vorgelegt, die eine diagnostische Beurteilung mit unterschiedlichen Schwerpunkten und unter Verwendung verschiedener Kategorien ermöglichte. Die angehenden Lehrkräfte wurden gebeten, die Lösung zu beurteilen und ihr Urteil zu begründen. Die offenen Antworten der angehenden Lehrkräfte wurden qualitativ-inhaltsanalytisch einem schrittweisen induktiven Ansatz folgend ausgewertet. Zunächst wurde ein zufällig ausgewähltes Drittel des Datensatzes von einem Forschenden mit mathematikdidaktischer Expertise auf Kategorien von Urteilkriterien hin analysiert, die in den Begründungen ersichtlich wurden. Diese Kategorien wurden anhand des Materials mit zwei weiteren Forschenden diskutiert. Anschließend wurden Codebezeichnungen, Definitionen und Ankerbeispiele sowie Abgrenzungsregeln dokumentiert und in drei Kodierungs- und Rekodierungsrunden überarbeitet, um zu einem zu den Daten passgenauen Kodiermanual zu gelangen (Mayring, 2015). Unter Anwendung der identifizierten Kategorien wurde das gesamte Datenmaterial durch zwei Forschende kodiert. Ein Satz oder mehrere Sätze, die einen Sinnzusammenhang ergaben, wurden als eine Bedeutungseinheit behandelt und einer Kategorie oder wenn nötig mehrerer Kategorien zugeordnet. Für den Kodierungsprozess wurde die Software

MAXQDA 2018 (VERBI Software 2017) verwendet. Die Intercoder-Übereinstimmung wurde berechnet und zeigte eine moderate bis starke Konsistenz mit $k = .79$ und höher für die vier Kategorien (McHugh, 2012). Die Auswertungen verwiesen auf vier Kategorien von Urteilkriterien, die angehende Lehrkräfte zugrundelegen, wenn sie die Schülerlösung beurteilen: Sie fokussierten das Verstehen des Schülers, die angewandte Prozedur, die formale Präsentation der Lösung und die Motivation des Schülers.

Ein Fokus auf das Verstehen wurden kodiert, wenn die angehenden Lehrkräfte das Vorhandensein oder das Fehlen eines Verstehens der Aufgabe oder des Problems direkt oder indirekt erwähnten. Angehende Lehrkräfte mit diesem Fokus betonten häufig, dass der Schüler das Problem begriffen hat und dadurch in der Lage war es richtig zu lösen, und stuften in vielen Fällen die Lösung des Schülers als klug, intelligent oder clever ein. Ein weiterer Fokus auf die Prozedur wurde kodiert, wenn sich die Beurteilungen der Lösung auf das angewendete Verfahren konzentrierte und darauf verwiesen wurde, dass der Schüler es richtig oder falsch bzw. vollständig oder unvollständig angewendet hat. Wenn angehende Lehrkräfte die Präsentation der Lösung benannten, wurde das als weitere Kategorie kodiert. In diesen Fällen verwiesen die angehenden Lehrkräfte darauf, wie der Lösungsprozess dargestellt wurde, dass der Schüler das Problem visualisiert hat oder seinen Lösungsweg dokumentiert und einen Antwortsatz formuliert hat. Sofern das Bemühen des Schülers eine Lösung zu finden anerkannt wurde, wurde dies als Fokus auf die Motivation kodiert. Die vier Kategorien wurden unabhängig davon kodiert, ob die angehenden Lehrkräfte eine defizit- oder eine potentialorientierte Perspektive einnahmen.

Mit dem Fokus auf Verstehen und „angewandte Prozedur“ nahmen also die angehenden Lehrkräfte fachbezogene Indikatoren in den Blick, die auf prozedurales und konzeptuelles Wissen des Schülers hindeuten. Darüber hinaus achteten die angehenden Lehrkräfte darauf, wie die Lösung formal präsentiert wurde. Häufig mündete der Fokus auf die Prozedur in der Beurteilung der formalen Präsentation dieser, also auf einer oberflächlichen Ebene verbleibend. Insofern bilden die beiden Kategorien möglicherweise eine gemeinsame Kategorie. Der Fokus auf die Motivation als weitere Kategorie schließt einen nicht-fachlichen Aspekt mit ein.

Ausgehend von den offenen Antworten der angehenden Lehrkräfte wurden für jede identifizierte Kategorie Items konstruiert. Dazu wurden alle Einzelaussagen in der jeweiligen Kategorie betrachtet und prototypische Aussagen ausgewählt und pointiert formuliert, z. B. „Die Lösung zeigt, dass der Schüler das Problem gut verstanden hat.“ (Kategorie „Verstehen“), „Die Lösung enthält nicht alle Schritte“ (Kategorie „angewandte Prozedur“), „Er hat einen Antwortsatz aufgeschrieben.“ (Kategorie „formale

Präsentation der Lösung“), „Er hat sich angestrengt, um eine Lösung zu finden.“ (Kategorie „Motivation“).

Für die Untersuchung der Fragestellung von Teilstudie I wurden die konstruierten Items mit einer sechsstufigen Ratingskala (von 1 = „stimme voll und ganz zu“ bis 6 = „stimme ganz und gar nicht zu“) kombiniert und für die Beurteilung einer Schülerlösung zu einer Aufgabe aus dem Bereich Wahrscheinlichkeiten herangezogen (siehe Abb. 1). Das Instrument wurde Studierenden zweier Universitäten ($n_1 = 168$) vorgelegt, die sich überwiegend im ersten oder zweiten Semester des Masterstudium befanden. Die Erhebung fand im Rahmen von Lehrveranstaltungen statt; die Teilnahme erfolgte freiwillig.


Aufgabe	Schülerlösung
<p>In einer Schachtel befinden sich 5 Kugeln. Vier Kugeln sind rot und eine Kugel ist schwarz. Mit geschlossenen Augen werden gleichzeitig zwei Kugeln aus der Schachtel gezogen.</p> <p>Wie groß ist die Wahrscheinlichkeit, dass die schwarze Kugel gezogen wird?</p>	 <p>$10 - 6 = 4$ also $\frac{4}{10} = 40\%$ Man bekommt mit 40% die schwarze Kugel.</p>

Abbildung 1. Aufgabe Wahrscheinlichkeiten und Schülerlösung.

Zur Analyse der dimensional Struktur der Items wurden konfirmatorische Faktorenanalysen (CFA) durchgeführt, wobei eine vierdimensionale Struktur mit den Dimensionen „Verstehen“, „angewandte Prozedur“, „formale Präsentation der Lösung“ und „Motivation“ gegen eine dreidimensionale Struktur „Verstehen“, „Präsentation der Prozedur“ und „Motivation“ getestet wurde. Items wurden ausgeschlossen, sofern diese eine geringe Ladung ($< .50$) auf die jeweilige Dimension aufwiesen.

Um zu prüfen, ob sich die gleiche dimensionale Struktur für Studierende in allen Phasen des Lehramtsstudiums zeigt, wurden in Teilstudie II die konstruierten Items ebenfalls mit einer Ratingskala kombiniert um eine Schülerlösung zu einer Aufgabe aus dem Bereich Arithmetik (siehe Abb. 2) von angehenden Lehrkräften ($n_2 = 209$) beurteilen zu lassen. Dieser Inhaltsbereich wurde gewählt, da dieser bereits zu Beginn des Lehramtsstudiums gelehrt wird und in weiteren Veranstaltungen aufgegriffen wird. Die Erhebung erfolgte online während eines digitalen Semesters. Die Teilnahme war freiwillig. Um zu prüfen, inwieweit das diagnostische Urteil in gleicher Weise strukturiert ist, wie durch die Ergebnisse aus Teilstudie I indiziert, wurden CFA durchgeführt.

Aufgabe

In einem Stall befinden sich 21 Tiere.
Es sind Pferde und Fliegen (mit sechs
Beinen). Zusammen haben sie 96
Beine.
Wie viele Pferde und wie viele
Fliegen sind es?

Schülerlösung

21 Tiere : ~~11~~ Pferde ~~10~~ Fliegen
7 Fliegen : 42 Beine
14 Pferde : 56 Beine, zu viel Beine
8 Fliegen : 48 Beine
13 Pferde : 52 Beine, noch mehr Beine
6 Fliegen : 36 Beine
15 Pferde : 60 Beine, sind so 36 Beine

Abbildung 2. Aufgabe Arithmetik und Schülerlösung.

Die Ergebnisse beider Studien verwiesen auf die gleiche dimensionale Struktur, legten jedoch eine genauere Untersuchung einer der Dimensionen nahe. Die Dimension Verstehen adressiert die fachspezifische Anforderung zu diagnostizieren, ob konzeptuelles Wissen erworben wurde bzw. der Schüler zu konzeptuellem Verständnis gelangt ist. Die Items die diese Dimension repräsentieren adressieren sowohl das Verstehen als auch die Qualität der Lösung und adressieren somit möglicherweise zwei unterschiedliche Urteilkriterien. Vor diesem Hintergrund wurden für Studie III weitere Items basierend auf typischen Aussagen der angehenden Lehrkräfte, die im Rahmen der vorangegangenen qualitativen Studie erhoben wurde, konstruiert. Das ursprüngliche Itemset wurde durch die neu konstruierten Items, die das Schülerverstehen bzw. die Qualität der Lösung adressieren ergänzt und in Kombination mit einer Ratingskala zur Beurteilung der Lösung der Wahrscheinlichkeitsaufgabe erneut bei einer weiteren Stichprobe angehender Lehrkräfte ($n_3 = 209$) am Ende ihres Bachelorstudiums eingesetzt. Die Erhebung fand im Rahmen einer Lehrveranstaltung statt, nachdem die Studierenden an einer Vorlesung zu diesem Inhaltsbereich teilgenommen hatten. Die Datenfreigabe erfolgte freiwillig. Auch hier wurde eine CFA durchgeführt, um die Struktur des Konstrukts zu prüfen.

Die Qualität der Modellanpassung wurde in den drei Teilstudien anhand etablierter Anpassungsindizes beurteilt (Hu & Bentler, 1999). *McDonald's ω* wurde als Maß für die Reliabilität berechnet (Hayes & Coutts, 2020).

5. Ergebnisse

Mittels CFA wurde die dreidimensionale Lösung gegen eine vierdimensionale Lösung getestet, vor dem Hintergrund, dass die zuvor durchgeführte qualitative Studie auf eine vierdimensionale Struktur hindeutete. Dazu wurden die Dimensionen „angewandte Prozedur“ und „formale Präsentation der Lösung“ neben den Dimensionen „Verstehen“ und „Motivation“ im vierdimensionalen Modell als getrennte und im dreidimensionalen Modell als gemeinsamer Faktor „Präsentation der Prozedur“ modelliert. Die Ergebnisse der CFA zeigen, dass ein dreidimensionales Modell mit einem gemeinsamen Faktor

„Präsentation der Prozedur“ eine bessere Passung zu den Daten aufweist (AIC = 3812,44, BIC = 3909,28, $X^2 = 37.25$, $df = 23$, $p = .031$, CFI = .96; RMSEA = .06 KI95% [.02 ; .09], SRMR = .06) als ein vierdimensionales Modell mit jeweils getrennten Faktoren für die „formale Präsentation der Lösung“ und die „angewandte Prozedur“ (AIC = 5411,88, BIC = 5546,21, $X^2 = 111,88$, $df = 47$, $p = .000$, CFI = .89, RMSEA = .09 KI95% [.07 ; .11], SRMR = .08). Mit *McDonald's* $\omega = .80$ wies der gemeinsame Faktor „Präsentation der Prozedur“ eine ausreichende Reliabilität auf, zusätzlich zu den Faktoren „Verstehen“ (McDonalds $\omega = .78$) und dem Faktor „Motivation“ (*McDonald's* $\omega = .79$). Jeder Faktor wurde durch die Items mit substantiellen Ladungen von mehr als .57 dargestellt. Die Faktoren korrelierten moderat miteinander, wobei die höchste Korrelation zwischen dem Faktor konzeptioneller Fokus und motivationaler Fokus bestand ($r = .55$, $p < .001$).

In Teilstudie II wurden mittels derselben Items, kombiniert mit einer Ratingskala, die Beurteilung der Lösung einer arithmetischen Aufgabe von Lehramtsstudierenden aller Phasen vorgenommen. Die Ergebnisse der CFA bestätigen auch hier eine dreidimensionale Struktur des Konstrukts ($X^2 = 37.29$, $df = 22$, $p = .022$; CFI = .96; RMSEA = .06 KI95% [.02; .09], SRMR = .05). Die Indizes deuteten auf eine gute Passung hin (Hu & Bentler, 1999). Den Analysen zufolge wurde jede der drei Dimensionen durch drei Items repräsentiert (mit ausreichender Reliabilität, *McDonald's* ω zwischen .75 und .84 (Hayes & Coutts, 2020; McDonald, 1999), alle Faktorladungen waren substantiell. Die Faktoren korrelieren moderat ($r = .74$).

Eine genauere Betrachtung der Itemqualität hinsichtlich der inhaltlichen Interpretierbarkeit legte eine Überarbeitung der Dimension des Faktors „Verstehen“ nahe. Zwei der Items der Dimension bezogen sich auf die Qualität der Lösung der Schüler (wie z. B. „Die Lösung ist clever.“). Ein Item bezog sich zwar explizit auf das konzeptuelle Verständnis der Schüler („(...) zeigte, dass er das Problem gut verstanden hat“). Daher wurden Items zu beiden Aspekten hinzugefügt, um zu prüfen, ob es sich um zwei verschiedene Dimensionen handelt. Die Items wurden auf der Grundlage der in der vorangegangenen qualitativen Studie gewonnenen Daten konstruiert.

Die Ergebnisse der CFA in Studie III zeigten, dass ein vierdimensionales Modell am besten zu den Daten passt (siehe Tab. 1). „Verstehen“ und „Qualität der Lösung“ bildeten nun zwei verschiedene Faktoren zusätzlich zum Faktor „Präsentation der Prozedur“ und dem Faktor „Motivation“ ($X^2 = 86.10$, $df = 48$, $p < .001$; CFI = .97, RMSEA = .06 KI95% [.04 ; .08], SRMR = .05) mit guter Modellanpassung. Den Analysen zufolge, wurden die vier Dimensionen durch drei Items mit einer standardisierten Ladung von jeweils mehr als .55 und einer ausreichenden Reliabilität des jeweiligen latenten Faktors (*McDonald's* ω zwischen .85 und .97) repräsentiert.

Dimension mit Reliabilität	Item	Standardisierte Ladungen	SE	p-Wert
Verstehen <i>McDonald's</i> $\omega = .97$	Die Lösung zeigt, dass der Schüler das Problem gut verstanden hat.	.96	.01	< .001
	Die Lösung deutet darauf hin, dass der Schüler die Aufgabe durchdrungen hat.	.95	.01	< .001
	Der Schüler hat die Problemstellung begriffen.	.96	.01	< .001
Qualität der Lösung <i>McDonald's</i> $\omega = .94$	Der Schüler hat sich das klug überlegt.	.91	.02	< .001
	Die Lösung ist clever.	.94	.01	< .001
	Der Schüler hat die Aufgabe geschickt gelöst.	.91	.03	< .001
Präsentation der Prozedur <i>McDonald's</i> $\omega = .85$	Der Schüler hätte die Lösung besser strukturieren sollen.	.55	.07	< .001
	Der Schüler hätte eine andere Schreibweise wählen sollen.	.86	.03	< .001
	Die Lösung zeigt nicht, wie der Schüler vorgegangen ist.	.96	.02	< .001
Motivation $\omega = .90$	Der Schüler hat sich bemüht, die Aufgabe zu verstehen.	.88	.04	< .001
	Der Schüler hat sich angestrengt, die Aufgabe zu lösen.	.83	.06	< .001
	Der Schüler hat sich Gedanken gemacht, um eine Lösung zu finden.	.88	.05	< .001

Tabelle 1. Dimensionen diagnostischer Urteile und Items und ihre Ladungen (SE = Standardfehler, p = Wahrscheinlichkeitsmaß).

Die Faktoren korrelieren moderat miteinander, mit Ausnahme des Faktors „Verstehen“ und des Faktors „Qualität des Lösungsansatzes“, die eine hohe Korrelation aufwiesen ($r = .82, p < .001$; siehe Tab. 2). Dennoch müssen diese beiden Dimensionen als zwei verschiedene Faktoren modelliert werden, da die CFA bestätigte, dass ein vierdimensionales Modell besser geeignet ist als ein dreidimensionales Modell ($\Delta CFI = .14$).

	Verstehen	Qualität der Lösung	Präsentation der Prozedur
Qualität der Lösung	.82***		
Präsentation der Prozedur	.52***	.59***	
Motivation	.54***	.51***	.21***

Tabelle 2. Korrelationen zwischen den Dimensionen. Anmerkung. *** $p < .001$.

6. Diskussion

In den drei Teilstudien wurden die relevanten Kriterien diagnostischer Urteile von Lehramtsstudierenden bei der exemplarischen Beurteilung von Schülerlösungen sowie die dimensionale Struktur des Konstrukts analysiert. Die angehenden Lehrkräfte beziehen sich sowohl auf inhaltsspezifische Aspekte wie das konzeptionelle Verständnis, die Qualität der Lösung und die Art und Weise, wie das Vorgehen präsentiert wurde, als auch auf allgemeine Aspekte, die die Motivation der Schülerinnen und Schüler adressieren. Die Ergebnisse der CFA bestätigten eine mehrdimensionale Faktorenstruktur diagnostischer Urteile und stehen im Einklang mit früheren Untersuchungen, wonach inhaltsspezifische Aspekte im Mittelpunkt stehen (Baldinger, 2020). Angehende Lehrkräfte scheinen dementsprechend Lernziele wie aufgabenadäquate Prozeduranwendung und Erwerb konzeptuellen Verständnisses im Blick zu haben (Rittle-Johnson & Schneider, 2015). Allerdings zeigen die Studien, dass der Faktor „angewandte Prozedur“ und „formale Präsentation der Lösung“ eine gemeinsame Dimension bilden, was mit früheren Erkenntnissen über diagnostische Urteile angehender Lehrkräfte übereinstimmt, wonach diese einen eher deskriptiven als interpretativen Ansatz verfolgen und ihre diagnostischen Urteile auf die Beschreibung der Lösung beschränken und keine Schlussfolgerungen zum Denken der Schülerinnen und Schüler und deren Verständnis ziehen (Crespo, 2000; Talanquer et al., 2015). Über die Ergebnisse früherer Studien hinaus, die ebenfalls inhaltsspezifische Dimensionen identifiziert haben, hat unsere Studie gezeigt, dass die Motivation der Schülerinnen und Schüler als weitere Dimension berücksichtigt werden sollte, um die diagnostischen Urteile angehender Lehrkräfte adäquat abzubilden. Mit den beiden Dimensionen „Motivation“ und „Präsentation der Prozedur“ werden eher Sicht- bzw. Oberflächenstrukturen der Lösung fokussiert (Kunter & Trautwein, 2013). Die Ergebnisse von Studie III weisen jedoch darauf hin, dass die Faktoren „Verstehen“ und „Qualität der Lösung“ als zwei verschiedene Dimensionen konzeptualisiert werden sollten, und demnach auch Tiefenstrukturen einer Lösung in den Blick genommen werden, indem eine differenzierte Beurteilung, die das Denken der Schülerinnen und Schüler fokussiert, vorgenommen wird (Kunter & Trautwein, 2013).

Mit Blick auf das theoretische Modell, dass als Heuristik diagnostischer Urteilsprozesse zu verstehen ist (Loibl et al., 2020) war zu erwarten, dass angehende Lehrkräfte in unterschiedlichen Phasen ihres Studiums unterschiedlich diagnostisch urteilen. Vor diesem Hintergrund wurde die dimensionale Struktur diagnostischer Urteile an unterschiedlichen Stichproben mit Studierenden in unterschiedlichen Phasen ihres Studiums untersucht und mithin an unterschiedlichen Gegenständen, um die fachlichen Voraussetzungen zur Beurteilung der Schülerlösungen sicherzustellen. Die Struktur der diagnostischen Urteile scheint den Ergebnissen nach in den einzelnen Phasen des

Lehramtsstudiums gleich zu bleiben. Die Ergebnisse von Teilstudie III legen eine differenzierte Modellierung der Faktoren „Verstehen“ und „Qualität der Lösung“ nahe, als zwei Dimensionen, die das Denken von Schülerinnen und Schüler fokussieren. Ein solcher Fokus könnte entsprechend der Ergebnisse der Studie von Monson et al. (2020) durch Lerngelegenheiten angeregt werden, die für eine Entwicklung der diagnostischen Kompetenz förderlich sind (Herppich et al., 2017, Koeppen et al., 2008). Da die Stichprobe in Teilstudie III sich bereits am Ende des Bachelorstudiums befand, kann angenommen werden, dass bereits Lerngelegenheiten wahrgenommen wurden. Es bleibt jedoch unklar, inwieweit diese Lerngelegenheiten auf den Erwerb diagnostischer Kompetenzen ausgerichtet waren. In weiteren Studien könnte geprüft werden, inwieweit der Erwerb diagnostischer Kompetenzen in den einzelnen Lehrveranstaltungen des Lehramtsstudiums curricular implementiert ist und inwieweit diese Lerngelegenheiten prädiktiv für die diagnostischen Urteile von angehenden Lehrkräften sind. Zudem wäre eine dimensionale Analyse unter Einbezug allgemein-kognitiver Fähigkeiten vorteilhaft, um die Domänenspezifität diagnostischer Urteile und die Abgrenzung gegenüber diesen entsprechend der Kompetenzforschung zu validieren (vgl. Blömeke & Jenßen, 2016; Jenßen et al., 2019).

7. Limitationen

Die Generalisierbarkeit der gewonnenen Erkenntnisse ist aufgrund der Durchführung der Teilstudien beschränkt. In allen Teilstudien bildeten Gelegenheitsstichproben die Datenbasis, weshalb keine Rückschlüsse auf die Gesamtheit der Studierenden im Studiengang oder einzelne Kohorten gezogen werden können. Die Teilnahme erfolgte freiwillig, so dass hier möglicherweise ein *Bias* in Bezug auf die Motivation oder den Kompetenzstand der Studierenden zu berücksichtigen ist, der sich möglicherweise in verzerrten Ergebnissen niederschlägt. Entsprechende Kontrollvariablen sollten in Anschlussstudien einbezogen werden.

Die Items sind aus dem Material, das im Rahmen der qualitativen Studie erhoben wurde, formuliert worden. Im Material wurden sehr ähnliche Aussagen zu den jeweiligen Dimensionen vorgefunden. Dies resultiert in einem jeweils engen Fokus der Faktoren. Damit wird möglicherweise nicht der möglichen Varianz innerhalb der Dimensionen Rechnung getragen. Dies wird im Sinne der Praktikabilität von Kurzskalen beim möglichen Einsatz in zukünftigen Studien in Kauf genommen.

8. Schlussfolgerungen

Insgesamt konnte mit den vorliegenden Teilstudien zum Einen zum Verständnis der Struktur diagnostischer Urteile angehender Lehrkräfte beigetragen werden. Zum anderen wurde ein standardisiertes Instrument zur Erhebung diagnostischer Beurteilungskriterien gewonnen. In Kombination mit zwei unterschiedlichen Aufgaben aus zwei Inhaltsgebieten der Mathematik wurde das Instrument bei angehenden Lehrkräften in unterschiedlichen Phasen des Lehramtsstudiums eingesetzt und zeigte eine gute Passung zu den Daten.

Literaturverzeichnis

- Anderson, J.R. (2005). *Cognitive psychology and its implications*. Worth Publishers, 6th edition.
- Baldinger, E. E. (2020). Reasoning about student written work through self-comparison: How pre-service secondary teachers use their own solutions to analyze student work. *Mathematical Thinking and Learning*, 22(1), 56–78. <https://doi.org/10.1080/10986065.2019.1624930>
- 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>
- Blömeke, S., & Jenßen, L. (2016). A question of validity: Clarifying the hierarchical nature of teacher cognition. In M. Rosén, K. Y. Hansen, & U. Wolff (Eds.), *Cognitive abilities and educational outcomes: A festschrift in honour of Jan-Eric Gustafsson* (S. 89–107). Springer. <https://doi.org/10.1007/978-3-319-43473-5>
- Cai, J., LaRochelle, R., Hwang, S., & Kaiser, G. (2021). Expert and preservice secondary teachers' competencies for noticing student thinking about modelling. *Educational Studies in Mathematics*. <https://doi.org/10.1007/s10649-021-10071-y>
- Crespo S. (2000). Seeing more than right and wrong answers: Prospective teachers' interpretations of students' mathematical work. *Journal of Mathematics Teacher Education*, 155–181.
- Durkin, K., Star, J.R. & Rittle-Johnson, B. (2017). Using comparison of multiple strategies in the mathematics classroom: lessons learned and next steps. *ZDM Mathematics Education* 49, 585–597. <https://doi.org/10.1007/s11858-017-0853-9>
- Fischbach, A., Budson, T. G., Preckel, F., Martin, R., & Brunner, M. (2013). Do teacher judgments of student intelligence predict life outcomes? *Learning and Individual Differences*, 27, 109–119. <https://doi.org/10.1016/j.lindif.2013.07.004>
- Goldin, G. A. (2018). Discrete mathematics and the affective dimension of mathematical learning and engagement. In E. W. Hart & J. Sandefur (Hrsg.), *Teaching and learning discrete mathematics worldwide: curriculum and research* (S. 53–65). Springer International Publishing.
- Hayes, A. F. & Coutts, J. J. (2020). Use Omega Rather than Cronbach's Alpha for Estimating Reliability. But... *Communication Methods and Measures*, 14(1), 1-24. <https://doi.org/10.1080/19312458.2020.1718629>
- Herppich, S., Praetorius, A.-K., Förster, N., Glogger-Frey, I., Karst, K., Leutner, D., Behrmann, L., Böhmer, M., Ufer, S., Klug, J., Hetmanek, A., Ohle, A., Böhmer, I., Karing, C., Kaiser, J., & Südkamp, A. (2018). Teachers' assessment competence: Integrating knowledge-, process-, and product-oriented approaches into a competence-oriented conceptual model. *Teaching and Teacher Education*, 76, 181–193. <https://doi.org/10.1016/j.tate.2017.12.001>
- Hu, L. & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1–55. <https://doi.org/10.1080/10705519909540118>

- Jenßen, L., Dunekacke, S., Gustafsson, J.-E., & Blömeke, S. (2019). Intelligence and knowledge: The relationship between preschool teachers' cognitive dispositions in the field of mathematics. *Zeitschrift für Erziehungswissenschaft*, 22, 1313–1332. <https://doi.org/10.1007/s11618-019-00911-2>
- Koepfen, K., Hartig, J., Klieme, E., & Leutner, D. (2008). Current issues in competence modeling and assessment. *Zeitschrift für Psychologie*, 216(2), 61-73. <https://doi.org/10.1027/0044-3409.216.2.61>
- Kuckartz, U. & Rädiker, S. (2019). *Analyzing qualitative data with MAXQDA*. Springer International Publishing.
- Kunter, M. & Trautwein, U. (2013). *Psychologie des Unterrichts*. Schöningh, Paderborn.
- Loibl, K., Leuders, T., & Dörfler, T. (2020). A framework for explaining teachers' diagnostic judgements by cognitive modeling (Diacom). *Teaching and Teacher Education*, 91, 1-10. <https://doi.org/10.1016/j.tate.2020.103059>
- McHugh, M. L. (2012). Interrater reliability: The kappa statistic. *Biochemia Medica*, 22(3), 276-282. <https://doi.org/10.11613/BM.2012.031>
- Monson, D., Krupa, E., Lesseig, K., & Casey, S. (2020). Developing secondary prospective teachers' ability to respond to student work. *Journal of Mathematics Teacher Education*, 23(2), 209–232. <https://doi.org/10.1007/s10857-018-9420-8>
- Ohle, A., & McElvany, N. (2015). Teachers' diagnostic competences and their practical relevance. Special issue editorial. *Journal for Educational Research Online*, 7(2), 5–10. <https://doi.org/10.25656/01>
- Quinn, D. M. (2020). Experimental evidence on teachers' racial bias in student evaluation: The role of grading scales. *Educational Evaluation and Policy Analysis*, 42, 375–392. <https://doi.org/10.3102/0162373720932188>
- Ready, D. D., & Wright, D. L. (2011). Accuracy and inaccuracy in teachers' perceptions of young children's cognitive abilities. The role of child background and classroom context. *American Educational Research Journal*, 48, 335–360. <https://doi.org/10.3102/0002831210374874>
- Rittle-Johnson, B., & Schneider, M. (2015). Developing conceptual and procedural knowledge of mathematics. In R. C. Kadosh & A. Dowker (Hrsg.), Oxford library of psychology. *The Oxford handbook of numerical cognition* (S. 1118–1134). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199642342.013.014>
- Scheffler, I. (1965). *Conditions of knowledge: An introduction to epistemology and education*. University of Chicago Press.
- Talanquer, V., Bolger, M., & Tomanek, D. (2015). Exploring prospective teachers' assessment practices: Noticing and interpreting student understanding in the assessment of written work. *Journal of Research in Science Teaching*, 52(5), 585–609. <https://doi.org/10.1002/tea.21209>
- Zhu, M., Urhahne, D., & Rubie-Davies, C. M. (2018). The longitudinal effects of teacher judgement and different teacher treatment on students' academic outcomes. *Educational Psychology*, 34(1), 1–21. <https://doi.org/10.1080/01443410.2017.1412399>

Informationen zur Förderung

Gefördert durch die Deutsche Forschungsgemeinschaft (DFG) – Projektnummer LA 5067/1-1

Ethische Richtlinien

Ethische (zu stellende Ethikanträge) und rechtliche Anforderungen (Datenschutz-Konzepte, Einverständniserklärungen etc.) wurden in allen Teilstudien beachtet. Es

waren entsprechend der Forderung von AERA, APA & NCME keine negativen Konsequenzen für die Studienteilnehmenden zu erwarten.

Autor:innen

Christin Laschke, Dr., ist Mitarbeiterin in der Abteilung fachbezogener Erkenntnistransfer des Leibniz-Instituts für die Pädagogik der Naturwissenschaften und Mathematik und beschäftigt sich mit den diagnostischen Kompetenzen von Lehrkräften und Lehramtsstudierenden. Sie hat an der Humboldt-Universität zu Berlin zum Thema „Angehende Mathematiklehrkräfte für die Sekundarstufe I in Deutschland und Taiwan - Empirische Befunde und Erklärungsansätze“ promoviert und hat in diesem Zusammenhang einen stipendiengeförderten Forschungsaufenthalt an der National Taiwan Normal University in Taipei realisiert.

Bettina Rösken-Winter, Prof. Dr., ist Professorin für Mathematik in der Primarstufe an der Humboldt-Universität zu Berlin und beschäftigt sich in der Forschung schwerpunktmäßig mit Aspekten der Professionalisierung von Lehrer:innen und Multiplikator:innen. Sie ist stellvertretende Leiterin des DZLM (Deutsches Zentrum für Lehrkräftebildung Mathematik)-Netzwerks der Abteilung „Fachbezogener Erkenntnistransfer“ des Leibniz-Instituts für die Pädagogik der Naturwissenschaften und Mathematik (IPN).

Lars Jenßen, Dr., studierte Psychologie (Diplom) an der Freien Universität Berlin. Zurzeit ist er Wissenschaftskoordinator des Arbeitsbereichs Mathematik in der Primarstufe sowie Vertreter der Professur Erziehungswissenschaftliche Methodenlehre an der Humboldt-Universität zu Berlin. Sein Forschungsschwerpunkt liegt auf der professionellen Kompetenz von Lehrkräften und frühpädagogischen Fachkräften mit einem Fokus auf professionelles Wissen und Emotionen.