

Gastherausgeber:innen des Themaheftes

Dr. Philipp Hendricks
(TU Dortmund) &
Dr. Anneke Langer
(EU Flensburg)

Themaheft: Diagnostische Kompetenz aus sportpädagogischer Perspektive in Theorie und Praxis

Hauptbeiträge

- | | |
|---|-----|
| Diagnostische Kompetenzen aus sportpädagogischer
Perspektive in Theorie und Empirie
<i>Anneke Langer & Philipp Hendricks</i> | 4 |
| Diagnostic Competence and Judgement Accuracy
of Primary School PE Teachers in Assessing Basic Motor
Competencies
<i>Simone Storni, Johanna Kress, Kathrin Bretz et al.</i> | 30 |
| Aufgabenanalytische Kompetenzen angehender
Sportlehrkräfte als eine Voraussetzung zur kognitiven
Aktivierung von Schüler*innen im Sportunterricht
<i>Philipp Hendricks</i> | 54 |
| Digital Transformation of a Diagnostic Procedure for
Swimming Skill Analysis: Teachers' Evaluation of a
Low-Fidelity App Prototype (Click Dummy) of the 'Owl's Eye'
<i>Inga Fokken, Ilka Staub, Michel Brinkschulte et al.</i> | 76 |
| Diagnostische Kompetenzen von Sportlehramtsstudierenden
durch authentische Situationen professionell entwickeln
<i>Anneke Langer & Nele Schlapkohl</i> | 98 |
| Ankündigungen | |
| Ankündigung der Jahrestagung der DGfE 2026 | 123 |

ZEITSCHRIFT FÜR SPORTPÄDAGOGISCHE FORSCHUNG (ZSF)

Organ der Deutschen Vereinigung für Sportwissenschaft (dvs)

Herausgegeben vom Verein zur Förderung sportpädagogischer Forschung e. V.

Herausgeber:innenkollegium

Ulrike Burrmann (Berlin)

Claus Krieger (Hamburg), Hilke Pallesen (Hamburg), Miriam Seyda
(geschäftsführend, Dortmund), Jan Sohnsmeier (Heidelberg), Vera Volkmann (Han-
nover)

Im Auftrag der

Kommission ›Sportpädagogik‹ in der Deutschen Gesellschaft für
Erziehungswissenschaft (DGfE) sowie der Sektion ›Sportpädagogik‹ der Deutschen
Vereinigung für Sportwissenschaft (dvs)

Mitbegründer:innen und ehemalige Mitglieder des Herausgeber:innenkollegiums:

Peter Frei (Hildesheim, 2013-2018), Erin Gerlach (Potsdam, 2017-2020), Elke Grimminger-
Seidensticker (2021-2015), Tim Heemsoth (2021-2025),
Rüdiger Heim (Heidelberg, 2013-2016; geschäftsführend, 2017-2020), Ilka Lüsebrink (Frei-
burg, 2019-2022), Roland Messmer (Basel, 2019-2022), Wolf-Dietrich Miethling (Kiel,
geschäftsführend 2013-2016),
Nils Neuber (Münster, 2013-2016), Jörg Thiele (Dortmund, 2017-2020),
Petra Wolters (Vechta, 2013-2018)

Inhalt

Editorial	3
-----------------	---

HAUPTBEITRÄGE

<i>Anneke Langer & Philipp Hendricks</i> Diagnostische Kompetenzen aus sportpädagogischer Perspektive in Theorie und Empirie	4
<i>Simone Storni, Johanna Kress, Kathrin Bretz et al.</i> Diagnostic Competence and Judgement Accuracy of Primary School PE Teachers in Assessing Basic Motor Competencies	30
<i>Philipp Hendricks</i> Aufgabenanalytische Kompetenzen angehender Sportlehrkräfte als eine Voraussetzung zur kognitiven Aktivierung von Schüler*innen im Sportunterricht	54
<i>Inga Fokken, Ilka Staub, Michel Brinkschulte et al.</i> Digital Transformation of a Diagnostic Procedure for Swimming Skill Analysis: Teachers' Evaluation of a Low-Fidelity App Prototype (Click Dummy) of the 'Owl's Eye'	76
<i>Anneke Langer & Nele Schlapkohl</i> Diagnostische Kompetenzen von Sportlehramtsstudierenden durch authentische Situationen professionell entwickeln	98

ANKÜNDIGUNG

Ankündigung der Jahrestagung der DGfE 2026 »Ehrliche Theorie – für die Sportdidaktik und die Sportpädagogik« 26.01.2027 – 29.01.2027 in Laax Graubünden, Schweiz	122
Ausschreibung	124
Autor:innenverzeichnis	125

Editorial

Liebe Leser:innen,

die aktuelle Ausgabe der *Zeitschrift für Sportpädagogische Forschung* stellt das erste »Special Issue« in der Geschichte der Zeitschrift dar und erscheint zudem ebenso zum ersten Mal im open-access-Format. Dies sind gleich zwei Neuerungen unserer Zeitschrift, die hoffentlich in Zukunft zur Selbstverständlichkeit werden.

Dass wir nun im 14. Jahrgang ein open-access-Format haben, verdanken wir in erster Linie dem Nomos-Verlag, welcher durch ein Subscribe-to-Open-Verfahren insgesamt vier seiner rund 80 Zeitschriften diese Möglichkeit eröffnet hat. Die ZSF ist eine dieser *ausgewählten* vier, welche Erfahrungen in dieser Pilotphase sammeln können. In zweiter Linie verdanken wir die Umsetzung der DGfE-Kommission Sportpädagogik und dem Verein zur Förderung des sportpädagogischen Nachwuchses, die für dieses Vorhaben einen Zuschuss ermöglicht haben. Und in dritter Linie sind wir allen Abonnent:innen dieser Zeitschrift, die für 2026 ihre Gebühr für die Zeitschrift entrichtet haben und auf eine Papierausgabe verzichten, zu großem Dank verpflichtet. Durch diesen Zusammenschluss war es möglich, den Autor:innen eine kostenfreie open access Publikation zu ermöglichen.

Diese Ausgabe in Form eines »Special Issue« trägt den Titel »Diagnostische Kompetenz aus sportpädagogischer Perspektive in Theorie und Empirie«. Als Gastherausgeber:innen fungieren Dr. Philipp Hendricks und Dr. Anneke Langer. Sie haben den Herausgabeprozess eigenständig verantwortet. In den Fällen, in denen sie selbst (Mit-)Autor:innen waren, wurde der Reviewprozess vom regulären Herausgeber:innengremium der ZSF übernommen. Insgesamt ist ein sehr vielschichtiges Heft entstanden, das die »Diagnostische Kompetenz« aus theoretischer, empirischer und handlungspraktischer Perspektive betrachtet. Da die Gastherausgeber:innen selbst eine Einführung in das Heft geben, soll dies an dieser Stelle ausreichen.

Wir möchten Sie abschließend noch auf zwei Dinge hinweisen. Zum einen auf die Ausschreibung von Nachfolge-Positionen im Herausgeber:innenkollegium, zum anderen auf die Ankündigung der nächsten Jahrestagung der DGfE-Kommission für Sportpädagogik am Ende dieses Heftes.

Für das Herausgeber:innengremium im Mai 2026,

Miriam Seyda

Diagnostische Kompetenzen aus sportpädagogischer Perspektive in Theorie und Empirie

von Anneke Langer & Philipp Hendricks*

Zusammenfassung: Diagnostische Kompetenzen können als Teil der professionellen Handlungskompetenzen von Sportlehrkräften verstanden werden. Sie spielen eine zentrale Rolle, um einen für die Schüler:innen lernwirksamen Sportunterricht zu gestalten. Allerdings zeigen bisherige empirische Befunde, dass Sportlehrkräfte aller Schulformen sowohl verschiedene lernrelevante Merkmale von Schüler:innen (z. B. motorische und motivationale) als auch die Merkmale von Aufgaben (z. B. die Schwierigkeit) nicht akkurat einschätzen können. Um dem entgegenzusteuern, ist eine Ausbildung diagnostischer Kompetenzen von Sportlehrkräften erforderlich, die bereits in der Hochschullehre ansetzen sollte. Im Beitrag wird einerseits der aktuelle Diskurs über diagnostische Kompetenzen von Sportlehrkräften auf theoretischer Ebene dargestellt und mit Bezug zur empirischen Herangehensweise in einem Modell strukturiert. Im Modell werden verschiedene Kompetenzfacetten diagnostischer Kompetenzen (angehender) Sportlehrkräfte dargestellt, die einen Einfluss auf die Kompetenzfacetten der Schüler:innen im Sportunterricht haben können. Andererseits werden im Beitrag bisherige Erkenntnisse sowie offene Felder in Theorie und Empirie hinsichtlich der Aus- und Fortbildung diagnostischer Kompetenzen von (angehenden) Sportlehrkräften vorgestellt.

Schlagwörter: diagnostische Kompetenzen, professionelle Handlungskompetenzen, Sportlehrkräfteprofessionalisierung, pädagogische Diagnose, Sportunterricht

Diagnostic Competencies from a Sport Pedagogical Perspective in Theory and Empirical Research

Abstract: Diagnostic competencies can be understood as part of the professional competencies of physical education (PE) teachers. They play a central role in designing physical education lessons with an impetus for student learning. However, empirical findings show that PE teachers at all types of schools have difficulties to accurately assess different learning prerequisites of their students (e.g. motor and motivational prerequisites) as well as the characteristics of tasks (e.g. difficulty). Fostering diagnostic competencies of PE teachers is necessary and should start at higher education. On the one hand, this article conceptualises on a theoretical level the current discourse on diagnostic competencies of PE teachers and structures it in a framework with reference to the empirical approach. The framework presents various facets of diagnostic competencies that can have an influence on student learning in PE. On the other hand, this article presents empirical findings as well as open fields in theory and empirical research with regard to fostering diagnostic competencies of (pre-service) PE teachers.

Keywords: diagnostic competencies, professional competence, physical education teacher education, assessment literacy, physical education

* Die Autor:innen waren gleichermaßen an der Erstellung dieses Beitrags beteiligt. Sie teilen sich die Erstautor:innenschaft.

Diagnostische Kompetenzen aus sportpädagogischer Perspektive in Theorie und Empirie

1 Problemstellung

Sportlehrkräfte sehen sich in ihrem Berufsalltag mit vielfältigen Aufgaben konfrontiert (z. B. Unterrichten, Erziehen, Beurteilen und Innovieren; KMK, 2022), die u. a. durch den leitenden Bildungsanspruch des Sportunterrichts konkretisiert werden. So wird in den Kernlehrplänen der Anspruch nach einem kompetenzorientierten, erziehenden Sportunterricht formuliert, wonach im Sportunterricht sowohl die Vermittlung von Bewegungskompetenzen als auch die Persönlichkeitsentwicklung intendiert werden sollen (u. a. MSW NRW, 2021; Neuber, 2020). Damit dies gelingt, sollten Sportlehrkräfte ihren Sportunterricht adaptiv gestalten, um fachliche Lernprozesse individuell anzuregen (Bräutigam, 2015) und so einen für die Schüler:innen lernwirksamen Sportunterricht zu gestalten. Hierfür müssen sie einerseits geeignete Unterrichtsmaßnahmen auswählen und im Unterricht einsetzen können (didaktische Kompetenzen) und andererseits die Lern- und Leistungsvoraussetzungen sowie die Lernprozesse ihrer Schüler:innen angemessen erfassen und diese im Hinblick auf die angestrebten Lernziele sinnvoll beurteilen können. Letzteres wird als diagnostische Kompetenzen bezeichnet (vgl. adaptive Lehrkompetenz bei Brühwiler, 2014).

Aus Studien anderer Fächer ist bekannt, dass sich hohe diagnostische Kompetenzen in Verbindung mit hohen didaktischen Kompetenzen positiv auf das Lernen von Schüler:innen auswirken (Brühwiler, 2014). Demnach scheinen diagnostische Kompetenzen als »Katalysatorvariable« positiv auf die Schulleistungen der Schüler:innen zu wirken (u. a. Helmke, 2017, S. 130).

Vor diesem Hintergrund sind in den letzten Jahren verschiedene sportpädagogische/-didaktische Arbeiten entstanden. Hierbei zeigt sich, dass die Qualität diagnostischer Urteile im Sportunterricht bestenfalls moderat zu sein scheint (Borghouts et al., 2017). Im Konkreten haben Sportlehrkräfte aller Schulstufen Schwierigkeiten, motorische und motivationale Lernvoraussetzungen (z. B. Ferrari et al., 2022; Niederkofler et al., 2018; Niederkofler, 2023) oder das Fähigkeitsselbstkonzept von Kindern akkurat einzuschätzen (Seyda, 2018). Darüber hinaus ist es für Sportlehrkräfte herausfordernd, die Anforderungen in Aufgaben akkurat zu beurteilen (u. a. Hendricks, 2024; Leukel et al., 2023). Dadurch ist fraglich, ob sie in der Lage sind, Aufgaben auszuwählen, die an den Lern- und Leistungsvoraussetzungen der Schüler:innen orientiert sind.

Die Förderung diagnostischer Kompetenzen scheint daher eine wichtige und notwendige Voraussetzung zu sein, um das professionelle Handeln einer Sportlehrkraft zu verbessern. Dem Studium wird hierfür eine hohe Bedeutung beigemessen (Röger-Offergeld et al., 2021). Aus diesem Grund ist es sinnvoll, wenn Förder- und Unter-

stützungsmaßnahmen für den Erwerb diagnostischer Kompetenzen bereits in der universitären Ausbildung, d. h. in der Sportlehrkräftebildung ansetzen (Latzko & Gottlebe, 2022). Die zentralen Fragen dieses Hauptbeitrags, inwiefern diagnostische Kompetenzen bei (angehenden) Sportlehrkräften angebahnt und gefördert werden können, wurden bislang jedoch sowohl in der auf Sportunterricht bezogenen als auch fachübergreifenden Professionsforschung nicht hinreichend beantwortet. An dieser Stelle setzt dieser Beitrag an, indem der aktuelle Diskurs um diagnostische Kompetenzen sowie offene Felder in Theorie und Empirie von Sportlehrkräften dargestellt und strukturiert werden. Der Beitrag verfolgt das Ziel, Möglichkeiten für die Aus- und Fortbildung diagnostischer Kompetenzen von (angehenden) Sportlehrkräften aufzuzeigen.

2 Theoretische Einführung in diagnostische Kompetenzen

Die eingangs dargestellte Forschungslücke resultiert u. a. daraus, dass in der (Sport-)Lehrkräftebildungsforschung vielfältige Auffassungen darüber vorliegen, was unter diagnostischen Kompetenzen genau zu verstehen ist (zsf. von Aufschnaiter et al., 2015). Dies zeigt sich exemplarisch in der Verwendung unterschiedlicher Begriffe (z. B. diagnostische Kompetenz im Singular bei Schrader, 2013 oder diagnostische Kompetenzen im Plural bei Hendricks, 2025). Diagnostische Kompetenz beschreibt nach Weinert (2000, S. 14) ein »Bündel von Fähigkeiten, um den Kenntnisstand, die Lernfortschritte und die Leistungsprobleme der einzelnen Schüler sowie die Schwierigkeit verschiedener Lernaufgaben im Unterricht fortlaufend beurteilen zu können, sodass das didaktische Handeln auf diagnostischen Einsichten aufgebaut werden kann«. In einer anderen Beschreibung werden diagnostische Kompetenzen als Diagnosekompetenz beschrieben, die als die Fähigkeit von Lehrkräften verstanden wird, »nach festgelegten Kriterien angemessene Urteile über das Lern- und Leistungsverhalten ihrer Schüler:innen abzugeben« (Paradies et al., 2007, S. 55). Schrader (2013) beschreibt diagnostische Kompetenz als die Fähigkeit, Merkmale von Personen und das Lernmaterial korrekt zu beurteilen. An diesen Beispielen zeigt sich beispielhaft, dass diagnostische Kompetenzen einerseits auf die Beurteilung von Personenmerkmalen (z. B. kognitive oder motorische Fähigkeiten, Fähigkeitsselbstkonzept) beschränkt werden (u. a. Anders et al., 2010; Helmke, 2017; Karst, 2009; Paradies et al., 2007). Andererseits umfassen sie in einer weiten Auslegung auch die Beurteilung des Lernmaterials (Karst, 2009; McElvany et al., 2009; Schrader, 2013).

Neben der bestehenden begrifflichen Unschärfe¹ werden diagnostische Kompetenzen zudem unterschiedlich im Diskurs um die professionellen Handlungskompetenzen von Lehrkräften eingeordnet. Professionelle Handlungskompetenzen beruhen auf dem kompetenztheoretischen Ansatz zur Beschreibung der Professionalität von (Sport-)Lehrkräften, in dem die zur Bewältigung der Anforderungen im Lehrkräfteberuf (z. B. Unterrichten, Erziehen, Diagnostizieren) notwendigen Kompetenzen beschrieben werden (Baumert & Kunter, 2006). In der deutschsprachigen Sportlehrkräftebildungsforschung haben insbesondere die Konzeptionen professioneller Handlungskompetenzen aus der COACTIV-Studie eine hohe Aufmerksamkeit erfahren (u. a. Baumert & Kunter, 2006; Brunner et al., 2011; Krauss et al., 2004). Hiernach werden professionelle Handlungskompetenzen unterteilt in das Professionswissen, Überzeugungen/Werthaltungen, motivationale Orientierungen und selbstregulative Fähigkeiten. Fachübergreifend wird betont, dass Lehrkräfte mit fortschreitender Expertise auf ein stark ausgeprägtes Professionswissen zurückgreifen (zsf. Heemsoth, 2016). Es umfasst das Fachwissen, das fachdidaktische Wissen, das pädagogische Wissen sowie das Organisations- und das Beratungswissen (Baumert & Kunter, 2006; Shulman, 1987). Diagnostische Kompetenzen werden in den unterschiedlichen Modellierungen professioneller Handlungskompetenzen sowohl fachspezifisch dem fachdidaktischen Wissen (u. a. Krauss et al., 2004) und dem allgemeinen pädagogischen Wissen (Baumert & Kunter, 2006) als auch fachübergreifend als Teil des fachdidaktischen und des pädagogischen Wissens (Brunner et al., 2011) zugeordnet. Daraus ergeben sich unterschiedliche Auffassungen darüber, inwiefern diagnostische Kompetenzen als fachunspezifische oder fachspezifische Kompetenzen betrachtet werden sollten. Je nach Auslegung ergeben sich unterschiedliche Konsequenzen für die Lehrkräftebildung.

Die Vielfalt der theoretischen Herangehensweisen zeigt sich auch in den unterschiedlichen theoretischen Modellierungen, in denen diagnostische Kompetenzen bspw. in produktorientierten Ansätzen als Urteilsakkuratheit (z. B. Südkamp et al., 2017) oder in prozessorientierten Ansätzen als diagnostischer Prozess in Form von Prozessmodellen dargestellt werden (u. a. Herppich et al., 2017; Klug et al., 2013; Loibl et al., 2020). In *produkt- bzw. genauigkeitsorientierten Ansätzen* wird unter-

- 1 Weitere Begriffe finden sich im internationalen Diskurs. Hier werden diagnostische Kompetenzen mit dem Begriff *assessment literacy* umschrieben, der die Qualität und Wirksamkeit der Beurteilung und Bewertung betrachtet. Damit sind alle Überzeugungen, Kenntnisse und Praktiken gemeint, die u. a. von Sportlehrpersonen für die Verbesserung des Lernens von Schüler:innen und deren Leistungen eingesetzt werden (AIESEP Position paper, 2020). In Anlehnung an Hay und Penney (2012) besteht sie aus vier interdependenten Elementen: Verständnis, Anwendung, Interpretation sowie kritische Auseinandersetzung mit der Beurteilung und Bewertung. »Die Beurteilungs- und Bewertungskompetenz der Lehrpersonen ist damit eine wichtige Voraussetzung für die Beurteilungsqualität« (AIESEP Positions Papier, 2020, deutsche Übersetzung, S. 5). Sie sind zu fördern, um gültige und akkurate Urteile über den Lernprozess und deren Ergebnis von Schüler:innen zu fällen. Während es erste internationale Befunde zur Qualität der Beurteilungen gibt, erhoben durch Lehrerfragebögen (z. B. Borghouts et al., 2017), scheinen internationale Studien zur Akkuratheit und Wirksamkeit der Beurteilungen im Sportunterricht noch zu fehlen.

sucht, wie akkurat Lehrkräfte die Merkmale eines Beurteilungsgegenstands einschätzen können. Die Urteilsakkuratheit meint hierbei die Übereinstimmung zwischen den Lehrkrafturteilen und den (objektiv) gemessenen Merkmalsausprägungen von Schüler:innen oder von Aufgaben. Um die Urteilsakkuratheit bestimmen zu können, wird zumeist auf die vereinfachte Komponentenanalyse zurückgegriffen, die zwischen Niveau-, Differenzierungs- und Rangkomponente unterscheidet (Schrader, 2014). Die *Niveauelemente* gibt an, inwiefern Lehrkräfte das mittlere Leistungsniveau der Klasse oder einzelner Schüler:innen beurteilen können. Die *Differenzierungskomponente* beschreibt, inwiefern Lehrkräfte Kenntnis über die Streuung zwischen den Schüler:innenleistungen haben, d. h. ob sie die Streuung der Leistung innerhalb der Klasse korrekt beurteilen können. Die *Rangkomponente* gibt an, inwiefern die Lehrkräfte ihre Schüler:innen in Bezug auf ein bestimmtes Merkmal in einer Rangreihe ordnen können (Schrader & Helmke, 1987). In *prozessorientierten Ansätzen* wird der Beurteilungsprozess untersucht (u. a. Herppich et al., 2017; Hoth et al., 2016; Klug et al., 2012; Klug et al., 2013; Langer, 2023). Auf diese Weise kann z. B. festgestellt werden, wie die Nutzung von Beurteilungsmerkmalen den Beurteilungsprozess beeinflusst (Klug et al., 2012).

Die Frage, welche Fähigkeiten Sportlehrkräfte für ein situationsangepasstes Handeln im Sportunterricht brauchen, werden auch in Arbeiten zum professionellen Blick aufgeworfen (u. a. Reuker & Schicklinski, 2024). Der professionelle Blick beschreibt die Fähigkeit von Lehrkräften zur selektiven Aufmerksamkeit, einer pädagogisch fundierten Deutung des Wahrgenommenen und einer daran anschließenden situationsspezifischen Entscheidungsfindung (Reuker & Schicklinski, 2024). Er steht damit in einem engen Zusammenhang mit prozessorientierten Ansätzen diagnostischer Kompetenzen. Trotz dieser Parallelen ist der professionelle Blick nur bedingt an den Diskurs um die professionellen Handlungskompetenzen anschlussfähig. Er beruft sich in aktuellen Auslegungen (Reuker & Schicklinski, 2024) auf eine sozial-kulturelle Perspektive, die sich im Diskurs um die Professionalität eher den berufsbiografischen oder strukturtheoretischen Ansätzen zuordnen lässt.

Diese drei exemplarischen Problemfelder² zeigen, dass diagnostische Kompetenzen unterschiedlich theoretisch beschrieben werden. Die zeitlich vergleichsweise erst spät einsetzende sportpädagogische/-didaktische Auseinandersetzung mit diagnostischen Kompetenzen knüpft daran an und folgt z. T. dieser unsystematischen Herangehensweise. Eine systematische Aufarbeitung des sportpädagogischen/-didaktischen Diskurses, mit der bestehende Erkenntnisse und konkrete Forschungslücken aufgezeigt werden, bleibt bislang aus. Durch die unterschiedlichen Begriffe, die nicht einheitliche Einordnung in den Diskurs um die professionellen Handlungskompetenzen und

2 Mit Problemfeldern sind in diesem Beitrag die begriffliche Unschärfe, die uneinheitliche Verortung in den Diskurs um die professionellen Handlungskompetenzen sowie die unterschiedlichen Modellierungen gemeint. Auf eine tiefergehende Aufarbeitung wird an dieser Stelle verzichtet und auf bereits bestehende Analysen verwiesen (u. a. Hendricks, 2024).

die vielfältigen Modellierungen ist es zudem nur schwer möglich, Konsequenzen für die Sportlehrkräftebildung abzuleiten. Hierzu bedarf es zunächst einer theoretischen Einordnung.

3 Theoretische Einordnung diagnostischer Kompetenzen für das Themenheft

Angesichts der im vorangegangenen Kapitel aufgeführten Problemfelder erscheint es notwendig, einen theoretischen Rahmen für das Themenheft festzulegen, der als Grundlage zur Einordnung des bestehenden Forschungsstandes dient und entlang dessen Forschungslücken konkret benannt werden können. Für die theoretische Einordnung diagnostischer Kompetenzen für das Themenheft ist Folgendes leitend: Der Fokus liegt auf der Aus- und Fortbildung von (angehenden) Sportlehrkräften. Es bedarf daher ein Modell, in dem verschiedene Ansatzpunkte für die Sportlehrkräftebildung abgebildet werden. Das Modell sollte so breit angelegt sein, dass verschiedene Sichtweisen auf diagnostische Kompetenzen möglich sind. Da Modellierungen diagnostischer Kompetenzen dadurch abhängen, welches begriffliche Verständnis angelegt wird (u. a. von Aufschnaiter, 2015), müssen zunächst die für das Verständnis notwendigen zentralen Begriffe beschrieben werden. Darauf aufbauend wird eine Modellierung diagnostischer Kompetenzen eingeführt, die für die Einordnung des bestehenden Forschungsstands sowie für die Rahmung der in diesem Themenheft zusammengestellten Beiträge leitend ist.

3.1 Zentrale Begriffe diagnostischer Kompetenzen

Pädagogische Diagnostik

Der Ausgangspunkt der im Themenheft zusammengefassten Arbeiten ist die pädagogische Diagnostik im schulischen Kontext. Pädagogische Diagnostik meint alle Tätigkeiten, mit denen die »[...] Voraussetzungen und Bedingungen planmäßiger Lehr-Lernprozesse ermittelt, Lernprozesse analysiert und Lernergebnisse festgestellt werden, um individuelles Lernen zu optimieren« (Ingenkamp & Lissmann, 2008, S. 13). Darauf aufbauend kann Unterricht geplant und es können pädagogisch-didaktische Anschlusshandlungen (z. B. Unterrichtsaktivitäten oder Fördermaßnahmen) an den Bedürfnissen der Schüler:innen ausgerichtet werden. Das Verständnis darüber, was unter Diagnostik genau zu fassen ist, unterscheidet sich teils deutlich voneinander (von Aufschnaiter et al., 2015; Schrader 2014). Von Aufschnaiter et al. (2015) stellen vier Arten der Diagnostik heraus: Unter dem Begriff *Statusdiagnostik* werden Maßnahmen zur Erfassung von aktuell vorliegenden Merkmalsausprägungen (z. B. hoch oder niedrig entwickelte motorische Fähigkeiten bei Schüler:innen) verstanden. Sie kann zu einem bestimmten Zeitpunkt im Lernprozess erfolgen (z. B. am Anfang oder am Ende). Die *Prozessdiagnostik* zielt darauf ab, den Lösungsprozess bei der Bearbeitung von Aufgaben aufzudecken. *Veränderungsdiagnostik*

umfasst Maßnahmen zur Erfassung von Veränderungen von Merkmalen und ihrer Ausprägung. Um Veränderungen aufzuzeigen, werden die Merkmale mehrmals erfasst (z. B. zur Überprüfbarkeit von Interventionen) und es werden mindestens zwei Urteile miteinander verglichen. Die *Verlaufsdiagnostik* bezieht sich auf Prozesse, die zwischen zwei Messzeitpunkten ablaufen. Sie ergänzt die Veränderungsdiagnostik somit um Handlungs- und Denkprozesse. Auf diese Weise kann herausgestellt werden, wodurch sich Veränderungen zwischen zwei Messzeitpunkten ergeben (von Aufschnaiter et al., 2015). Eine weitere Unterteilung von Diagnostik findet sich im angelsächsischen Sprachraum. Diagnostik kann einerseits als abschließendes Ergebnis am Ende eines Lernanlasses (summativ) und andererseits im Verlauf des Lernprozesses (formativ) erfolgen (Schrader, 2014). Die summativ Diagnostik weist eine Ähnlichkeit zur Statusdiagnostik; die formative Diagnostik zur Prozessdiagnostik auf. In der Unterrichtspraxis hat die formative Diagnostik eine höhere Relevanz, weil eine Hauptaufgabe von (Sport-)Lehrkräften darin besteht, Lernprozesse anzuregen und zu begleiten (Langfeld, 2014; Langer, 2023). Diagnostik findet im Unterricht zumeist zirkulär, d. h. in mehreren unterschiedlichen Situationen statt (Niederkofler et al., 2018). Auf diese Weise können diagnostische Informationen schrittweise gesammelt werden. Die mit der Diagnostik gewonnen Informationen können sehr vielfältig sein. Sie können u. a. Merkmale der Schüler:innen (z. B. kognitive, motorische, soziale oder motivational-emotionale Kompetenzen) oder Informationen über Aufgabenmerkmale (z. B. die motorische Schwierigkeit) umfassen (Hendricks, 2024; van Ophuysen, 2010). Diese Merkmale werden als lernrelevante Merkmale bezeichnet. Ob Sportunterricht lernförderlich sein kann, ist eng damit verbunden, inwiefern Sportlehrkräfte Informationen über diese lernrelevanten Merkmale verarbeiten, d. h. inwiefern sie angemessene Diagnosen stellen können.

Diagnosen

Diagnosen sind im schulischen Kontext Urteile über lernrelevante Merkmale, die entlang von Kategorien, Begriffen oder Konzepten getätigt werden (Betsch et al., 2011; Helmke, 2017). Bei Urteilen schreiben (Sport-)Lehrkräfte Schüler:innen in Bezug auf lernrelevante Merkmale eine Ausprägung zu (z. B. hohe oder niedrige Ausprägung der motorischen Fähigkeiten). Die Urteile können sich hierbei sowohl auf die Erfassung von Merkmalen und deren Ausprägung zu einem bestimmten Zeitpunkt (i. S. einer Statusdiagnostik) als auch auf die Erfassung von Lernverläufen und -geschwindigkeiten beziehen (i. S. einer Verlaufs- oder Veränderungsdiagnostik). Im (sport-)unterrichtlichen Kontext reicht die Spannweite von informellen bis zu formellen Diagnosen (Hascher, 2008). Sie unterscheiden sich darin, dass informelle Diagnosen nicht standardisiert sind, sondern subjektiv und eher beiläufig erfolgen, während formelle Diagnosen durch standardisierte methodisch kontrollierte Prozesse gekennzeichnet sind (z. B. standardisierte und empirisch geprüfte

Erhebungsinstrumente, wie das MOBAK-Instrument zur Erhebung motorischer Basiskompetenzen; Herrmann & Seelig, 2020). Semiformelle Diagnosen erfolgen weniger standardisiert als formelle Diagnosen, sind aber dennoch gezielt (Hesse & Latzko, 2011). Im Berufsalltag von Sportlehrkräften treten verschiedene Situationen auf, in denen Diagnosen relevant sind.

Diagnosesituationen

Diagnosesituationen sind Situationen im Berufsalltag von Lehrkräften, in denen die Lehrkräfte Diagnosen tätigen. Sie kennzeichnen sich durch das Zusammenspiel zwischen den zu beurteilenden Merkmalen und den zeitlich-räumlichen Bedingungen, unter denen Diagnosen getätigt werden. Karst (2012) stellt drei idealtypische Situationen heraus, die sich aus der Perspektive auf Lernende ergeben, die Lehrende bei der Einschätzung ihrer Schüler:innen einnehmen. In *klassenbezogenen Situationen* wird die Klasse als Gesamtes betrachtet (z. B. bei der Frage, ob ein Thema bereits neu eingeführt werden kann, oder erst zuvor notwendige Inhalte erarbeitet werden müssen). In *schüler:innenglobalen Situationen* werden die Schüler:innen in Gruppen aufgeteilt. Die Einschätzung bezieht sich dann auf die einzelnen Gruppen, wodurch eine globale Sicht auf die Schüler:innen der Gruppe eingenommen wird. Bei *schüler:innenspezifischen Situationen* werden einzelne Schüler:innen beurteilt (z. B. im Rahmen individuell förderlicher Unterrichtsmaßnahmen) (Karst, 2012). Darüber hinaus unterscheiden sich die Diagnosesituationen hinsichtlich ihrer zeitlichen Abfolge zum Unterrichtsgeschehen (vor bzw. während des Unterrichts). In der Unterrichtsvorbereitung werden Diagnosen meist reflektiert und ohne Zeitdruck erstellt (Ostermann et al., 2015), wohingegen Diagnosen während des Unterrichts unter einem hohen Zeitdruck erfolgen. Dies betrifft im besonderen Maße sportunterrichtliche Diagnosesituationen, da sich die Schüler:innen durch die zeitlich-räumliche Organisationsstruktur meist frei in der Sporthalle bewegen (Neuber, 2020). Darüber hinaus weisen Bewegungen häufig einen prozessorientierten Charakter auf (Neumaier, 2009), weshalb ein zusätzlicher Zeitdruck entsteht. Es kann davon ausgegangen werden, dass Diagnosen während des Sportunterrichtsgeschehens meist informell oder semiformell sind. Um in den unterschiedlichen Diagnosesituationen angemessene Diagnosen stellen zu können, benötigen Sportlehrkräfte diagnostische Kompetenzen.

3.2 Modellierung diagnostischer Kompetenzen für das Themenheft

Im vorliegenden Beitrag wird auf das erweiterte Verständnis diagnostischer Kompetenzen zurückgegriffen, wonach die Fähigkeit verstanden wird, Merkmale von Personen und das Lernmaterial korrekt zu beurteilen (u. a. Schrader, 2013). Diagnostische Kompetenzen werden in diesem Beitrag als fachspezifische Kompetenzen von (Sport)Lehrkräften aufgefasst. Mit Bezug zur kompetenzorientierten Sportlehrkräftebildungsforschung (Baumgartner, 2022) können sie in einem heuristischen

Rahmenmodell modelliert werden (Abb. 1). Das KompeSport-Modell (Baumgartner, 2022) basiert auf dem Kompetenzkontinuum (Blömeke et al., 2015), welches in Bezug auf diagnostische Kompetenzen bereits adaptiert wurde (in Mathematik bspw. Hoth et al., 2016; im Sport u. a. Hendricks, 2024; 2025).

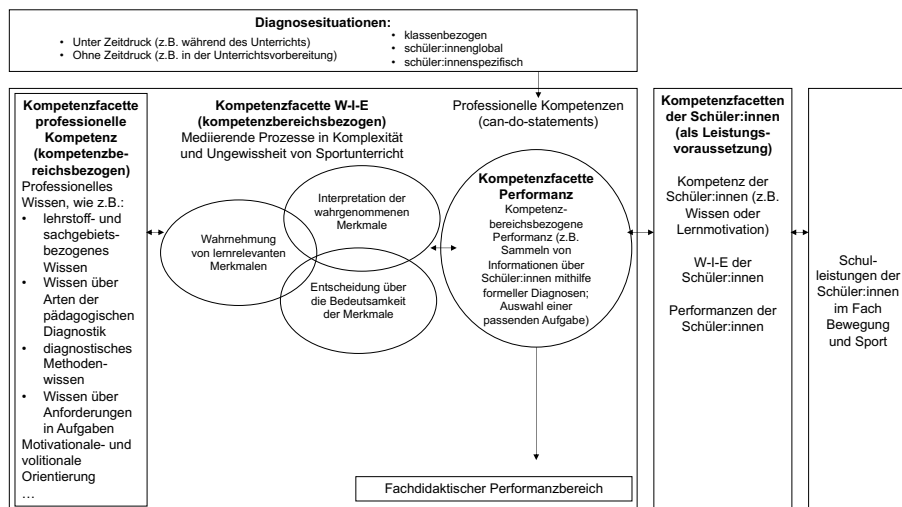


Abb. 1: Diagnostische Kompetenzen als Teil des KompeSport-Modells (angelehnt an Baumgartner, 2022, S. 41; Blömeke et al., 2015)

Ausgangspunkt zur Beschreibung professioneller Kompetenzen sind im KompeSport-Modell (Baumgartner, 2022) berufliche Anforderungssituationen. Diese umfassen in Bezug auf diagnostische Kompetenzen die drei idealtypischen Diagnosesituationen in Anlehnung an Karst (2012). Diagnosen können während des Sportunterrichts unter Zeitdruck oder in der Unterrichtsvorbereitung ohne größeren Zeitdruck erfolgen (vgl. Kapitel 3.1). Das Modell unterteilt diagnostische Kompetenzen in drei Kompetenzfacetten (Abb. 1). Hierdurch kann der Fokus entweder auf die zugrundeliegenden Dispositionen (z. B. Professionswissen), die informationsverarbeitenden Prozesse (Interpretation, Wahrnehmung, Entscheidung) oder die Performanz, d.h. das Verhalten in realen Anforderungssituationen gelegt werden (Baumgartner, 2022; Blömeke et al., 2015).

Die *Kompetenzfacette professionelle Kompetenz* setzt sich aus dem Professionswissen, Überzeugungen/Werthaltungen, motivationale Orientierungen und selbstregulative Fähigkeiten zusammen, von denen angenommen wird, dass sich eine hohe Ausprägung positiv auf die Performanz auswirkt (Baumgartner, 2022). Mit der Einordnung diagnostischer Kompetenzen als fachspezifische Kompetenzen folgen wir der Annahme, dass die Beurteilung fachlicher Kompetenzen von Schüler:innen, ihre Lernprozesse sowie die fachbezogenen Anforderungen in Aufgaben mit dem fachdi-

daktischen Wissen zusammenhängt (Krauss et al., 2004). In Bezug auf diagnostische Kompetenzen ist damit insbesondere lehrstoff- und sachgebietsbezogenes Wissen, diagnostisches Methodenwissen oder Wissen über Aufgaben gemeint (Schrader, 2017). Darüber hinaus umfasst die Kompetenzfacette professionelle Kompetenz ebenso Wissen über die vier Arten der Diagnostik oder methodische Verfahren der semiformalen oder formalen Diagnosen (Kapitel 3.1). Dieses Wissen kann im KompeSport-Modell als Disposition der Lehrkräfte aufgefasst werden. Sportlehrkräfte müssen darüber hinaus auch gewillt sein, Diagnosen zu stellen (Schrader, 2017). Hierfür benötigen sie u. a. motivationale Orientierungen. Die Kompetenzfacette professionelle Kompetenz stellt demnach eine notwendige Voraussetzung für akkurate Urteile dar. Hierfür müssen Informationen über die beurteilten Merkmale in einer spezifischen Diagnosesituation angemessen verarbeitet werden (Hoth et al., 2016).

Die Informationsverarbeitungsprozesse werden in der *Kompetenzfacette W-I-E* beschrieben. Die Wahrnehmungs-, Interpretations- und Entscheidungsprozesse vermitteln als medierende Prozesse zwischen professioneller Kompetenz und Performanz. Sportlehrkräfte müssen auf Grundlage ihres professionellen Wissens lernrelevante Merkmale wahrnehmen und interpretieren. Darauf aufbauend kann eine Entscheidung hinsichtlich unterrichtlicher Maßnahmen getroffen werden.

Die *Kompetenzfacette Performanz* beschreibt das (beobachtbare) diagnostische Handeln in einer konkreten Diagnosesituation. Die Performanz zeigt sich z. B. in der Anwendung semiformaler Diagnosen, der Auswahl von Aufgaben oder in akkuraten Urteilen.

Dem Modell folgend kann angenommen werden, dass diagnostische Kompetenzen neben weiteren Kompetenzbereichen der Lehrkraft (z. B. im didaktischen und fachwissenschaftlichen Bereich) einen Einfluss auf die Kompetenzentwicklung der Schüler:innen haben. Die Unterteilung diagnostischer Kompetenzen in die drei Kompetenzfacetten bietet die Möglichkeit, die teils deutlich unterschiedlichen theoretischen und empirischen Herangehensweisen in der sportpädagogischen/-didaktischen Erforschung diagnostischer Kompetenzen in einem heuristischen Modell systematisch zu ordnen und Zusammenhänge herauszustellen. So lassen sich beispielsweise die produkt- bzw. genauigkeitsorientierten und prozessorientierten Ansätze im Modell zuordnen (die Urteilsakkuratheit verstanden als Ergebnis der Performanz und der diagnostische Prozess verstanden als die situationsspezifischen Fähigkeiten).

Die im Modell (Abb. 1) dargestellten Zusammenhänge sind zwar bislang nicht ausreichend empirisch geprüft (für das ursprüngliche KompeSport-Modell u. a. Baumgartner et al., 2023), es deutet sich jedoch an, dass die Ausprägung professioneller Kompetenzen in der Sportlehrkräftebildung empirisch untersucht und in Interventionen gezielt verbessert werden können (Baumgartner, 2017; Baumgartner et al., 2021). Zudem können über das Modell Aussagen abgeleitet werden, was eine Sport-

lehrkraft mit hohen diagnostischen Kompetenzen auszeichnet: Sie muss über Dispositionen (professionelles Wissen, motivationale Orientierungen etc.) verfügen. Diese sind die Grundlage für die Wahrnehmungs-, Interpretations- und Entscheidungsprozesse, die wiederum die Performanz beeinflussen. Hohe diagnostische Kompetenzen zeichnen sich demnach nicht nur durch gut ausgeprägte einzelne Kompetenzfacetten aus, sondern durch ihr Zusammenspiel in spezifischen Diagnosesituationen.

Für die Analyse des Forschungsstands liegt der Fokus auf Arbeiten, die sich im Kontext der professionellen Handlungskompetenzen von Sportlehrkräften verorten lassen und eine sportpädagogische/-didaktische Auslegung fokussieren. In diesem Feld sind empirische Arbeiten entstanden, die sich hinsichtlich ihrer theoretischen sowie empirischen Herangehensweise teils deutlich voneinander unterscheiden und diagnostische Kompetenzen unter verschiedenen Perspektiven in den Blick nehmen.

4 Empirische Befunde zu diagnostischen Kompetenzen von (angehenden) Sportlehrkräften

Die zentralen Fragen dieses Beitrags lauten, inwiefern diagnostische Kompetenzen bei (angehenden) Sportlehrkräften angebahnt und gefördert werden können und wie dies in der Sportlehrkräftebildung erfolgen kann. Zur Beantwortung dieser Fragen bietet das angepasste KompeSport-Modell (Abb. 1) einen heuristischen Rahmen. Die einzelnen Kompetenzfacetten stellen Ansatzpunkte zur konkreten Förderung dar (z. B. die Vermittlung von Wissen oder die Schulung der Wahrnehmungs-, Interpretations- und Entscheidungsprozesse). Darüber hinaus lässt sich durch die angenommenen Wirkungszusammenhänge die Verknüpfung zwischen den einzelnen Kompetenzfacetten beschreiben (z. B. der Zusammenhang zwischen einem gut ausgeprägten professionellen Wissen und der Urteilsakkuratheit). Um Maßnahmen zur Förderung und Anbahnung diagnostischer Kompetenzen gezielt an den Lern- und Leistungsvoraussetzungen der (angehenden) Sportlehrkräfte ausrichten zu können, ist es zunächst wichtig, den bestehenden Forschungsstand zu analysieren.

Für die nachfolgende Analyse wurden aufbauend auf dem KompeSport-Modell folgende Teilfragen abgeleitet:

- (1) Welche Erkenntnisse über die Bedeutung der Kompetenzfacetten von (angehenden) Sportlehrkräften bei der Diagnose von lernrelevanten Merkmalen lassen sich aus den bisher veröffentlichten Forschungsarbeiten extrahieren?
- (2) Welche Erkenntnisse über Maßnahmen zur Förderung der einzelnen Kompetenzfacetten von (angehenden) Sportlehrkräften bei der Diagnose von lernrelevanten Merkmalen lassen sich aus den bisher veröffentlichten Forschungsarbeiten extrahieren?

Es wurden nur Studien in englischer und deutscher Sprache in nationalen und internationalen Datenbanken (Bisp Surf, PEDOCs, ERIC; SCOPUS) recherchiert, die diagnostische Kompetenzen sportunterrichtender Lehrkräfte sowie Sportlehramts-

studierender und ihre Schüler:innen ab der 1. Klasse untersuchen. Studien, die z. B. die diagnostischen Kompetenzen von Lehrpersonen im Kindergarten erhoben haben, wurden ausgeschlossen. Eingeschlossen wurden alle Studien, die nach einem peer-review Verfahren oder in Monographien veröffentlicht wurden (Kategorien A und B; dvs, 2011). Die Studien wurden durch zwei Wissenschaftler:innen im konsensuellen Austausch den Kompetenzfacetten zugeordnet und die vorliegenden Erkenntnisse extrahiert.

Teilforschungsfrage 1: Welche Erkenntnisse über die Bedeutung der Kompetenzfacetten von (angehenden) Sportlehrkräften bei der Diagnose von lernrelevanten Merkmalen lassen sich aus den bisher veröffentlichten Forschungsarbeiten extrahieren?

In den bisher veröffentlichten Forschungsarbeiten lassen sich insbesondere drei Forschungsstränge über die Bedeutung der Kompetenzfacetten von (angehenden) Sportlehrkräften für die Diagnose von lernrelevanten Merkmalen identifizieren. Im ersten Forschungsstrang steht die grundsätzliche Frage im Fokus, inwiefern Sportlehrkräfte lernrelevante Merkmale von Schüler:innen (motorische Basiskompetenzen, Fähigkeitsselbstkonzept) akkurat einschätzen können (Urteilsakkuratheit als Ergebnis der Performanz). Zum anderen wurde in diesen Studien auch der Zusammenhang zwischen verschiedenen Berufsmerkmalen (Qualifikation, Berufsbzw. Unterrichtserfahrung) und der Performanz untersucht (z. B. Hendricks, 2024; Seyda, 2018). Dies stellt den zweiten Forschungsstrang dar. Die verschiedenen Berufsmerkmale stehen in einem engen Zusammenhang mit der Kompetenzfacette professionelle Kompetenz. In Arbeiten, in denen die Qualifikation untersucht wurde (Seyda, 2018), wird die Facultas als indirekter Indikator für das professionelle Wissen herangezogen. Studien, in denen die Berufserfahrung (im Sinne einer längeren Berufspraxis) untersucht wurde, gehen davon aus, dass die Berufserfahrung als Moderatorvariable einen indirekten Einfluss auf die Urteilsakkuratheit hat. Zwar seien nicht die Berufsjahre entscheidend, sondern die Ausprägung des fachdidaktischen Wissens (zsf. Hendricks, 2024). Dieses wird jedoch im Studium und in Fortbildungen erworben, weshalb davon ausgegangen werden kann, dass Berufserfahrung eine notwendige, aber keine hinreichende Bedingung für die Entwicklung der Kompetenzfacette professionelle Kompetenz darstellt. Der Zusammenhang zwischen der Kompetenzfacette W-I-E und der Performanz stellt den dritten Forschungsstrang dar.

Erster Forschungsstrang: Urteilsakkuratheit als Ergebnis der Performanz

Hinsichtlich des ersten Forschungsstrangs zeigt sich, dass Sportlehrkräfte aller Schulstufen Schwierigkeiten haben, motorische Lernvoraussetzungen von Kindern akkurat zu beurteilen (Niederkofler, 2023; Niederkofler et al., 2018). Für die Beurteilung motorischer Basiskompetenzen von Schüler:innen der Grundschule und der Sekundarstufe I durch Sportlehrpersonen haben Niederkofler et al. (2018) und Nie-

derkofler (2023) herausgefunden, dass Lehrkräfte die Rangfolge der Schulkinder mit einer mittleren bis schwachen Qualität einschätzen. Darüber hinaus ergaben die Ergebnisse zur Differenzierungskomponente, dass sowohl Lehrkräfte der Sekundarstufe (Niederkofler et al., 2018) als auch der Grundschule (Niederkofler, 2023) die leistungsbezogene Heterogenität unterschätzen. Lediglich in der Primarstufe (Niederkofler, 2023) deuten die Ergebnisse der Differenzierungskomponente im Kompetenzbereich *Etwas Bewegen* darauf hin, dass Grundschullehrkräfte gute Kenntnisse über die Heterogenität der motorischen Leistungen ihrer Grundschulkinder haben. Die Ergebnisse beider Studien zeigen zudem eine Überschätzung des absoluten motorischen Niveaus (Niveauelemente). Somit gehen Lehrpersonen möglicherweise von einer besseren motorischen Leistung aus, als es die Schulkinder in der Testsituation zeigen konnten (Niederkofler, 2023). Ähnliche signifikante Überschätzungstendenzen des Niveaus zeigen sich auch bei Grundschulsportlehrkräften im Hinblick auf das Fähigkeitsselbstkonzept von Kindern (Seyda, 2018). Die Ergebnisse zur Differenzierungskomponente zeigen zudem, dass die Grundschulsportlehrkräfte im Durchschnitt signifikant die Spannweite der Merkmalsausprägung in ihrer Klasse überschätzten. Des Weiteren konnten die untersuchten Sportlehrkräfte im Durchschnitt die individuellen Unterschiede des Fähigkeitsselbstkonzepts der Kinder ihrer Klasse nicht ausreichend akkurat beurteilen (Rangkomponente). Weitere Ergebnisse deuten zudem auf eine interindividuelle Streuung der Beurteilungen zwischen Sportlehrpersonen hin (u. a. Hendricks, 2024; Niederkofler et al., 2018; Seyda, 2018). Sportlehrpersonen sind entsprechend unterschiedlich gut darin, die Lernvoraussetzungen von Schüler:innen und die Aufgaben hinsichtlich ihrer Anforderungen zu diagnostizieren. Mit Blick auf die Beurteilung kognitiver und motorischer Anforderungen in Aufgaben hat Hendricks (2024) in einer Mixed-Methods Studie herausgefunden, dass die Mehrzahl der untersuchten Sportlehrpersonen im Vergleich zwischen zwei Aufgaben die jeweils anspruchsvollere Aufgabe identifizieren konnten. Jedoch merkt er kritisch an, dass eine normative Interpretation der Ergebnisse aufgrund fehlender Vergleichswerte schwierig ist und das Ergebnis in einem Studiendesign mit einer geringeren Ratewahrscheinlichkeit vermutlich anders ausfallen könnte.

Zweiter Forschungsstrang: Zusammenhang zwischen professioneller Kompetenz und Performanz

In Bezug auf den zweiten Forschungsstrang zeigt sich, dass die Berufsmerkmale (verstanden als Teil der Kompetenzfacette professionelle Kompetenz) nicht mit der Performanz zusammenhängen. So konnte Seyda (2018) zeigen, dass die Qualifikation von Grundschulsportlehrkräften (Fakultas/keine Fakultas) keinen signifikanten Einfluss für ein akkurates Lehrkrafturteil über das Fähigkeitsselbstkonzept von Schüler:innen hat. Auch zwischen der Berufserfahrung von Sportlehrkräften (an Haupt-, Real-, Sekundar- und Gesamtschulen) und der Beurteilungsleistung von kognitiven und motorischen Anforderungen bei Aufgaben besteht kein signifikan-

ter Zusammenhang (Hendricks, 2024). Ebenso zeigt sich kein Zusammenhang zwischen der Berufserfahrung von Sportlehrkräften und der Einschätzung motorischer Basiskompetenzen ihrer Schüler:innen (Niederkofler et al., 2018).

Dritter Forschungsstrang: Zusammenhang zwischen W-I-E und Performanz

Die Fragen, welche Informationen in diagnostischen Situationen wahrgenommen und wie diese im Hinblick auf das daraus entstehende Urteil verarbeitet werden, sind bislang noch wenig untersucht (Leukel et al., 2025). Die bisher veröffentlichten Forschungsarbeiten analysieren vorrangig, welche Beurteilungsformen verwendet werden (Performanz) und welche Schüler:innenmerkmale damit wahrgenommen und interpretiert werden können.

Die empirischen Befunde zeigen, dass Sportlehrkräfte meist messbare performative Leistungen ihrer Schüler:innen beurteilen. Es dominieren summative Bewertungen sowie Benotungen im Sportunterricht (Backman et al., 2023; Dinan-Thompson & Penney, 2015; Moura et al., 2021). Fragebogenerhebungen haben ergeben, dass formative Beurteilungen, wie z. B. den Schüler:innen durch die Verwendung von Rubriken einen Einblick in ihren Lernprozess geben, selten stattfindet (Slingerland et al., 2024). Quantitative Forschungen zeigen auch, dass die Beurteilungs- und Benotungspraktiken von Sportlehrpersonen meist intuitiv und subjektiv sind (Bezeau et al., 2023; Borghouts et al., 2017). Sportlehrkräfte verwenden für die Beurteilung motorischer Kompetenzen ihrer Schüler:innen selten evidenzbasierte Bewertungsinstrumente, sondern entwickeln oft ihre eigenen Beurteilungsmechanismen, wie Snow et al. (2025) mit einer qualitativen Interviewstudie herausgefunden haben. Für den kompetenzorientierten Sportunterricht in Deutschland haben Schittkowski et al. (2024) qualitativ herausgefunden, dass sich Sportlehrpersonen die Umsetzung eines kompetenzorientierten Sportunterrichts als herausfordernd wahrnehmen, aber auch eine Vielzahl von Strategien anwenden, um Transparenz in den Beurteilungen zu schaffen und Bewertungsfehler zu minimieren. Die untersuchten Sportlehrkräfte bewerten vorrangig summativ die motorischen Leistungen ihrer Schüler:innen und seltener ihre soziale Leistungen (z. B. für Mitarbeit, Unterstützung von Mitschüler:innen). Allerdings ist in internationalen Forschungen eine Tendenz erkennbar, dass bei den Lehr- und Bewertungspraktiken im Sportunterricht auch formative Beurteilungen Einzug erhalten (Bezeau et al., 2023; Borghouts et al., 2017; Chróinin & Cosgrave, 2013; Leirhaug & Annerstedt, 2016). Qualitative Befunde deuten darauf hin, dass (erfahrene) Sportlehrkräfte formative Beurteilungen im Sportunterricht umsetzen, wenn sie diese bei der Unterrichtsplanung berücksichtigen und wenn die Sportlehrkräfte eine positive Einstellung gegenüber formativen Beurteilungsformen haben bzw. entwickelt haben. Die Ergebnisse zeigen aber auch, dass Sportlehrkräfte fehlende Unterstützung und mangelnde Fähigkeiten als hinderlich bei einer Umsetzung formativer Beurteilungen wahrnehmen (Barrientos et al., 2023; Moura et al., 2021).

In einer qualitativen Studie wurden bei der Untersuchung der Verknüpfung zwischen didaktischen und diagnostischen Handlungen von Sportlehrkräfte der Grundschule auch die Entscheidungsprozesse (Kompetenzfacette W-I-E) berücksichtigt. Es zeigte sich, dass Verknüpfungen zwischen didaktischem und diagnostischem Handeln nur selten in der Planungsphase vorbereitet werden (z. B. Beobauungskriterien in der Planung auswählen, mit denen Schüler:innen während einer Stationsarbeit beurteilt werden können) (Langer, 2023). Weitergehend unterscheidet sich die Art und Weise, wie Grundschulsportlehrpersonen mit unterschiedlicher diagnostischer Kompetenz (i. S. der Urteilsakkuratheit berechnet über die Rangkomponente) diagnostische Handlungen mit didaktischen Entscheidungen verbinden. Beispielsweise schaffen sich Sportlehrpersonen mit hohen und moderaten diagnostischen Kompetenzen durch ihre strukturierten, didaktischen Handlungen Zeit für das Diagnostizieren im Unterricht, während sich bei Sportlehrpersonen mit niedrigen diagnostischen Kompetenzen verschiedene Unterrichtstätigkeiten überlagern (Langer, 2023).

Leukel et al. (2023) untersuchten, inwiefern sich Sportlehrkräfte und Trainer:innen im Turnen hinsichtlich der kognitiven Prozesse bei der Beurteilung von Aufgaben und ihrer Urteilsgenauigkeit unterscheiden. Für die Untersuchung haben beide Gruppen Bewegungen im Turnen bei Schüler:innen beobachtet und mittels Event Segmentation strukturiert und die Urteile erklärt. Im Gegensatz zu den Trainer:innen wiesen Sportlehrkräfte eine signifikant niedrigere Urteilsgenauigkeit auf und stuften den Schweregrad der Bewegungsfehler weniger schwerwiegend ein. Darüber hinaus verwiesen die Sportlehrkräfte in ihren Begründungen seltener auf relevante kinematische Merkmale der Aufgabe.

Teilforschungsfrage 2: Welche Erkenntnisse über Maßnahmen zur Förderung der einzelnen Kompetenzfacetten von (angehenden) Sportlehrkräften bei der Diagnose von lernrelevanten Merkmalen lassen sich aus den bisher veröffentlichten Forschungsarbeiten extrahieren?

Studien zur Ausbildung diagnostischer Kompetenzen im Sportlehramt konzentrieren sich vermehrt darauf, wie Sportlehrkräfte bestärkt werden können, formative Beurteilungsformen im Sportunterricht umzusetzen. Diese Studien können demnach im Bereich der Kompetenzfacette professionelle Kompetenz eingeordnet werden, da sie einerseits auf die Vermittlung diagnostischen Methodenwissens und andererseits auf motivational-volitionale Aspekte abzielen.

Mit Blick auf die Förderung von Sportlehrkräften, formative Beurteilungsformen in der Unterrichtspraxis umzusetzen, zeigen die Ergebnisse systematischer Literaturreviews (z. B. Herrero-González et al, 2024; Moura et al., 2021), dass sowohl formative Beurteilungsformen als Teil der Sportlehramtsausbildung zu integrieren, als auch Fortbildungen mit Unterstützungsmaßnahmen zur Umsetzung während der Berufspraxis hilfreich sind. Slingerland et al. (2024) haben die Erfahrungen von Sportlehrkräften qualitativ untersucht, die an einem 18-monatigen Fortbildungs-

programm zur formativen Beurteilung teilnahmen. Die Lehrkräfte entwarfen und implementierten verschiedene formative Beurteilungsformate in ihrem Sportunterricht. Die Ergebnisse zu den Einschätzungen der Lehrkräfte zeigen, dass die Teilnahme zu positiven Erfahrungen mit formativen Beurteilungssituationen führten, wie z. B. eine verstärkte Beteiligung der Schüler:innen. Allerdings erlebten die Sportlehrpersonen auch komplexe Probleme, wie z. B. die sich verändernde Lehrkraftrolle. Ähnliche Ergebnisse zeigen sich auch bei Chróinin und Cosgrave (2013), die die Implementierung formativen Assessments von Grundschulsportlehrpersonen qualitativ untersuchten. Mit Blick auf die Förderung der Umsetzung formativer Beurteilungsformen im Rahmen des Sportlehramtsstudiums (u. a. während Praktika im Praxissemester) haben Backman et al. (2024) den Einsatz von Peer-Assessment im Schulpraktikum aus Sicht von Sportlehramtsstudierenden qualitativ untersucht. Die Ergebnisse zeigen, dass die Sportlehramtsstudierenden den Einsatz von Peer-Assessment u. a. für den Aufbau sozialer Beziehungen bedeutsam finden und um das Lernziel transparent zu machen. Allerdings nehmen sie beim Peer-Assessment auch komplexe Herausforderungen wahr, z. B. wie mit sensiblen und geschlechtsspezifischen Kommentaren, i. S. von wertenden Kommentaren von Schüler:innen untereinander zu Körpern und Bewegungen umgegangen werden soll. Mit Blick auf die Selbstwahrnehmung von Sportlehramtsstudierenden zu ihren pädagogischen Diagnosen untersuchten Atienza et al. (2023) mit einer qualitativen Analyse von Tagebucheinträgen die Auswirkungen des Einsatzes von Lernprozessdiagnosen im Sportlehramtsstudium. Die Ergebnisse deuten darauf hin, dass die Sportlehramtsstudierenden zu Beginn des Studiums keine Erfahrung mit Lernprozessdiagnosen hatten und die Durchführung der Lernprozessdiagnosen im Studium als positiv wahrnehmen. Als Gründe werden genannt, dass sie gelernt haben, wie ein kontinuierliches, sinnvolles, reflektiertes und angepasstes Lernen bewirkt werden kann. Allerdings nehmen sie durch die Lernprozessdiagnosen eine höhere Arbeitsbelastung wahr, sowohl für die Hochschullehrenden als auch für sich selbst. Die Mehrheit der Teilnehmenden äußerte die Absicht, Lernprozessdiagnosen in ihrer zukünftigen beruflichen Praxis anzuwenden. Darüber hinaus gibt es bisher nur vereinzelte Langzeitstudien, die den Übergang von an der Universität gelehrtene Inhaltsbereiche zu der beruflichen Praxis als Sportlehrperson in der Schule untersuchen (Tolgfors et al., 2023). Erste Befunde zeigen, dass die im Sportlehramtsstudium erworbenen theoretischen Erkenntnisse und im Praktikum gemachten Erfahrungen zum formativen Assessment die Lehrpersonen bestärken können, dies auch in der späteren Berufspraxis umzusetzen. In der Berufspraxis werden in der Umsetzung Herausforderungen erlebt, wie Widerstände seitens der Schüler:innen und andere Beurteilungsansätze von Kollegen (Tolgfors et al., 2023).

Hendricks (2025) konnte zeigen, dass die Leistung von Sportlehramtsstudierenden bei der Beurteilung kognitiver und motorischer Anforderungen in Aufgaben nicht mit der Studiendauer oder der Anzahl absolvierter Praktika zusammenhängt. Er schlussfolgert daraus, dass der Besuch von universitären Lehrveranstaltungen oder

Schulpraktika nicht per se zu einer Förderung der Aufgabenanalysefähigkeit beiträgt, sondern dass diese spezielle Ausbildungsmomente erfordert. Insgesamt existieren bislang kaum (Interventions-)studien, die Verbesserungsmöglichkeiten der Kompetenzfacette W-I-E oder der Urteilsakkuratheit (Performanz) von (angehenden) Sportlehrkräften untersuchen.

5 Diskussion und Identifikation zentraler Forschungslücken

Dieser Beitrag verfolgt das Ziel, diagnostische Kompetenzen aus einer sportpädagogischen Perspektive auf theoretischer Ebene zu konzeptualisieren, um darauf aufbauend bestehende Forschungsarbeiten zu systematisieren und Desiderate hinsichtlich der Aus- und Fortbildung von (angehenden) Sportlehrkräften aufzuzeigen. Es kann herausgestellt werden, dass diagnostische Kompetenzen als Teilbereich professioneller Handlungskompetenzen von Sportlehrkräften eine zentrale Rolle für die Gestaltung eines lernwirksamen Sportunterrichts spielen. Diese Relevanz zeichnet sich in den inzwischen zahlreichen Forschungsbeiträgen zu diagnostischen Kompetenzen von Sportlehrkräften ab. Die Beiträge unterscheiden sich in Bezug auf ihre theoretische Grundlage und dem eingenommenen Fokus teils deutlich voneinander. Auf Grundlage des angepassten KompeSport-Modells (Abb.1; Baumgartner, 2022) können die Beiträge systematisiert werden. Im Folgenden werden die Ergebnisse vor dem Hintergrund des angelegten theoretischen Hintergrunds eingeordnet, um darauf aufbauend entlang des KompeSport-Modells offene Forschungsfelder abzuleiten.

Durch die Orientierung am KompeSport-Modell (Baumgartner, 2022) können eindeutige Kompetenzfacetten benannt werden, mit der diagnostische Kompetenzen operationalisiert werden können. Dadurch ist es möglich, die bestehenden Studien den Kompetenzfacetten zuzuordnen. So können Ansatzpunkte für die Sportlehrkräftebildungsforschung sowie für die Aus- und Fortbildung aufgezeigt werden. Durch die eingenommene kompetenztheoretische Sichtweise besteht jedoch ein eingeschränkter Blick auf die Professionalisierung von Sportlehrkräften, da andere theoretische Herangehensweisen nicht berücksichtigt werden (z. B. der professionelle Blick bei Reuker & Schicklinski, 2024). Durch eine andere theoretische Sichtweise ließen sich ggf. unterschiedliche Forschungslücken und Möglichkeiten der Aus- und Fortbildung aufzeigen.

Insgesamt deuten sich die einleitend herausgestellten Problemfelder auch im sportpädagogischen/-didaktischen Diskurs an. Dies zeigt sich weniger in der begrifflichen Vielfalt oder der Auslegung der professionellen Handlungskompetenzen, sondern eher im Bereich der Modellierung diagnostischer Kompetenzen. Die sportpädagogischen/-didaktischen Studien greifen meist auf fachübergreifende Modelle zurück (u. a. Blömeke et al., 2015; Brühwiler, 2014; Herppich et al., 2017; Südkamp et al., 2017) und deuten sie sportspezifisch (u. a. Hendricks, 2024; Langer, 2023). Damit

folgen sie einheitlich der Auffassung, dass diagnostische Kompetenzen fachspezifische Kompetenzen von Sportlehrkräften darstellen.

In der Analyse zur Beantwortung der ersten Frage, welche Erkenntnisse zur Bedeutung der Kompetenzfacetten vorliegen, können drei zentrale Forschungsstränge identifiziert werden. Hierbei wird deutlich, dass ein Schwerpunkt der Studien auf der Untersuchung der Urteilsakkuratheit als Ergebnis der Performanz liegt (erster Forschungsstrang). Aus den bisherigen Befunden zur geringen Akkuratheit von Lehrerurteilen (sowohl zur Rang-, Niveau- und Differenzierungskomponente) im Fach Sport lässt sich der Bedarf ableiten, Lehrpersonen aller Schulstufen sowohl in der Beurteilung motorischer Lernvoraussetzungen (Niederkofler et al., 2018; Niederkofler, 2023) und Fähigkeitsselbstkonzepte ihrer Schüler:innen (Seyda, 2018) als auch der Einschätzung von Aufgabenschwierigkeiten (Hendricks, 2024) zu fördern. Hinsichtlich des zweiten Forschungsstrangs stellt sich die Frage, inwiefern neben den bisher untersuchten Berufsmerkmalen von Sportlehrkräften weitere Faktoren der professionellen Kompetenz (z. B. die Ausprägung des Professionswissens) oder Personenmerkmale (z. B. Einstellung) mit der Urteilsakkuratheit zusammenhängen und wie diese in der Sportlehrkräfteprofessionalisierung gefördert werden können. Beides bleibt bislang weitestgehend unberücksichtigt.

Die Analyse zum dritten Forschungsstrang zeigt zudem, dass in der Sportpädagogik/-didaktik neben Studien zur *Urteilsakkuratheit* ebenso *prozessorientierte Ansätze* zahlreich verfolgt werden. Mit Blick auf prozessorientierte Ansätze scheint die Sportdidaktik der Forderung nach einer stärker am Prozess orientierten Forschung von Kompetenzen (Baumert & Kunter, 2006, Blömeke et al., 2015) im Gegensatz zu anderen Fächern eher gerecht zu werden. Dies könnte daran liegen, dass der Diskurs um diagnostische Kompetenzen im Sport im Vergleich zu anderen Fächern (bspw. der Mathematik) recht spät an Bedeutung gewonnen hat. Dadurch konnte auf zahlreiche Vorarbeiten anderer Fächer zurückgegriffen werden, in denen die Bedeutung prozessorientierter Studien betont wird (z. B. Klug et al., 2013). Die prozessorientierten Studien greifen im Sport meist auf eine retrospektive Erläuterung der eigenen diagnostischen Tätigkeiten durch die Sportlehrkräfte (u. a. Endacott, 2016; Langer, 2023; Leukel et al., 2023) oder auf eine Einschätzung zur persönlichen Relevanz des Diagnostizierens zurück (u. a. Atienza et al., 2023). Die Informationsverarbeitungsprozesse während des Diagnostizierens wurden bislang noch nicht berücksichtigt. Es mangelnd demnach an Studien, die untersuchen, welche Informationen in einer Diagnosesituation von Lehrpersonen wahrgenommen werden, wie diese im Hinblick auf das daraus resultierende Urteil interpretiert werden und welche Entscheidungen daraus abgeleitet werden. Eine stärkere Orientierung am Diagnoseprozess ist notwendig, um ihn besser zu verstehen, in der Aus- und Fortbildung zu adressieren und in der Unterrichtspraxis unterstützen zu können. Dies ist insbesondere dann wichtig, wenn der Zusammenhang zwischen diagnostischen und didaktischen Kompetenzen untersucht werden soll (u. a. Seyda & Langer, 2020).

Ergänzend gibt es erste Studien, die vor dem Hintergrund der Unterrichtsadaptivität das diagnostische Urteil zusammen mit der didaktischen Entscheidung betrachten (Langer, 2023).

Hingegen ist gut erforscht, welche Beurteilungsformen von Sportlehrkräften verwendet werden. Die Ergebnisse zeigen, dass Sportlehrkräfte vorrangig summative Beurteilungen verwenden (Borghouts et al., 2017; Moura et al., 2021). Zudem basieren ihre Diagnosen auf subjektiven Einschätzungen, während evidenz-basierte Beurteilungen selten genutzt werden (Snow et al., 2025). Daraus lässt sich ableiten, dass Sportlehrkräfte darin bestärkt werden sollten, den Lernprozess ihrer Schüler:innen zu diagnostizieren und damit auch das selbstständige Lernen ihrer Schüler:innen zu fördern. Dies wird vor allem in internationalen Studien zur Förderung diagnostischer Kompetenzen bereits aufgegriffen. Es fehlen in diesem Zusammenhang jedoch Studien, die untersuchen, inwiefern der Lernprozess auch akkurat beurteilt wird.

Aus der Analyse zur Beantwortung der zweiten Frage, welche Erkenntnisse über Maßnahmen zur Förderung der einzelnen Kompetenzfacetten (angehender) Sportlehrkräfte bei der Beurteilung von Schüler:innenmerkmalen vorliegen, lassen sich unterschiedliche Fördermaßnahmen herausstellen. Es zeigt sich, dass die Teilnahme an kompetenzbereichsbezogenen Fort- und Ausbildungen zu formativen Beurteilungen (angehende) Sportlehrkräfte darin bestärken kann, derartige Beurteilungsformen im Sportunterricht umzusetzen. Allerdings fehlen bislang noch empirische Befunde, ob sich diagnostische Kompetenzen (angehender) Sportlehrkräfte durch die Teilnahme tatsächlich verbessern (i. S. einer gesteigerten Urteilsakkuratheit in der Performanz). Bislang existieren empirische Befunde, die die gesteigerte Selbstwahrnehmungen (angehender) Sportlehrkräfte zeigen. Erste Pilotstudien setzten an diesem Desiderat an. Leukel et al. (2025) untersuchten, ob durch eine Fortbildung, Sportlehrkräfte unterstützt werden können, Bewegungsmerkmale beim Schlagwurf akkurat zu diagnostizieren. Als ein weiteres Desiderat lässt sich aus den bisherigen Befunden ableiten, Sportlehrkräfte durch Fort- und Ausbildungen darin zu bestärken, evidenzbasierte, formelle Beurteilungen im Sportunterricht zu nutzen (wie z. B. den MOBAK, Herrmann et al., 2025). Darüber hinaus wird die Aufgabenanalysefähigkeit als Teilbereich diagnostischer Kompetenzen in der Aus- und Fortbildung bisher wenig berücksichtigt. »Dem kompetenztheoretischen Ansatz der Lehrkraftprofessionalisierung folgend ist dies jedoch eine wichtige Voraussetzung, damit Lehrkräfte Kompetenzen aufbauen« (Hendricks, 2024, S. 344). Eine weitere offene Frage ist, wie Diagnostizieren gelehrt und gelernt wird. Es ist bislang weitestgehend unbekannt, wie effektive Lehr-Lernsettings zur Vermittlung diagnostischer Kompetenzen in den unterschiedlichen Bewegungsfeldern konzipiert sein sollten, damit (angehende) Sportlehrkräfte auf die Beurteilungsaufgaben in ihrem späteren Sportunterricht vorbereitet werden.

Neben den identifizierten Forschungslücken innerhalb der drei Forschungssträngen und den Erkenntnissen zur Aus- und Fortbildung lassen sich aus dem Modell auf inhaltlicher Ebene weitere Forschungsstränge ableiten, die im sportpädagogischen/-didaktischen Diskurs bislang noch nicht berücksichtigt werden. Bislang liegen noch keine Studien vor, die den Einfluss der Diagnosesituation auf die Beurteilung untersucht haben. Dem KompeSport-Modell folgend wäre denkbar, dass sich der Diagnoseprozess und die Performanz in Abhängigkeit von Situationen mit oder ohne Zeitdruck unterscheiden. Erste Überlegungen hierzu werden von Hendricks (2024) in Grundzügen zwar mitgedacht, aber nicht tiefergehend empirisch untersucht. Des Weiteren stellt sich die Frage, inwiefern sich Merkmale der Lerngruppe (z. B. der Lerngruppengröße; Zusammensetzung hinsichtlich Heterogenität oder BMI der Schüler:innen) auf den Diagnoseprozess und die Performanz auswirken.

Darüber hinaus ist bislang nicht bekannt, wie die Ausprägung diagnostischer Kompetenzen von Sportlehrkräften mit der Kompetenzentwicklung der Schüler:innen zusammenhängt. Es stellen sich bspw. die Fragen, ob sich besonders stark ausgeprägte diagnostische Kompetenzen positiv auf die Kompetenzentwicklung der Schüler:innen auswirken oder inwiefern es für einen Lernerfolg der Schüler:innen eine Verknüpfung hoher diagnostischer Kompetenzen mit Strukturierungsmaßnahmen bedarf (z. B. wie in der Mathematik; Schrader & Helmke, 1987).

Darüber hinaus ergeben sich ebenso offene Forschungsstränge auf methodischer Ebene. Mit Blick auf die Kompetenzfacette professioneller Kompetenz fehlen bislang Wissenstests z. B. über diagnostisches Methodenwissen (angehender) Sportlehrkräfte und inwiefern es die Akkuratheit von Diagnosen beeinflusst. Für die Kompetenzfacette W-I-E bedarf es u. a. der Untersuchung mediiender Prozesse z. B. über Unterrichtsvignetten (Bartel & Roth, 2015). Mit Blick auf die Performanz sind zukünftig Beobachtungsinstrumente für den Sportunterricht oder Mixed-Method-Ansätze sinnvoll, mit denen die Qualität der Diagnosen gemessen werden kann (Baumgartner, 2017; Baumgartner, 2022). Darüber hinaus scheinen Mixed-Method-Ansätze oder Mediationsanalysen sinnvoll, mit denen das Zusammenspiel der drei Kompetenzfacetten untersucht werden kann, um herauszufinden, was diagnostische Kompetenzen kompetenter Sportlehrkräfte auszeichnet.

Der vorliegende Beitrag verfolgt das Ziel, den Forschungsstand mit der spezifischen Sichtweise auf eine kompetenzorientierte Sportlehrkräftebildung umfassend aufzuzeigen. Es wurde demnach keine empirische Aufarbeitung des Forschungsstands in Form eines systematischen Reviews durchgeführt, was als Einschränkung des Vorgehens gesehen werden muss. Zusammenfassend kann festgehalten werden, dass die Vermittlung diagnostischer Kompetenzen mit Blick auf den Forschungsstand eine der zentralen Ausbildungsaufgaben in der universitären Sportlehrkräfteprofessionalisierung zu sein scheint. Damit können angehende Sportlehrkräfte befähigt werden, einen lernförderlichen Sportunterricht zu gestalten. Durch die Zuordnung

der einzelnen Beiträge in das KompeSport-Modell (Baumgartner, 2022) konnten die folgenden zentralen Forschungslücken identifiziert werden:

- a. Inwiefern beeinflusst die Diagnosesituation (mit und ohne Zeitdruck; Schüler:innenmerkmale wie der BMI) den Beurteilungsprozess und die Performanz?
- b. Inwiefern hängt bei diagnostischen Kompetenzen die Ausprägung der Kompetenzfacette professionelle Kompetenz mit den Wahrnehmungs-, Interpretations- und Entscheidungsprozessen sowie mit der Performanz zusammen?
- c. Welche Informationsverarbeitungsprozesse laufen während des Diagnostizierens ab?
- d. Inwiefern hängt eine hohe Ausprägung diagnostischer Kompetenzen (im Sinne einer Katalysatorvariable) mit einem Kompetenzaufbau aufseiten der Schüler:innen zusammen?
- e. Wie kann die Entwicklung diagnostischer Kompetenzen von (angehenden) Sportlehrkräften entlang der Kompetenzfacetten in der Aus- und Fortbildung erfolgreich adressiert werden?

6 Einordnung der Beiträge im Themenheft

Die Beiträge im Themenheft orientieren sich an den in diesem Hauptbeitrag abgeleiteten zentralen Forschungslücken. Storni et al. gehen im Beitrag *The Diagnostic Competencies of Physical Education Teachers in Assessing Basic Motor Competencies in Primary School* u. a. auf die Frage ein, inwiefern Schüler:innenmerkmale wie der BMI die Urteilsakkuratheit beeinflussen (Forschungslücke a). Ziel des Beitrags ist es, die Genauigkeit dieser Einschätzung in einer quantitativen Querschnittsstudie zu untersuchen und Faktoren zu identifizieren, die mit der Urteilsakkuratheit in Zusammenhang stehen.

Hendricks geht im Beitrag *Aufgabenanalytische Kompetenzen angehender Sportlehrkräfte als eine Voraussetzung zur kognitiven Aktivierung von Schüler:innen im Sportunterricht* auf die Frage ein, welche Informationsverarbeitungsprozesse während des Diagnostizierens von kognitiven Anforderungen in Aufgaben ablaufen (Forschungslücke c). In Form einer Mixed-Methods Studie wird untersucht, inwiefern angehende Sportlehrkräfte die kognitiven Anforderungen in Aufgaben akkurat einschätzen können und ob sich Unterschiede in den informationsverarbeitenden Prozessen während der Aufgabenanalyse bei den angehenden Sportlehrkräften ergeben. Das Ziel der Studie liegt darin, Informationen über die der Urteilsbildung zugrundeliegenden Wahrnehmungs-, Interpretations-, und Entscheidungsprozesse zu erhalten.

Fokken et al. gehen im dritten Beitrag *Digital Transformation of a Diagnostic Procedure for Swimming Skill Analysis: Teachers' Evaluation of a Low-Fidelity App Prototype (Click Dummy) of the 'Owl's Eye'* auf die Frage ein, wie Lehrkräfte die Nutzbarkeit des App-Prototypen der App »Owl's Eye«, die zur Simulation digitaler

Unterstützung bei der Analyse der Schwimmfähigkeiten von Lernenden verwendet wird, beurteilen und welchen Mehrwert sie in der Verwendung der App für ihre diagnostische Arbeit sehen (Forschungslücke e). Dazu wurde der Umgang mit einem Click Dummy mithilfe von »Think-Aloud«-Protokollen erhoben und anschließend anhand einer qualitativen Inhaltsanalyse ausgewertet.

Langer und Schlapkohl gehen im vierten Beitrag *Diagnostische Kompetenzen von Sportlehramtsstudierenden durch authentische Situationen professionell entwickeln* auf die Frage ein, wie Sportlehramtsstudierende im Bewegungsfeld Bewegen im Wasser die Umsetzung authentischer Diagnosesituationen (Peer Assessment als Teil des formativen Assessment und Schüler:innen diagnostizieren mit einem summativen Assessment) erleben (Forschungslücke e). Dafür werden in einer qualitativen Querschnittsstudie die erlebten Potentiale und Herausforderungen der Sportlehramtsstudierenden nach der Teilnahme im Seminar »Bewegen im Wasser« analysiert und diskutiert. Das Ziel der Studie liegt darin, Informationen über die Entwicklungsmöglichkeiten der Wahrnehmungs-, Interpretations-, und Entscheidungsprozesse sowie der Performanz zu erhalten.

Die Beiträge lassen sich hinsichtlich des KompeSport-Modells sowohl der Kompetenzfacette Performanz (Storni et al.), in der Schnittstelle zwischen der Kompetenzfacette W-I-E und der Performanz (Hendricks), der Kompetenzfacette W-I-E (Fokken et al.) sowie in der Schnittstelle zwischen professioneller Kompetenz und Kompetenzfacette W-I-E (Langer & Schlapkohl) zuordnen.

Der Zusammenhang zwischen der Ausprägung der drei Kompetenzfacetten (Forschungslücke b) sowie der Zusammenhang zwischen den Kompetenzfacetten der Lehrkräfte und der Ausprägung der Kompetenzfacetten der Schüler:innen (Forschungslücke d) bleibt weiterhin ein Desiderat, was mit diesem Themenheft nicht behandelt wird. Das Themenheft adressiert somit offene Fragen diagnostischer Kompetenzen von Sportlehrkräften. Gleichzeitig zeigt es zukünftigen empirischen Forschungsbedarf auf und liefert Denkanstöße für die Weiterentwicklung eines zentralen Aspektes der professionellen Handlungskompetenzen von (angehenden) Sportlehrkräften.

Literatur

- AIESEP (2020). *AIESEP Position Statement on Physical Education Assessment*. Abruf unter <https://aiesep.org/scientific-meetings/position-statements/>
- 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.
- Atienza, R., Valencia-Peris, A., & López-Pastor, V. M. (2023). Formative assessment and pre-service teacher education: previous, current and prospective experiences. *Cultura, Ciencia y Deporte*, 18(55), 133-156. <https://doi.org/10.12800/ccd.v18i55.1914>
- Aufschnaiter, C. v., Cappell, J., Dübbele, G., Ennemoser, M., Mayer, J., Stiensmeier-Pelster, J., Sträßer, R., & Wolgast, A. (2015). Diagnostische Kompetenz: Theoretische Über-

- legungen zu einem zentralen Konstrukt der Lehrerbildung. *Zeitschrift für Pädagogik*, 61(5), 738-757.
- Backman, E., Quennerstedt, M., Tolgfors, B., & Nyberg, G. (2024). Peer assessment in physical education teacher education – a complex process making social and physical capital visible. *Curriculum Studies in Health and Physical Education*, 15(3), 274–288. <https://doi.org/10.1080/25742981.2023.2256327>
- Backman, E., Tolgfors, B., Nyberg, G., & Quennerstedt, M. (2023). How does physical education teacher education matter? A methodological approach to understanding transitions from PETE to school physical education. *Physical Education & Sport Pedagogy*, 28(4), 411–424. <https://doi.org/10.1080/17408989.2021.1990248>
- Barrientos, E., Lopez-Pastor, E., Lorente-Catalán, E., & Kirk, D. (2023). Challenges with using formative and authentic assessment in physical education teaching from experienced teachers' perspectives. *Curriculum Studies in Health and Physical Education*, 14(2). <https://doi.org/10.1080/25742981.2022.2060118>
- Bartel, M-E., & Roth, J. (2015). Diagnostische Kompetenz durch Videovignetten fördern. In F. Caluori, H. Linneweber-Lammerskitten & C. Streit (Hrsg.), *Beiträge zum Mathematikunterricht*. WTM-Verlag
- Baumert, J., & Kunter, M. (2006). Stichwort: Professionelle Kompetenz von Lehrkräften. *Zeitschrift für Erziehungswissenschaft*, 9(4), 469-520.
- Baumgartner, M. (2017). Performanzentwicklung in der Ausbildung von Lehrkräften: Eine Interventionsstudie zur Verbesserung des Feedbacks bei angehenden Sport-lehrkräften. Dissertation. *Empirische Erziehungswissenschaft: Band 65*.
- Baumgartner, M. (2022). Professional competence(s) of physical education teachers: terms, traditions, modelling and perspectives. *German Journal of Exercise and Sport Research*, 52, 550-557.
- Baumgartner, M., Owassapian, D., & Perret Guldemann, J. (2021). Zur Verbesserung der Klassenführung von sportunterrichtenden Lehrkräften auf der E-Learning Plattform bewegunglesen.ch. *Sportunterricht*, 70(3), 111-116.
- Betsch, T., Funke, J., & Plessner, H. (2011). *Denken – Urteilen Entscheiden Problemlösen*. Springer.
- Bezeau, D., Chevrier, J. R., & Savard, M. (2023). Opinions on grading and difficulties encountered by physical education and health teachers in relation to assessment. *Physical Education & Sport Pedagogy*, 1–14. <https://doi.org/10.1080/17408989.2023.2269959>
- Blömeke, S., Gustafsson, J. E., & Shavelson, R. J. (2015). Beyond Dichotomies Competence Viewed as a Continuum. *Zeitschrift für Psychologie*, 223(1), 3–13.
- Borghouts, L., Slingerland, M., & Haerens, L. (2017). Assessment quality and practices in secondary PE in the Netherlands. *Physical Education and Sport Pedagogy*, 22(5), 473–489. <http://dx.doi.org/10.1080/17408989.2016.1241226>
- Bräutigam, M. (2015). *Sportdidaktik. Ein Lehrbuch in 12 Lektionen (Sportwissenschaft studieren, 3)*. Meyer & Meyer.
- Brühwiler, Ch. (2014). *Adaptive Lehrkompetenz und schulisches Lernen. Effekte handlungssteuernder Kognitionen von Lehrpersonen auf Unterrichtsprozesse und Lernergebnisse der Schülerinnen und Schüler*. Waxmann.
- Brunner, M., Anders, Y., Hachfeld, A., & Krauss, S. (2011). Diagnostische Fähigkeiten von Mathematiklehrkräften. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss, & M. Neubrand (Hrsg.), *Professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV* (S. 215-234). Waxmann.
- Chróinin, D. N., & Cosgrave, C. (2013). Implementing formative assessment in primary physical education. Teacher Perspectives and Experiences. *Physical Education and Sport Pedagogy*, 18(2), 219–233.
- Dinan-Thompson, M., & Penney, D. (2015). Assessment literacy in primary physical education. *European Physical Education Review*, 21(4), 485–503. <https://doi.org/10.1177/1356336X15584087>
- Deutsche Vereinigung für Sportwissenschaft (dvs). (2011). *Beschlüsse des dvs-Präsidiums vom 10.06.2011* (zur Gewichtung von Publikationen). Zugriff am 19.01.2026 unter: https://www.sportwissenschaft.de/fileadmin/pdf/download/CHE-Beschluesse_140611.pdf

- Endacott, J. L. (2016). Using video-stimulated recall to enhance preservice-teacher reflection. *The New Educator*, 12(1), 28–47. <https://doi.org/10.1080/1547688X.2015.1113351>
- Ferrari, I., Kühnis, J., Bretz, K., & Herrmann, Ch. (2022). Diagnostische Kompetenz der Lehrpersonen und deren didaktischen Implikationen in der Förderung motorischer Basiskompetenzen. In R. Messmer & C. Krieger (Hrsg.), *Narrative zwischen Wissen und Können Aktuelle Befunde aus Sportdidaktik und Sportpädagogik* (S. 177-196). Academia.
- Hascher, T. (2008). Diagnostische Kompetenzen im Lehrberuf. In C. Kraler & M. Schratz (Hrsg.), *Wissen erwerben, Kompetenzen entwickeln. Modelle zur kompetenzorientierten Lehrerbildung* (S. 71-86). Waxmann.
- Hay, P., & Penney, D. (2012). Assessment in physical education. Routledge. <https://doi.org/10.4324/9780203133163>
- Heemsoth, T. (2016). Fachspezifisches Wissen von Sportlehrkräften. Ein Überblick über fachübergreifende und fachfremde Ansätze und Perspektiven für die Professionsforschung von Sportlehrkräften. *Zeitschrift für sportpädagogische Forschung*, 4(2), 41-60.
- Helmke, A. (2017). *Unterrichtsqualität und Lehrerprofessionalität*. Klett Friedrich Verlag.
- Hendricks, C.P. (2024). *Aufgabenanalysefähigkeit als Teil der professionellen Handlungskompetenz von Sportlehrkräften – Theoretische Grundlegung und empirische Zugänge*. Springer VS.
- Hendricks, P. (2025). Aufgabenanalysefähigkeit von angehenden Sportlehrkräften – Eine quantitative Studie zu einem bisher vernachlässigten Thema in der Sportlehrkräftebildungsforschung. *Zeitschrift für sportpädagogische Forschung*, 13(2), 113-134.
- Herppich, S., Praetorius, A.-K., Hetmanek, A., Glogger-Frey, I., Ufer, S., Leutner, D., Behrmann, L., Böhmer, I., Böhmer, M., Förster, N., Kaiser, J., Karin, C., Karst, K., Klug, J., Ohle, A., & Südkamp, A. (2017). Ein Arbeitsmodell für die empirische Erforschung der diagnostischen Kompetenz von Lehrkräften. In A. Südkamp & A.-K. Praetorius (Hrsg.), *Diagnostische Kompetenz von Lehrkräften. Theoretische und methodische Weiterentwicklungen* (S. 75-94). Waxmann.
- Herrero-González, D., López-Pastor, V. M., Manrique-Arribas, J. C., & Moura, A. (2024). Formative and shared assessment: Literature review on the main contributions in physical education and physical education teacher education. *European Physical Education Review*, 30(3), 493-510. [10.1177/1356336X231220995](https://doi.org/10.1177/1356336X231220995)
- Herrmann, C., & Seelig, H. (2020). Motorische Basiskompetenzen. In P. Neumann, & E. Balz (Hrsg.), *Grundschulsport: empirische Einblicke und pädagogische Empfehlungen* (S. 17-30). Meyer & Meyer.
- Herrmann, C., Ferrari, I., & Bretz, K. (2025). Assessing Basic Motor Competencies in 1st- and 2nd-Grade Primary School Children - The MOBAC-1-2 test. <https://doi.org/10.5281/zenodo.17140251>
- Hesse, I., & Latzko, B. (2011). *Diagnostik für Lehrkräfte*. Beltz.
- Hoth, J., Döhrmann, M., Kaiser, G., Busse, A., König, J., & Blömeke, S. (2016). Diagnostic competence of primary school mathematics teachers during classroom situations. *ZDM Mathematics Education*, 49, 41-53.
- Ingenkamp, K., & Lissmann, U. (2008). *Lehrbuch der pädagogischen Diagnostik*. Beltz.
- Karst, K. (2009). *Kompetenzmodellierung des diagnostischen Urteils von Grundschullehrern*. Dissertation, Universität Kassel.
- Karst, K. (2012). *Kompetenzmodellierung des diagnostischen Urteils von Grundschullehrern*. Empirische Erziehungswissenschaft, Band 35: Bd. 35. Waxmann.
- Klug, J., Bruder, S., Kelava, A., Spiel, C., & Schmitz, B. (2013). Diagnostic Competence of Teachers: A Process Model That Accounts for Diagnosing Learning Behavior Tested by Means of a Case Scenario. *Teaching and Teacher Education*, 30, 38-46.
- Klug, J., Bruder, S., Keller, S., & Schmitz, B. (2012). Hängen Diagnostische Kompetenz und Beratungskompetenz von Lehrkräften zusammen? In J. Fleischer & D. Leutner (Hrsg.), Themenheft: Modellierung von Kompetenzen im Bereich der Bildung. Die psychologische Perspektive. *Psychologische Rundschau*, 63, 3-10.
- KMK. (2022). *Standards Lehrerbildung: Bildungswissenschaften*. KMK.
- Korban, S., & Künzell, S. (2019). Verbesserung der diagnostischen Kompetenz mit iPads – Ein Ausbildungskonzept, *Zeitschrift für Studium und Lehre in der Sportwissenschaft*, 2(2), 5–13. <https://doi.org/10.25847/zsls.2018.012>

- Krauss, S., Kunter, M., Brunner, M., Baumert, J., Blum, W., Neubrand, M., Jordan, A., & Löwen, K. (2004). COACTIV: Professionswissen von Lehrkräften, kognitiv aktivierender Mathematikunterricht und die Entwicklung von mathematischer Kompetenz. In J. Doll & M. Prenzel (Hrsg.), *Bildungsqualität von Schule. Lehrerprofessionalisierung, Unterrichtsentwicklung und Schülerförderung als Strategien der Qualitätsverbesserung* (S. 31-53). Waxmann.
- Langer, A. (2023). *Pädagogische Diagnosen im Kontext der Unterrichtsgestaltung im Sport*. Dissertation, Hofmann Verlag.
- Langfeld, H. P. (2014). Diagnosekompetenzen von Lehrerinnen und Lehrern. In H. P. Langfeld (Hrsg.), *Psychologie für die Schule* (S. 195-210). Beltz.
- Latzko, B., & Gottlebe, K. (2022). Der Nutzen der pädagogisch-psychologischen Diagnostik für eine diversitätssensitive und inklusive, pädagogische Praxis. In K. Kramer & B. Hoyer (Hrsg.), *FAU Lehren und Lernen: Bd. 8. "Individuell fördern"*. *Wissenschaftlicher Hintergrund sowie Ansatzpunkte aus und für die Praxis* (S. 43-67). FAU University Press.
- Leirhaug, P. E., & Annerstedt, C. (2016). Assessing with new eyes? Assessment for learning in Norwegian physical education. *Physical Education & Sport Pedagogy*, 21(6), 616–631. <https://doi.org/10.1080/17408989.2015.1095871>
- Leukel, C., Leuders, T., Bessi, F., & Loibl, K. (2023). Unveiling Cognitive Aspects and Accuracy of Diagnostic Judgments in Physical Education Teachers Assessing Different Tasks in Gymnastics. *Front. Educ.*, 8(1162499), 1-14.
- Leukel, C., Loibl, K., & Leuders, T. (2025). Förderung diagnostischer Kompetenzen von Sportlehrkräften – Entwicklung einer 4C/ID-basierten Fortbildung. *HLZ*, 8(1), 122–135. <https://doi.org/10.11576/hlz-7391>
- 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>.
- McElvany, N., Schroeder, S., Hachfeld, A., Baumert, J., Richter, T., Schnotz, W., Horz, H., & Ullrich, M. (2009). Diagnostische Fähigkeiten von Lehrkräften bei der Einschätzung von Schülerleistungen und Aufgabenschwierigkeiten bei Lernmedien mit instruktionalen Bildern. *Zeitschrift für Pädagogische Psychologie*, 23(34), 223-235.
- Ministerium für Schule und Bildung des Landes Nordrhein-Westfalen (2021). Lehrplan für die Primarstufe in Nordrhein-Westfalen Fach Sport. MSW.
- Moura, A., Graca, A., MacPhail, A. & Batista, P. (2021). Aligning the Principles of Assessment for Learning to Learning in Physical Education: A Review of Literature. *Physical Education and Sport Pedagogy*, 26(4). <https://doi.org/10.1080/17408989.2020.1834528>
- Neuber, N. (2020). *Fachdidaktische Konzepte Sport*. Springer VS.
- Neumaier, A. (2009). *Koordinatives Anforderungsprofil und Koordinationstraining* (3. Überarbeitete Auflage). Sportverlag Strauß.
- Niederkofler, B. (2023). Urteilsakkuratheit von Grundschullehrpersonen in der Einschätzung von motorischen Basiskompetenzen. *Motorik*, 1(23), 26–34. <https://doi.org/10.2378/mo12023.art05d>.
- Niederkofler, B., Herrmann, C., & Amesberger, G. (2018). Diagnosekompetenz von Sportlehrkräften – Semiformelle Diagnose von motorischen Basiskompetenzen. *Zeitschrift für sportpädagogische Forschung*, 6, 72–96.
- Ostermann, A., Leuders, T., & Nückles, M. (2015). Wissen, was Schülerinnen und Schülern schwer fällt. Welche Faktoren beeinflussen die Schwierigkeitseinschätzung von Mathematikaufgaben? *Journal für Mathematik-Didaktik*, 36(1), 45–76.
- Paradies, L., Linser, H. J., & Greving, J. (2007). *Diagnostizieren, Fordern und Fördern*. Cornelsen Scriptor.
- Praetorius, A.-K., & Südkamp, A. (2017). Eine Einführung in das Thema der diagnostischen Kompetenz von Lehrkräften. In A. Südkamp & A.-K. Praetorius (Hrsg.), *Pädagogische Psychologie und Entwicklungspsychologie: Bd. 94. Diagnostische Kompetenz von Lehrkräften: Theoretische und methodische Weiterentwicklungen* (S. 13-18). Waxmann.
- Reuker, S. & Schicklinski, K. (2024). Der Professionelle Blick von Sportlehrkräften – Revisited. *German Journal of Exercise and Sport Research*, 54(4), 531–540. <https://doi.org/10.1007/s12662-023-00935-1>.

- Röger-Offergeld, U., Kirsch, S., & Sygusch, R. (2021). Kompetenzorientierte Sportlehrkräftebildung am Beispiel Gesundheit. *Zeitschrift für Studium und Lehre in der Sportwissenschaft*, 4(1), 29-37.
- Schittkowski, B., Woll, A., & Wagner, I. (2024). Multidimensionale Herausforderungen und Strategien der Leistungsbewertung im Sportunterricht: eine qualitative Analyse. *German Journal of Exercise and Sport Research*. <https://doi.org/10.1007/s12662-024-01004-x>
- Schrader, F.-W. (2013). *Diagnostische Kompetenz von Lehrpersonen*. Beiträge zur Lehrerinnen- und Lehrerbildung, 31(2), 154–165.
- Schrader, F.-W. (2014). Lehrer als Diagnostiker. In E. Terhart, H. Bennewitz, & M. Rothland (Hrsg.), *Handbuch der Forschung zum Lehrerberuf* (S. 865-882). Waxmann.
- Schrader, F.-W. (2017). Diagnostische Kompetenz von Lehrkräften – Anmerkungen zur Weiterentwicklung des Konstrukts. In A. Südkamp & A.-K. Praetorius (Hrsg.), *Diagnostische Kompetenz von Lehrkräften* (S. 245-255). Waxmann.
- Schrader, F.-W., & Helmke, A. (1987). Diagnostische Kompetenz von Lehrern. Komponenten und Wirkung. *Empirische Pädagogik*, 1, 27-52.
- Seyda, M. (2018). Können Sportlehrkräfte die Perspektive ihrer Schülerinnen und Schüler einnehmen? Eine Untersuchung über die Akkuratheit von Beurteilungen physischer Fähigkeitsselbstwahrnehmungen. *Unterrichtswissenschaft*, 46(2), 215-231.
- Seyda, M., & Langer, A. (2020). Die Orchestrierung des diagnostischen und didaktischen Handelns im Sportunterricht. Eine Frage der Kompetenz! *Sportunterricht*, 69(8), 353-357.
- Shulman, L. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review*, 1(57), 1–22.
- Slingerland, M., Weeldenburg, G., & Borghouts, L. (2024). Formative assessment in physical education: teachers' experiences when designing and implementing formative assessment activities. *European Physical Education Review*, 30(4), 620–637. <https://doi.org/10.1177/1356336X241237398>
- Snow, S., Brown, T., Barnett, L., Salmon, J. & Lander, N. (2025). Understanding the Use of Motor Competence Assessments in Australian Elementary Physical Education: A Qualitative Study. *Journal of Teaching in Physical Education*, 1-9. <https://doi.org/10.1123/jtpe.2024-0370>
- 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. 33-37). Waxmann.
- Tolgfors, B., Quennerstedt, M., Backman, E., & Nyberg, G. (2023). A PE teacher's tale: journeying from teacher education to teaching practice in physical education. *Sport, Education and Society*, 30(1), 29–41. <https://doi.org/10.1080/13573322.2023.2281389>
- Van Ophuysen, S. (2010). Professionelle pädagogisch-diagnostische Kompetenz – eine theoretische und empirische Annäherung. In N. Berkemeyer, W. Bos, H. G. Holtappels, N. McElvany, & R. Schulz-Zander (Hrsg.), *Jahrbuch der Schulentwicklung. Daten, Beispiele und Perspektiven* (S. 203-234). Juventa.
- Weinert, F. E. (2000). Lehren und Lernen für die Zukunft – Ansprüche an das Lernen in der Schule. *Pädagogische Nachrichten Rheinland-Pfalz*, 2, 1–16.

Diagnostic Competence and Judgement Accuracy of Primary School PE Teachers in Assessing Basic Motor Competencies

von Simone Storni, Johanna Kress, Kathrin Bretz, Christian Herrmann, Benoît Lenzen, Ilaria Ferrari

Abstract: This study analyses the judgement accuracy of primary school physical education teachers in assessing their pupils' basic motor competencies, comparing their subjective evaluations (LEMOK: Lehrpersoneneinschätzung motorischer Kompetenzen) to the competence levels measured by the MOBAK (Motorische Basiskompetenzen) test, using a sample of 689 pupils (339 pupils in year four of primary school aged: $M = 9.94$ years, $SD = 0.37$, 51.0% girls; and 350 pupils in year four of primary school aged: $M = 10.94$ years, $SD = 0.38$, 50.9% girls). Results show a significant overestimation, especially with regard to Self Movement. Structural Equation Modelling analysis suggests that neither pupils' gender nor BMI directly influence the teachers' judgements, but that they have an indirect effect, mediated by the pupils' effective basic motor competencies. The findings highlight the importance of promoting initial and continuing professional development that integrates objective diagnostic tools and metacognitive awareness to support fairer, more informed and conscious evaluation practices.

Keywords: Basic motor competencies, diagnostic competence, judgement accuracy, physical education teachers.

Diagnostische Kompetenzen und Urteilsgenauigkeit von Sportlehrpersonen bei der Einschätzung motorischer Basiskompetenzen in der Grundschule

Zusammenfassung: Die Studie analysierte die Urteilsgenauigkeit von Sportlehrpersonen an Primarschulen, indem die subjektiven Lehrpersoneneinschätzungen motorischer Kompetenzen (LEMOK) mit den objektiven Ergebnissen der motorischen Basiskompetenzen (MOBAK) von 689 Kindern (339 Viertklässler:innen im Alter $M = 9,94$ Jahre, $SD = 0,37$, 51,0% Mädchen; und 350 Fünftklässler:innen im Alter $M = 10,94$ Jahre, $SD = 0,38$, 50,9% Mädchen) verglichen wurden. Die Daten zeigten eine signifikante Überschätzung der motorischen Basiskompetenzen durch die Lehrpersonen, insbesondere im Bereich des Sich-Bewegens. Mediationsanalysen zeigten, dass weder das Geschlecht noch der BMI die Einschätzungen der Lehrpersonen beeinflussen, sondern indirekt über die tatsächlichen motorischen Basiskompetenzen mediiert werden. Die Ergebnisse betonen die Relevanz der Förderung der beruflichen Aus- und Weiterbildung, wobei die Integration objektiver Diagnoseinstrumente und metakognitiver Reflexionsfähigkeit zur Unterstützung fairerer, fundierterer und bewussterer Bewertungsverfahren von entscheidender Bedeutung ist.

Schlagwörter: Motorische Basiskompetenzen, Diagnostische Kompetenz, Genauigkeit diagnostischer Urteile, Sportlehrpersonen.

Diagnostic Competence and Judgement Accuracy of Primary School PE Teachers in Assessing Basic Motor Competencies

1 Introduction

As early as the 1960s, Swiss pedagogue Moor emphasised that education requires understanding the child before attempting to teach (Moor, 1965). This early pedagogical view anticipated later developments in educational diagnostics, which conceptualise teachers' capacity to assess pupils' cognitive, emotional, and motivational states as a central aspect of professional competence (Shavelson, 1978).

In everyday school life, teachers need a wide range of assessment literacy to correctly assess pupils' characteristics, learning, and task demands (Herppich et al., 2018). In this regard, diagnostic competence has long been recognised as a fundamental component of professional competence (Weinert, 2001). In the German research literature, this construct corresponds to «*Diagnostische Kompetenz*», which is conceptually close to assessment competence but more explicitly grounded in the pedagogical perspective (Hascher, 2008; Ingenkamp & Lissmann, 2008). This view naturally leads to a pedagogical-didactical conception of teachers' professional work, one that emphasises how diagnostic insights inform and adapt instruction. Following this tradition, diagnostic competence serves as a bridge between understanding and promoting learning, enabling teachers to translate diagnostic observations into adaptive teaching actions aligned with pupils' developmental needs.

However, despite growing attention to diagnostic competence, little is known about how it manifests in physical education (PE). Only few studies have examined how accurately PE teachers assess their pupils' basic motor competencies (BMC) (Ferrari et al., 2022; Herrmann & Seelig, 2020; Niederkofler et al., 2024).

In this study, the focus is on BMC, which constitute an essential prerequisite for pupils' participation in sport and physical activity within the school context (Herrmann & Seelig, 2020). Against this background, the study examines how accurately primary school PE teachers assess their pupils' BMC by comparing teachers' anticipatory subjective judgements (LEMOK) with pupils' objectively measured performances (MOBAK). Additionally, it investigates whether pupils' gender and body mass index (BMI) are associated with systematic biases in teachers' judgements. This dual focus allows for an examination of both judgement accuracy and potential sources of diagnostic bias.

1.1 Diagnostic competence and teachers' judgement accuracy

Following Weinert (2001), competencies are conceived as the dynamic interplay of cognitive, motivational, volitional, and social abilities that enable individuals to

act successfully in domain-specific contexts. This definition provides a broad and widely recognised theoretical foundation for understanding teacher competence.

In the field of teacher professional competence, Blömeke et al. (2015) conceptualise in their PID model competence as a continuum linking latent dispositions with situation-specific skills (perception, interpretation and decision-making), which result in observable performance. The PID model describes how teachers transform professional knowledge and underlying attitudes into action by perceiving, interpreting, and deciding in response to specific situations. High-quality PID processes make teaching more adaptive, increasing learning time and supporting pupils' motivation and learning outcomes.

Diagnostic competence is understood as a broader construct of skills and knowledge that includes the collection, interpretation, and use this information in authentic classroom situations (Blömeke et al., 2015; Voss et al., 2011). In this study, diagnostic competence refers to the underlying professional capacity that enables teachers to make accurate, reflective, and pedagogically meaningful judgements, whereas judgement accuracy represents the observable outcome of these diagnostic processes (Blömeke et al., 2015; Südkamp et al., 2012), operationalised as the degree of congruence between teachers' judgements and pupils' objectively measured BMC. This distinction aligns with Artelt and Gräsel's (2009) view that judgement accuracy constitutes a central expression of teachers' professional competence.

In this framework, «diagnostic competence» refers to the overarching professional capacity to make accurate and pedagogically meaningful judgements, while «diagnostic competencies» refer to the specific, observable skills like observing motor performance and adapting instruction. The singular form (competence) designates the integrated construct, whereas the plural (competencies) refers to its observable subcomponents (Barnett et al., 2010; Weinert, 2001). This conceptual distinction is particularly relevant in PE, where diagnostic work involves both underlying dispositions (e.g., perception and interpretation) and concrete, observable actions such as analysing motor performance, identifying errors and adjusting tasks as part of situation-specific professional decision making (Blömeke et al., 2015).

In this study, judgement accuracy serves as an empirical indicator of how effectively teachers apply their diagnostic competence in practice. It is operationalised as the degree of correspondence between teachers' anticipatory LEMOK judgement and pupils' objectively measured MOBAC performance, representing a measurable manifestation of competence in action (Blömeke et al., 2015; Herppich et al., 2018; Herrmann & Seelig, 2020; Südkamp et al., 2012). Through this lens, judgement accuracy represents the performance end of the competence continuum introduced by Blömeke et al. (2015), linking diagnostic dispositions and PID processes to observable evaluative outcomes relevant for effective instructional decision-making.

In PE, this framework has specific relevance, as diagnostic competence is exercised through the observation of embodied motor actions rather than purely cognitive outputs. Consequently, judgement accuracy offers a valid and context-sensitive indicator of diagnostic processes in practice (Südkamp et al., 2012). This perspective underscores the situated, performance-based nature of diagnostic work in PE, where perceptual sensitivity and interpretation are tightly coupled with teachers' ability to make adaptive and evidence-informed decisions that support both learning and instructional alignment.

1.2 Diagnostic judgement as a situated and accuracy-oriented process

Diagnostic judgement is best understood as a situated, accuracy-oriented process that integrates cognitive, motivational, and situational dimensions of teachers' professional action (Blömeke et al., 2015; Brühwiler, 2017; Helmke, 2017). It represents the applied dimension of diagnostic competence, where teachers' professional knowledge and situational skills are transformed into observable evaluative behaviour and thus become accessible through measures of judgement accuracy.

These interrelated components enable teachers to analyse pupils' learning processes, identify difficulties, and design targeted interventions (Brühwiler, 2017; Grotegut & Klingsieck, 2025). Accordingly, diagnostic competence goes beyond mere observation: it involves analytical reasoning and pedagogical interpretation that lead to adaptive and evidence-based decisions (Grotegut & Klingsieck, 2025; McElvany et al., 2009). Observation remains a cornerstone of diagnostic practice. Hascher (2008), Klug et al. (2013) and Kunter et al. (2013) highlight that classroom observation allows teachers to collect rich, situated information essential for informed pedagogical decisions. Moreover, semi-formal learning situations, less structured than standardised assessment situations but remaining pedagogically meaningful, enable teachers to collect complementary data that individualise instructional strategies (Südkamp et al., 2012), supporting more calibrated and accurate judgements.

Recent models (Baumert & Kunter, 2013; Loibl et al., 2020) emphasise that diagnostic judgement is cyclical and self-regulative, requiring teachers to perceive, interpret, and evaluate learning cues while actively managing cognitive biases.

1.3 Diagnostic judgement accuracy in PE

In PE research, diagnostic competence assumes specific characteristics due to the inherently performative, visual, and often subjective nature of classroom observations (Ferrari et al., 2022; Herrmann & Seelig, 2020). These conditions make PE an ideal domain for studying how diagnostic competence manifests through perception and decision-making in complex, dynamic contexts. Within this context, the focus

lies on BMC, which represent an important foundation for children's participation in sport and physical activity within the school setting (Herrmann & Seelig, 2020).

While Niederkofler et al. (2018) conceptualise PE teachers' diagnostic work as involving the observation of motor behaviour, the assessment of BMC, and the adaptation of instruction, Baumgartner (2022) emphasises the reflective and interpretative dimensions of teachers' evaluative practice, highlighting the role of subjectivity and pedagogical meaning in diagnostic processes. Their perspectives are complementary in that both underline the interplay between perception, interpretation, and pedagogical decision-making. In this framework, judgement accuracy represents the critical link between observation and pedagogical adaptation, as only sufficiently accurate judgement can support instructional decisions and learners' motor development (Hulteen et al., 2018; Stodden et al., 2008).

These interconnected processes operationalise diagnostic competence in PE, requiring not only theoretical and practical knowledge but also the ability to make accurate and context-sensitive judgements of pupils' motor performance. Effective diagnostic reasoning also requires teachers to attend to the affective and motivational factors that influence motor behaviour (Hascher, 2008; Spinath, 2005). Moreover, teachers' diagnostic work should remain aligned with the broader educational aims of PE as articulated in official curricula (Deutschschweizer Erziehungsdirektoren-Konferenz [D-EDK], 2017; Divisione della Scuola [DECS], 2015).

Unlike other school subjects assessed through written or standardised tests, PE judgements rely mainly on visual observation, making them more susceptible to bias and perceptual distortion (Herrmann & Seelig, 2020). The accuracy and pedagogical value of such judgements depend largely on teachers' diagnostic competence, specifically their capacity to observe motor behaviour systematically, assess physical abilities, and interpret the psychosocial dynamics influencing engagement (Ferrari et al., 2022; Spinath, 2005). This complexity underscores why judgement accuracy is particularly crucial in PE: it minimises interpretative errors, mitigates bias, and ensures equitable learning opportunities (Südkamp et al., 2012). Accordingly, judgement accuracy serves as a key indicator of diagnostic competence in PE, enabling teachers to identify pupils' motor strengths and difficulties based on a judgement that is sufficiently accurate for educational practice, rather than perfectly precise (Südkamp et al., 2012).

Among the several factors influencing judgement accuracy in PE, gender and body mass index (BMI) warrant particular attention, as both may shape teachers' expectations and perceptions and thus introduce systematic bias (Barnett et al., 2010). Empirical evidence shows that teachers often associate boys with higher competence in Object Movement (OM) tasks (e.g., throwing, catching, bouncing and dribbling) and girls with better performance in Self Movement (SM) tasks (e.g., balancing, rolling, jumping and running), reflecting enduring gender stereotypes that can bias

evaluations prior to observation (Barnett et al., 2010; Seyda & Langer, 2020). Similarly, pupils with higher BMI are often perceived as less competent regardless of their actual motor performance — a bias more closely linked to assumptions about body size than to objective ability (Baumgartner, 2022; Ferrari et al., 2022; Niederkofler et al., 2018).

Research consistently shows that teachers often struggle to assess pupils' performance accurately (Helmke, 2017; Südkamp et al., 2012), a finding that is also well documented in sport- and PE-specific research (Niederkofler et al., 2018). Many tend to overestimate both pupils' competence levels and their own judgement accuracy, thereby limiting the precision of their evaluations. Even experienced teachers are subject to cognitive distortions such as the halo effect or confirmation bias, which may compromise objectivity (Ferrari & Baumgartner, 2021; Van Ophuysen & Behrmann, 2015). In PE, where evaluation is often based on spontaneous and largely unstructured visual observation of pupils' motor behaviour during ongoing instructional activities (Urhahne & Wijnia, 2021), the predominance of perceptual impressions may increase the risk of biased or inconsistent judgements, potentially affecting diagnostic reliability. Overconfidence in teachers perceived judgement accuracy poses a significant risk, as it can lead to distorted assessments and expectations that may negatively influence pupils' motivation and development through self-fulfilling processes (Ferrari et al., 2022; Good et al., 2018). To minimise such biases, teachers must develop metacognitive awareness of the subjective elements inherent in diagnostic decision-making. In this context, self-efficacy — understood as the belief in one's ability to accurately observe, assess, and interpret pupils' competencies — emerges as a central determinant of judgement accuracy (Bandura, 1997). While contextual factors such as experience, pedagogical training, and class size may influence diagnostic competence (Baumert & Kunter, 2013; Klug et al., 2013; Südkamp et al., 2012), they fall beyond the scope of this study, which focuses primarily on judgement accuracy.

1.4 Research questions and Hypotheses

From a conceptual standpoint, diagnostic competence in teachers and motor competence in pupils can be seen as parallel and mutually informing constructs: both integrate cognitive, motivational, and performance-related dimensions. While pupils demonstrate their motor competence through observable performance, teachers express their diagnostic competence through the accuracy of their professional judgements, captured in this study through anticipatory LEMOK ratings completed before the MOBAK tests. This dual perspective reflects the reciprocal nature of teaching and learning processes in PE, where diagnostic and motor competences interact dynamically.

Building on this theoretical and empirical foundation, the present study investigates how accurately primary school PE teachers assess their pupils' BMC by comparing their anticipatory LEMOK judgements with objectively measured MOBAK performance. The analyses draw on a large dataset from pupils in year four and five of primary school in the Swiss Canton of Ticino, combining the MOBAK test (Herrmann, 2018) with teachers' LEMOK ratings for comparison.

It also examines whether judgement accuracy varies across the two MOBAK domains of OM and SM, which reflect children's ability to handle objects and to perform movements with their own body. Finally, the study explores the extent to which gender and BMI shape teachers' judgements, considering both potential direct influences and indirect effects mediated through pupils' actual motor performance.

Drawing on the theoretical framework and research questions, three hypotheses guide the present study.

- 1) Teachers' judgements (LEMOK) are expected to show a positive correlation with objectively measured MOBAK scores, indicating diagnostic competence through judgement accuracy (Blömeke et al., 2015; Brühwiler, 2017).
- 2) Teachers are assumed to demonstrate greater judgement accuracy in the OM domain than in the SM domain, as OM tasks are typically more structured and observable (Ferrari et al., 2022; Herrmann & Seelig, 2020).
- 3) Gender and BMI are not expected to directly influence teachers' judgements, but rather to exert indirect effects mediated by pupils' actual motor performance (Barnett et al., 2010; Südkamp et al., 2012).

Together, these hypotheses translate the conceptual model into empirically testable propositions, linking diagnostic competence to measurable indicators of judgement accuracy.

2 Material and Methods

The presented cross-sectional study focuses on data from the fourth measurement point (2024) of the longitudinal study «Development of Basic Motor Competencies in Children (EMOKK: Entwicklung Motorischer Kompetenzen in der Kindheit)» which took place in the Swiss Canton of Ticino. The Switzerland-wide study was funded by the Swiss National Science Foundation (SNSF; Grant no. 200840). The study was conducted in accordance with the Declaration of Helsinki and was approved by the Directorate of Education, Culture and Sport of the Canton of Ticino (No. 2021-00252, Rif CE Ti 3819) and by the University Commission for Ethical Research in Geneva (CUREG-2023-01-05). Prior to the testing, both teachers and parents or legal guardians of the children provided written consent for participation. The children verbally agreed to participate.

The school system in the Canton of Ticino has a peculiar feature: PE teachers in primary schools are specialised professionals, with a university degree (Bachelor's or Master's) and pedagogical preparation at the University for Teacher Education aimed at teaching PE exclusively. Each PE teacher supervises several classes, often in different schools, delivering three teaching units (45min lessons) per week in each class (2h15min per class in total).

2.1 Data collection

Data were collected from three different sources: (1) pupils' MOBAK scores, assessed using the MOBAK test instrument (Herrmann, 2018; Herrmann & Seelig, 2020); (2) the LEMOK, based on a questionnaire completed by the PE teacher for each child; and (3) teacher-related information obtained through a separate teacher questionnaire (Herrmann et al., 2024).

BMC (MOBAK test)

The MOBAK test instrument assesses children's BMC. In this study, BMC are operationalised through the validated MOBAK test battery. In the MOBAK test, these competencies are assessed based on eight items that can be divided into two competency areas: SM tested through the four items balancing, rolling, jumping, and running; and OM tested with the items throwing, catching, dribbling and bouncing.

The MOBAK test instruments were developed for preschool (Herrmann et al., 2020), as well as year one to six of primary school (Herrmann, 2018; Herrmann & Seelig, 2020). The test is designed with an appropriate age level of difficulty. The tasks were developed to assess motor competencies in accordance with the objectives defined in the official school curricula (D-EDK, 2017; DECS, 2015). The MOBAK tests were validated in several studies with confirmatory factor analyses (Herrmann et al., 2016; Herrmann & Seelig, 2020).

The assessment of the MOBAK test items is easily implemented due to a dichotomous coding (1 = pass or 0 = fail) for each test item and trial based on clear standardised criteria. No practice trial was allowed. Children performed each test item twice except for the items *throwing* and *catching*. Therefore, for each test item, a maximum of 2 points could be reached. The number of successful attempts was recorded, ranging between 0 = no successful attempts, 1 = one successful attempt and 2 = two successful attempts. For the items *throwing* and *catching* children had six attempts and the number of hits or caught balls was recorded. 0-2 hits or caught balls were given 0 points, 3-4 hits or caught balls were coded as 1 point, 5-6 hits were coded as 2 points. For the test evaluation, the sum of the scores of the

individual items in the OM and SM competence areas was calculated, resulting in a score of 0-8 points.

Data was collected between February and March 2024, during regular PE classes. Pupils were divided into groups of 3-4 children and performed the MOBAK 3-4 or the MOBAK 5-6 tests depending on the class level, under the supervision of trained staff.

Teachers' judgement (LEMOK)

Teachers made individual judgements about pupils' motor performance approximately one week before the MOBAK testing sessions. Each PE teacher received the LEMOK questionnaire in advance and was asked to complete one form for each child based on their prior knowledge and observations during regular PE lessons. To assess teachers' judgement (LEMOK), teachers were asked to indicate, for each child in the class, how they judged the child's motor competence in eight tasks based on the MOBAK test items (Herrmann et al., 2024). Since this teacher-based LEMOK version is an internal judgement instrument developed specifically within the EMOKK project and is not yet published as an independent tool, no external source is currently available¹. In the questionnaire, the eight test items that the teachers had to evaluate were described based on the required competence. For example, one item asked: *«Can the child hit a target with a ball?»*

All participating PE teachers had access to detailed descriptions of the MOBAK tasks to ensure a consistent understanding of the required competences. Moreover, they were already familiar with the MOBAK test procedures, as they had participated in previous EMOKK data collections. The answer for each item was given on a trichotomous scale with the options "No", "Maybe" and "Yes".

The assumption behind this form of questioning children's motor competencies is that teachers also evaluate the children's competencies within the construct of the competence areas of SM and OM, each represented by four test items, resulting in a total of eight items. In previous studies, teachers carried out the judgement at class level (Ferrari et al., 2022). To align the child-centred judgement with the MOBAK test, the teachers' ratings were scored on a 3-point scale. The answers were recorded with 0 («No»), 1 («Maybe»), and 2 («Yes») analogous to the MOBAK test coding, which is assessed with no, one and two successful attempts, resulting in a maximum of 8 points per competence area.

1 For further information on LEMOK, the authors can be contacted directly.

2.2 Sample

In total, $N = 339$ children² (age: $M = 9.94$ years, $SD = 0.37$, 51.0% girls) in year four of primary school from 21 classes and $N = 350$ (age: $M = 10.94$ years, $SD = 0.38$, 50.9% girls) children in year five of primary school from 22 classes were included. The average class size was $n = 16.10$ children in year four of primary school and $n = 15.91$ children in year five of primary school. The assessed participants came from ten different schools, evenly distributed across the Canton of Ticino. The 43 classes were taught by 14 different PE teachers (age: $M = 42.48$ years, $SD = 9.17$, 71.4% female). Data about the MOBAK tests of children in year four and five of primary school, the teachers' judgements of BMC (LEMOK) and further information about the teachers were collected. Table 1 illustrates the extent of data available for the LEMOK and MOBAK tests, divided by school years and the two competence areas.

Table 1. Overview of the sample characteristics of classes in year four and five with respect to availability. It reports the number of children from whom basic motor competencies (MOBAK) were assessed, the number of children with teacher judgements of basic motor competencies (LEMOK), and the number of children from whom both data sources are available.

		MOBAK	LEMOK	MOBAK and LEMOK
		n	n	n
Year 4	Object Movement	315	336	313
	Self Movement	315	337	314
Year 5	Object Movement	329	345	324
	Self Movement	329	342	321

2.3 Data analysis

After data collection and processing, the datasets of the MOBAK test, the teacher's judgement (LEMOK) and the information about the teachers were merged. Multi-variate analysis with Mplus 8.8 (Muthen & Muthen, 2017) was then conducted as follows. The modelling of latent structures was carried out using a three-step approach. First, confirmatory factor analyses (CFA) were conducted to assess the factorial validity of the MOBAK and LEMOK instruments separately. Second, the relationship between MOBAK scores and teachers' judgements in the two motor competence areas was examined via a structural equation model (SEM) which serve as the basis for assessing judgement accuracy. Third, a structural equation

2 Number of children from whom at least one of the test forms was collected.

modelling was used to examine the associations between judgement accuracy and the covariates gender and BMI, in order to determine whether gender and BMI influence teachers' judgements directly or indirectly through their relationship with BMC. The motor competence areas SM and OM differ regarding the covariates gender and BMI (Ferrari et al., 2022; Herrmann & Seelig, 2020). Based on this knowledge, an analysis was conducted to determine whether gender and BMI represent a type of bias in the LEMOK assessment by teachers. A mediation analysis was carried out between the LEMOK (dependent variable (DV)), gender and BMI (both as independent variables (IV)) with the mediator variable MOBAK. These analyses were calculated separately for the two competence areas SM and OM.

The following recommended criteria were used to evaluate a good model fit: a comparative fit index of CFI > .95, although values of CFI > .90 are considered acceptable; and a root mean square error of approximation RMSEA < .06 (Hu & Bentler, 1999). Significance tests were conducted and reported two-sided and *p* values below .05 were regarded as statistically significant. The multi-level structure was considered by correcting the standard errors using the «type = complex» function.

Further descriptive and correlational analyses were conducted using SPSS 28 (IBM Corp., 2021). On a manifest level, descriptive statistics regarding MOBAK and LEMOK were calculated. For this, the sum values of MOBAK OM and SM as well as the LEMOK OM and SM were calculated. The teachers' judgement accuracy was calculated by the difference between the LEMOK and the MOBAK value in the respective area. Differences in mean values were analysed using independent samples *t*-tests. In addition to the mean values, standard deviations (*SD*) and effect sizes³ (Cohen's *d*) were calculated for the teachers' judgement accuracy. This comparison between teacher ratings and performance scores reflects the level component of judgement accuracy, as conceptualised by Urhahne and Wijnia (2021), and serves as an indicator of diagnostic realism rather than discriminatory precision. Due to the change in the MOBAK test instrument between pupils in year four and five of primary school, all mean values and analyses were done separately for each year group.

Missing values occurred due to the use of different data sources (MOBAK assessment, LEMOK assessment) and variations in survey days (see Table 1). In some cases, children participated in the MOBKA assessment, but the teachers did not assess the LEMOK, and vice versa. Among children who took part in the MOBKA assessment, the proportion of missing values ranged from 7.1% (MOBKA catching) to 9.1% (MOBKA rolling) in year four of primary school and from 6.0% (MOBKA throwing) to 7.4% (MOBKA dribbling) in year five of primary school. In the

3 The interpretation of Cohen's *d* followed Cohen (1988) guidelines, categorising effect sizes as small (*d* = 0.10), medium (*d* = 0.50), and large (*d* = 0.80).

LEMOK assessment, missing values were observed in all items ranging from 0.6% (LEMOK jumping and running) to 1.5% (LEMOK catching) in year four of primary school and 1.4% (LEMOK bouncing) to 2.9% (LEMOK running) in year five of primary school.

The full information maximum likelihood (FIML) algorithm was used for latent analysis to estimate missing values. This procedure minimises bias in the sample composition by preserving the full sample size and is widely used in educational research (Urban & Mayerl, 2014). In contrast, for descriptive manifest statistics, missing data were handled using listwise deletion meaning that only children who presented MOBAK and LEMOK data in both competence areas were included (see Table 1).

3 Results

3.1 Correlation between LEMOK and MOBAK

The confirmatory factor analysis (CFA) confirmed an acceptable factorial validity of the LEMOK instrument for the two factors SM and OM in the classes in year four of primary school ($\chi^2 = 50.312$, $df = 19$, $p < .001$, CFI = .977, RMSEA = .070, $N = 337$) with excellent factor reliabilities (FR SM = .87; FR OM = .93). In classes in year five of primary school, the CFA also confirmed the factorial validity of the LEMOK instrument ($\chi^2 = 37.599$, $df = 19$, $p = .007$, CFI = .993, RMSEA = .053, $N = 345$), with excellent factor reliabilities (FR SM = .88; FR OM = .93).

The test of the two-factor structure of the MOBAK 3-4 instrument showed good model fit ($\chi^2 = 21.04$, $df = 19$, $p = .33$, CFI = .988, RMSEA = .018, $N = 315$) in year four of primary school. Both latent factors demonstrated satisfactory factor reliabilities (FR), which is notable given the small number of items per factor (FR SM = .58; FR OM = .66; see Weiber & Mülhhaus, 2014). The model fits and factor reliabilities of the selected subsample are consistent with those of the overall EMOKK sample (Herrmann et al., 2025).

In year five of primary school, the factorial validity ($\chi^2 = 22.84$, $df = 19$, $p = .24$, CFI = .977, RMSEA = .025, $N = 329$) and factor reliabilities (FR SM = .65; FR OM = .63) of the MOBAK 5-6 test instrument were also confirmed.

To investigate correlations between pupils' MOBAK scores and the teachers' judgements (LEMOK), as indicators of judgement accuracy, SEM was performed between the two latent factors of MOBAK (OM and SM) and the two factors of LEMOK (OM and SM) for both year groups separately (see Appendix figures 1 and 2). The four-factor SEM resulted for both year groups in a good model fit (year four of primary school: $\chi^2 = 137.350$, $df = 98$, $p = .005$, CFI = .980, RMSEA = .034, $N = 350$; year five of primary school: $\chi^2 = 138.285$, $df = 98$, $p = .0046$, CFI = .980, RMSEA = .035, $N = 338$).

The correlation between the factors MOBAK OM and LEMOK OM was $r = .66$ ($p < .001$) in year four of primary school and $r = .72$ ($p < .001$) in year five of primary school (Table 2). For the factors MOBAK SM and LEMOK SM the correlations were higher in year four of primary school $r = .76$ ($p < .001$) than in year five of primary school $r = .46$ ($p < .001$). The correlation between MOBAK OM and LEMOK SM was significant in both year groups (year four of primary school: $r = .65$, $p < .001$; year five of primary school: $r = .39$, $p < .001$), whereas the correlation between MOBAK SM and LEMOK OM was moderate but only significant in year four of primary school (year four of primary school: $r = .51$, $p < .001$; year five of primary school: $r = .18$, $p = .092$). The correlations within the MOBAK construct between OM and SM were $r = .48$ ($p < .001$) in the year four of primary school and $r = .34$ ($p < .001$) in the year five of primary school. The correlations between LEMOK SM and OM resulted as follows: $r = .89$ ($p < .001$) in year four of primary school and $r = .81$ ($p < .001$) in year five of primary school.

Table 2. Latent correlations of the SEM of the two basic motor competence areas (MOBAK OM and MOBAK SM) and the teachers' judgements of the motor competencies in the two competence areas (LEMOK OM and LEMOK SM) for year four and five.

	(1)	(2)	(3)	(4)
(1) LEMOK OM		.89**	.66**	.51**
(2) LEMOK SM	.81**		.65**	.76**
(3) MOBAK OM	.72**	.39**		.48**
(4) MOBAK SM	.18	.46**	.34**	

Note. Above the diagonal values from the year four of primary school, below from the year five of primary school.
** $p < .01$

3.2 Judgement accuracy and the role of gender, BMI, and BMC

Figure 1 shows SEM with LEMOK OM as the outcome (DV), while the standardized path coefficients (β) and their significance for both year four and five of primary school are presented. The analysis showed that the two independent variables gender (boys = 1 and girls = 2) and BMI were not directly significantly associated with LEMOK OM with the only exception in year five of primary school where the BMI had a small significant effect ($\beta = .12$, $p = .038$) on LEMOK OM. However, gender had a moderate to large significant effect on the MOBAK OM (year four of primary school $\beta = -.38$, $p < .001$; year five of primary school $\beta = -.51$, $p < .001$) with boys performing better. Also, the MOBAK OM had a significant association with the LEMOK OM (year four of primary school $\beta = .62$, $p < .001$; year five of primary school $\beta = .70$, $p < .001$). Gender and BMI seem not to be directly

associated with LEMOK OM. However, gender seems to be important for the MOBAK OM assessment, and this is in turn important for the teacher’s judgement accuracy LEMOK OM. The association between gender and LEMOK OM was partially and significantly mediated by MOBAK OM. This results in an indirect effect of gender over MOBAK OM on LEMOK OM of $\beta = -.234^{**}$, $p < .001$ in year four of primary school and $\beta = -.354^{**}$, $p < .001$ in year five of primary school. The indirect effects of BMI over the MOBAK OM on the LEMOK OM were for both years small and not significant (year four of primary school $\beta = -.062$, $p = .121$; year five of primary school $\beta = -.040$, $p = .377$).

The model had acceptable model fits for both school classes (year four of primary school: $\chi^2 = 68.111$, $df = 31$, $p < .001$, CFI = .967, RMSEA = .059, $N = 339$; year five of primary school: $\chi^2 = 65.590$, $df = 31$, $p < .001$, CFI = .977, RMSEA = .056, $N = 350$). Overall, the model with gender, BMI and MOBAK OM accounted for 43% in year four of primary school and 47% in year five of primary school of the variance in LEMOK OM.

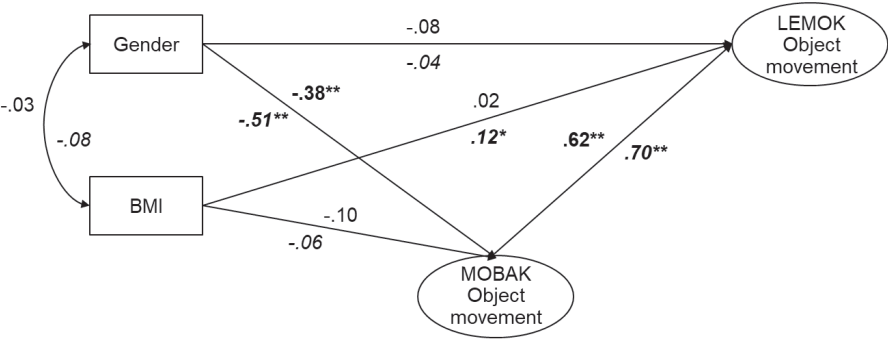


Figure 1. Mediation model of gender and BMI (IV), MOBAK OM (mediator), and LEMOK OM (DV).

Note. Above the arrows are the standardised regression weights (Std B) of year four of primary school, below in italics those of year five of primary school; significance is indicated by $p < .05$ (*), $p < .001$ (**).

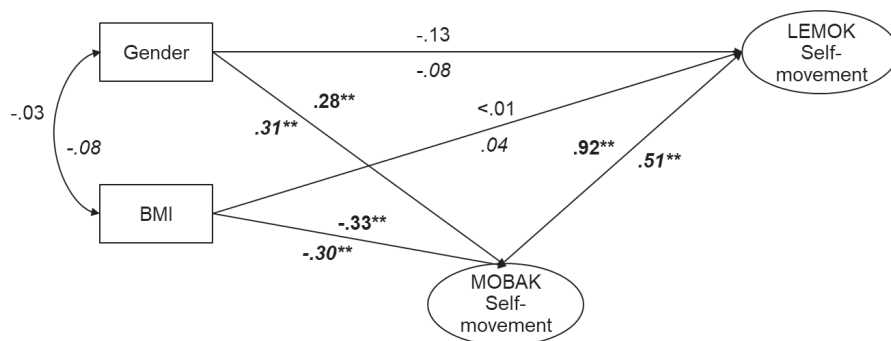


Figure 2. Mediation model of gender and BMI (IV), MOBAK SM (mediator), and LEMOK SM (DV).

Note. Above the arrows are the standardised regression weights (Std B) of year four of primary school, below in italics those of year five of primary school; significance is indicated by $p < .05$ (*), $p < .001$ (**).

The SEM with LEMOK SM as outcome is shown in Figure 2. Analogue to the calculation in Figure 1 the path coefficients indicate that gender (boys = 1 and girls = 2) and BMI are not directly associated with LEMOK SM. However, gender (year four of primary school $\beta = .28$, $p < .001$; year five of primary school $\beta = .31$, $p < .001$) and BMI (year four of primary school $\beta = -.33$, $p < .001$; year five of primary school $\beta = -.30$, $p < .001$) are significantly associated with MOBAK SM. In turn, the direct effects of MOBAK SM show an important significant association with LEMOK SM (year four of primary school $\beta = .92$, $p < .001$; year five of primary school $\beta = .51$, $p < .001$). The indirect effects of the mediation show, that the association between gender and LEMOK SM was mediated partially by MOBAK SM (year four of primary school $\beta = .26$, $p < .001$; year five of primary school $\beta = .16$, $p = .001$). This finding is consistent with the association of BMI and LEMOK SM, which is also significantly mediated by MOBAK SM (year four of primary school $\beta = -.31$, $p < .001$; year five of primary school $\beta = -.15$, $p < .001$).

The model presented a good model fit (year four of primary school: $\chi^2 = 50.870$, $df = 31$, $p < .001$, CFI = .925, RMSEA = .053, $N = 339$; year five of primary school: $\chi^2 = 70.512$, $df = 31$, $p < .001$, CFI = .934, RMSEA = .060, $N = 350$). Overall, the model with gender, BMI and MOBAK SM accounted for 79% in year four and 23% in year five of primary school of the variance in LEMOK SM.

3.3 Description of the sample at student level

The teachers generally estimated the children's BMC in both SM and OM to be high, with an average score of more than 6 points on a scale from 0 to 8. This trend was consistent across both class levels and competence areas (Table 3). In

comparison, the MOBAK scores of the pupils’ were lower in both competence areas for both classes. Additionally, the difference between the teachers’ judgements and pupils’ MOBAK scores - for each child is indicated as the teachers’ judgement accuracy.

The results indicate that teachers significantly overestimate the competence of the children in both year groups and in both competence domains OM (year four of primary school $d = 1.09$; year five of primary school $d = 1.61$) and SM (year four of primary school $d = 1.45$; year five of primary school $d = .95$).

Table 3. Mean values of the teachers’ judgement (LEMOK) and children’s basic motor competencies (MOBAK test), as well as the mean differences between the two measures across the competence areas self-movement and object movement in year four and five of primary school.

			Teacher Judge- ment (LEMOK)		MOBAK test		Teachers' judgement accuracy (difference between the teacher's assessment and the children's actual motor competence)		
		n	M	SD	M	SD	M	95%-CI	d
Year 4	Object Movement	313	6.66	1.69	4.50	1.97	2.16**	[1.94; 2.38]	1.09
	Self-Movement	314	7.18	1.28	4.62	1.85	2.56**	[2.37; 2.76]	1.45
Year 5	Object Movement	324	6.78	1.64	3.58	2.04	3.20**	[2.99; 3.42]	1.61
	Self-Movement	321	6.14	1.90	3.87	2.15	2.26**	[2.00; 2.52]	.95

Note. Children included in this table (n) have a value for LEMOK and MOBAK.
** The t-test is significant at the 0.01 level (two-sided).

4 Discussion

4.1 Interpreting main findings and hypotheses

The results of this study contribute to a clearer understanding of the judgement accuracy of primary school PE teachers, i.e. their capacity to make accurate and reliable assessments of children’s BMC measured by the MOBAK test. The following discussion interprets the main findings with regard to the research questions and theoretical assumptions.

A central finding of this study is the significant correlation between pupils’ MOBAK scores and teachers’ judgements (LEMOK). This can be interpreted as an indicator of teachers’ diagnostic competence, insofar as it suggests that teachers’ judgements

are at least partially grounded in systematic observation and reflective processing. This aligns with international research on diagnostic competence (Herppich et al., 2018; Südkamp et al., 2012), which emphasises the importance of systematic observation and reflexive information processing. The correspondence also reflects the idea that judgement accuracy represents the observable outcome of teachers' diagnostic competence.

However, the analysis also reveals a systematic tendency for PE teachers to overestimate motor competencies: mean LEMOK scores exceed MOBAK scores across domains and in both year groups. This significant overestimation (see Table 3), already observed in the literature at class level (Ferrari et al., 2022), may reflect overconfidence in one's evaluative capacity, or an evaluation influenced by general impressions not verified through structured observation possibly shaped by cognitive biases such as the halo effect or confirmation bias (Seyda & Langer, 2020; Urhahne & Wijnia, 2021). These evaluative distortions can affect expectations, planning, and pupil support (Good et al., 2018). This underscores the importance of professional learning opportunities that help teachers calibrate their perceptions based on empirical evidence rather than intuition. This tendency is particularly evident in the OM domain, where correlations of $r = .66$ in year four of primary school and $r = .72$ in year five of primary school suggest a moderate but consistent level of judgement accuracy. A possible explanatory factor is the greater exposure to structured and clearly defined OM tasks in upper year groups, such as throwing, catching, or dribbling, which offer teachers more consistent and comparable opportunities for observation. These repeated and observable performances make it easier for teachers to judge pupils' competence accurately (Herrmann & Seelig, 2020).

In the SM domain, a decrease in judgement accuracy appears from year four of primary school ($r = .76$) to year five of primary school ($r = .46$). While this may suggest a reduced ability to assess BMC, it should be interpreted cautiously. Two pedagogical aspects may have contributed to this trend: (1) fewer opportunities to observe SM activities in higher classes/year groups (Herrmann et al., 2021), and (2) a shift from structured observation to more global impressions as pupils grow older (Hascher, 2008), which increases the risk of cognitive bias in teacher judgements (Ferrari et al., 2022). Additionally, the increased difficulty of the MOBAK 5-6 test may have made the BMC in pupils in year five of primary school less predictable for teachers and thus harder to anticipate accurately.

Another important finding is the high correlation between teachers' LEMOK ratings in the OM and SM domains. This indicates teachers' tendency to rely on a general impression of pupils' overall motor ability rather than forming distinct judgements for OM and SM. In practice, this suggests that teachers integrate multiple aspects of motor performance into a single global evaluation, reflecting a holistic but potentially less differentiated diagnostic approach. Such an approach is common in PE

contexts (Seyda & Langer, 2020) but may compromise judgement accuracy. Recent studies show that even PE teachers need structured tools and targeted training to accurately differentiate between the two domains (Gonzalez-Huenulef et al., 2023; Heemsoth & Wibowo, 2020).

A particularly encouraging finding concerns the mediation analysis, which shows no direct influence of pupils' gender or BMI on teachers' judgement accuracy. Instead, both variables act indirectly through pupils' MOBAK scores. The results confirm well-known trends: boys score higher on MOBAK OM tasks, while girls perform better in MOBAK SM tasks. This gap is reflected in teachers' ratings which tend to assess boys' motor competencies more positively, particularly in the domain of OM (Barnett et al., 2010; Herrmann & Seelig, 2020). However, SEM analysis showed that this relationship is entirely mediated by pupils' MOBAK scores: teachers do not appear to base their judgements directly on gender but rather on pupil's observed performance. This finding is encouraging from an educational perspective, as it highlights an assessment practice based more on observation and empirical data than on gender stereotypes (Gonzalez-Huenulef et al., 2023; Seyda, 2018).

Overall, the judgement accuracy of teachers is moderate and domain-specific. However, the data reveals two main challenges, including a general tendency to overestimate pupils' motor competencies and difficulties in distinguishing between the two motor domains. These aspects highlight an ongoing need to strengthen diagnostic training in both initial and continuing professional development, focusing on the use of structured observation tools (e.g., forms, rubrics, checklists) and fostering critical reflection regarding one's own assessment practices (Klug et al., 2013). To mitigate the effects of bias and subjectivity, teachers should be equipped with validated diagnostic tools and methodological training that foster objective, evidence-based assessment (Tschannen-Moran & Hoy, 2001). Therefore, the education and continuing professional development of PE teachers should not only strengthen their diagnostic and pedagogical professional knowledge but also develop a critical awareness of the cognitive processes that influence diagnostic judgement. Consequently, the development of diagnostic competence should be a central component of PE teacher training, with a focus on both structured observation and the use of tools and strategies that support fair and accurate assessment grounded in professional criteria as well as objective and scientifically validated data.

4.2 Methodological considerations and limitations

One methodological strength of this study is the combined use of objective data (MOBAK) and teachers' judgements (LEMOK), analysed by means of SEM models that made it possible to verify not only direct correlations but also mediating mechanisms between variables. Furthermore, the level of analysis is that of the

individual pupil, which gives the results greater precision and internal validity. The use of a large and representative sample of pupils from the school context in the Canton of Ticino increases the applicative relevance of the results. However, the relatively small number of participating teachers invites caution when generalising these findings to broader populations.

However, this study has some limitations. The cross-sectional design does not allow causal conclusions to be drawn, nor does it allow an assessment of how judgement accuracy develops over time. In addition, the use of the LEMOK instrument, which is based on a subjective rating scale, may have been influenced by implicit biases affecting teachers' judgements. This concern is consistent with findings from recent literature highlighting how subjective ratings can be influenced by factors unrelated to pupils' actual performance (Urhahne & Wijnia, 2021). A further limitation is that anticipatory ratings may reinforce expectation-based biases.

While the comparison between teachers' evaluations (LEMOK) and pupils' performances (MOBAK) offers a meaningful measure of judgement accuracy, its scope must be interpreted with caution. The LEMOK reflects teachers' anticipatory judgements before the MOBAK test, whereas the MOBAK tests captures only performance outcomes, not the underlying learning processes. Thus, judgement accuracy reflects a correspondence between subjective and objective indicators rather than the dynamic reasoning through which teachers interpret and promote learning.

A further methodological consideration concerns the MOBAK instrument itself. Although this is a validated and standardised measure of pupils' BMC, its structure focuses primarily on observable performance outcomes rather than on the diagnostic thought processes underlying motor learning and adaptation.

Consequently, the comparison between LEMOK and MOBAK captures only one dimension of teachers' professional judgement: the measurable correspondence between teachers' subjective, anticipatory evaluations and pupils' objective, performance-based results. In line with critical reflections (Herrmann & Seelig, 2020; Niederkofler et al., 2018), the validity of MOBAK as a diagnostic reference should be understood within the boundaries of what it measures. Future studies could integrate more process-oriented indicators to complement performance-based measures.

4.3 Practical and educational implications

From a practical standpoint, the results highlight the value of encouraging teachers to reflect critically on the accuracy of their own judgement. Beyond accuracy, diagnostic competence also involves understanding pupils' emotional and motivational states, which shape their engagement and learning in PE (Spinath, 2005). Supporting teachers in integrating both performance and contextual cues may strengthen diagnostic practice.

Furthermore, the tendency toward overestimation observed in this study deserves careful consideration. Although it may appear to be a simple quantitative bias, it can have a significant impact on teachers' expectations and educational planning. From a diagnostic perspective, overestimating pupils' motor competencies may lead to the selection of tasks whose demands are not aligned with pupils' actual competence levels, potentially resulting in overload and reduced opportunities for targeted support. Conversely, underestimation may lead to overly simplified tasks that limit challenge and learning opportunities. From this perspective, improving diagnostic competence means not only refining observational accuracy but also maintaining a balanced perception of pupils' actual learning needs. Simple structured tools (e.g., rubrics) may help reduce reliance on global impressions.

5 Conclusion

This study analysed the judgement accuracy of primary school PE teachers in assessing pupils' BMC, also considering the indirect influence of gender and BMI. Results reveal moderately accurate judgements, with significant correlations between teachers' judgements of pupils' BMC (LEMOK) and pupils' MOBAK performance, but also a systematic overestimation across domains (OM, SM) and school levels (year four and five of primary school).

At the same time, it is important to acknowledge that the MOBAK test, while validated, captures only pupils' BMC, operationalised through eight standardised performance tasks. However, diagnostic competence in PE also involves interpretative, reflective and context-sensitive processes (e.g. consideration of situational and motivational factors) that are not directly measurable through performance tests. Therefore, while judgement accuracy constitutes a valid empirical indicator, it captures only the observable outcome of teachers' diagnostic reasoning, which results from latent processes of perception, interpretation and decision making, without providing insight into how teachers reflect on their judgements or adapt them to situational and pedagogical contexts (Blömeke et al., 2015). Moreover, as Baumgartner (2022) notes, even standardised diagnostic instruments can only partially reflect the situational and interpretative nature of professional judgement. Future research should therefore integrate complementary multimodal and qualitative approaches to better capture the dynamic interplay between perception, interpretation, and pedagogical decision-making.

From a broader epistemological perspective, these findings reaffirm the reciprocal link between diagnostic and motor competencies. Teachers' ability to make accurate judgements not only reflects their professional expertise but also contributes to pupils' competence development through adaptive teaching. Strengthening diagnostic competence in PE thus represents both a pedagogical and an epistemic commitment to connect evidence-based assessment with meaningful learning in con-

texts of movement. Beyond its empirical contribution, this study also highlights an epistemological challenge: diagnostic competence, as a professional construct, lies at the intersection between measurable performance and interpretative understanding. Its assessment through judgement accuracy provides valuable insights, yet it cannot capture the reflective and situational nature of teachers' diagnostic reasoning. Recognising this tension is essential for advancing both theoretical models and empirical approaches in research on teacher competence.

6 References

- Artelt, C., & Gräsel, C. (2009). Diagnostische Kompetenz von Lehrkräften. *Zeitschrift Für Pädagogische Psychologie*, 23(3–4), 157–160. <https://doi.org/10.1024/1010-0652.23.34.157>
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. W. H. Freeman.
- Barnett, L. M., Van Beurden, E., Morgan, P. J., Brooks, L. O., & Beard, J. R. (2010). Gender differences in motor skill proficiency from childhood to adolescence: A longitudinal study. *Research Quarterly for Exercise and Sport*, 81(2), 162–170. <https://doi.org/10.1080/02701367.2010.10599663>
- Baumert, J., & Kunter, M. (2013). Professionelle Kompetenz von Lehrkräften: Ergebnisse des Forschungsprogramms COACTIV. *Zeitschrift Für Erziehungswissenschaft*, 16(2), 233–253.
- Baumgartner, M. (2022). Professionelle Kompetenz(en) von Sportlehrkräften: Begriffe, Traditionen, Modellierungen und Perspektiven. In R. Sygusch, J. Hapke, S. Liebl, & C. Töpfer (Hrsg.), *Kompetenzorientierung im Sport: Grundlagen, Modellentwurf und Anwendungsbeispiele* (pp. 27–41). Hofmann.
- Blömeke, S., Gustafsson, J.-E., & Shavelson, R. J. (2015). Beyond dichotomies. *Zeitschrift Für Psychologie*, 223(1), 3–13. <https://doi.org/10.1027/2151-2604/a000194>
- Brühwiler, C. (2017). Diagnostische und didaktische Kompetenz als Kern adaptiver Lehrkompetenz. In A. Südkamp & A.-K. Praetorius (Hrsg.), *Diagnostische Kompetenz von Lehrkräften: Theoretische und methodische Weiterentwicklungen*. Waxmann.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. L. Erlbaum Associates.
- Deutschschweizer Erziehungsdirektoren-Konferenz (D-EDK). (2017). *Lehrplan 21*. <https://zh.lehrplan.ch/>
- Divisione della Scuola. (2015). *Piano di studio della scuola dell'obbligo del Canton Ticino*. Repubblica e Cantone Ticino, Dipartimento dell'Educazione, della Cultura e dello Sport.
- Ferrari, I., & Baumgartner, M. (2021). Der Kompetenzansatz im Bewegungs- und Sportunterricht. In Ch. Hayoz, N. Lanthemann, G. Patelli, & G. Grossrieder (Hrsg.), *Kompetenzorientiertes Lernen und Lehren im Bewegungs- und Sportunterricht: Fachdidaktisches Referenzmodell*. HEP Verlag.
- Ferrari, I., Kühnis, J., Bretz, K., & Herrmann, C. (2022). Diagnostische Kompetenz der Lehrpersonen und deren Bedeutung für die Förderung motorischer Basiskompetenzen. In R. Messmer & C. Krieger (Eds), *Narrative zwischen Wissen und Können* (pp. 175–194). Nomos. <https://doi.org/10.5771/9783985720118-175>
- Gonzalez-Huenulef, Y., Martino Fuentealba, P., Bretz, K., Ferbol, C., & Carcamo-Oyarzun, J. (2023). Competencia motriz real y percibida en niños y niñas de la Patagonia chilena: La percepción del estudiantado y del profesorado de educación física. *Retos*, 50, 290–297. <https://doi.org/10.47197/retos.v50.99781>
- Good, T. L., Sterzinger, N., & Lavigne, A. (2018). Expectation effects: Pygmalion and the initial 20 years of research. *Educational Research and Evaluation*, 24(3–5), 99–123. <https://doi.org/10.1080/13803611.2018.1548817>
- Grotegut, L., & Klingsieck, K. B. (2025). Wie können unterschiedliche Aspekte diagnostischer Kompetenz gefördert werden? Drei Maßnahmen im Vergleich. *Zeitschrift Für Pädagogische Psychologie*, 39(1–2), 95–109.

- Hascher, T. (2008). Diagnostische Kompetenz im Lehrberuf. In R. F. B. Schulte, T. Hascher, & K. O. Bauer (Hrsg.), *Wissen erwerben, Kompetenzen entwickeln: Modelle zur kompetenzorientierten Lehrerbildung* (pp. 71–86). Waxmann.
- Heemsoth, T., & Wibowo, J. (2020). Fachdidaktisches Wissen von angehenden Sportlehrkräften messen. *German Journal of Exercise and Sport Research*, 50(2), 308–319. <https://doi.org/10.1007/s12662-020-00643-0>
- Helmke, A. (2017). *Unterrichtsqualität und Lehrerprofessionalität: Diagnose, Evaluation und Verbesserung des Unterrichts*. Klett-Kallmeyer.
- 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>
- Herrmann, C. (2018). *MOBAK 1-4: Test zur Erfassung motorischer Basiskompetenzen für die Klassen 1–4*. Hogrefe.
- Herrmann, C., Bretz, K., Kress, J., & Seelig, H. (2024). *Development of basic motor competencies during childhood (EMOKK): Documentation of items and scales. Survey 2024* [Technical report].
- Herrmann, C., Bretz, K., Kühnis, J., Keller, R., Seelig, H., Kress, J., & Ferrari, I. (2021). *MOBAK – Monitoring of Basic Motor Competencies of 4- to 8-year-old Children in Switzerland: Documentation of Items and Scales*. Pädagogische Hochschule Zürich. <https://doi.org/10.5281/zenodo.5726943>
- Herrmann, C., Ferrari Ehrensberger, I., & Bretz, K. (2025). *Assessing basic motor competencies in 3rd- and 4th-grade primary school children: The MOBAK-3-4 test* [Technical report]. <https://doi.org/10.5281/zenodo.17140288>
- Herrmann, C., Ferrari Ehrensberger, I., Wälti, M., Wacker, S., & Kühnis, J. (2020). *MOBAK-KG: Motorische Basiskompetenzen im Kindergarten – Testmanual (3. Aufl.)*. <https://doi.org/10.5281/ZENODO.3774435>
- Herrmann, C., Gerlach, E., & Seelig, H. (2016). Motorische Basiskompetenzen in der Grundschule. *Sportwissenschaft*, 46(2), 60–73. <https://doi.org/10.1007/s12662-015-0378-8>
- Herrmann, C., & Seelig, H. (2020). *MOBAK-5-6: Motorische Basiskompetenzen in der 5. und 6. Klasse – Testmanual (3. Aufl.)*. <https://doi.org/10.5281/ZENODO.3774445>
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6(1), 1–55. <https://doi.org/10.1080/10705519909540118>
- Hulteen, R., Morgan, P., Barnett, L., Stodden, D., & Lubans, D. (2018). Development of Foundational Movement Skills: A Conceptual Model for Physical Activity Across the Lifespan. *Sports Medicine*, 48(7), 1533–1540. <https://doi.org/10.1007/s40279-018-0892-6>
- Ingenkamp, K., & Lissmann, U. (2008). *Lehrbuch der pädagogischen Diagnostik (6. Aufl.)*. Beltz.
- Klug, J., Bruder, S., Kelava, A., Spiel, C., & Schmitz, B. (2013). Diagnostic competence of teachers: A process model that accounts for diagnosing learning behavior tested by means of a case scenario. *Teaching and Teacher Education*, 30, 38–46. <https://doi.org/10.1016/j.tate.2012.10.004>
- Kunter, M., Klusmann, U., Baumert, J., Richter, D., Voss, T., & Hachfeld, A. (2013). Professional competence of teachers: Effects on instructional quality and student development. *Journal of Educational Psychology*, 105, 805–820. <https://doi.org/10.1037/a0032583>
- Loibl, K., Leuders, T., & Dörfler, T. (2020). A Framework for Explaining Teachers' Diagnostic Judgements by Cognitive Modeling (DiaCoM). *Teaching and Teacher Education*, 91, 103059. <https://doi.org/10.1016/j.tate.2020.103059>
- McElvany, N., Schroeder, S., Hachfeld, A., Baumert, J., Richter, T., Schnotz, W., Horz, H., & Ullrich, M. (2009). Diagnostische Fähigkeiten von Lehrkräften bei der Einschätzung von Schülerleistungen und Aufgabenschwierigkeiten bei Lernmedien mit instruktionalen Bildern. *Zeitschrift Für Pädagogische Psychologie*, 23(3–4), 223–235.

- Moor, P. (1965). *Heilpädagogik: Ein pädagogisches Lehrbuch*. Huber. <https://books.google.ch/books?id=0DbnGAAACAAJ>
- Muthen, L. K., & Muthen, B. (2017). *Mplus (Version 8) user's guide*. Muthén & Muthén.
- Niederkofler, B., Ferrari, I., Kühnis, J., & Herrmann, C. (2024). Diagnostische Kompetenz von Lehrpersonen in der Einschätzung der motorischen Basiskompetenzen von Kindern. In *Motorische Basiskompetenzen: Konstrukt, Erfassung und Anwendung* (pp. 209–225). Springer. https://doi.org/10.1007/978-3-658-45759-4_11
- Niederkofler, B., Herrmann, C., & Amesberger, G. (2018). Diagnosekompetenz von Sportlehrkräften: Semiformelle Diagnose von motorischen Basiskompetenzen. *Zeitschrift für sportpädagogische Forschung*, 6(2), 72–96. <https://doi.org/10.5771/2196-5218-2018-2-72>
- Seyda, M. (2018). Können Sportlehrkräfte die Perspektive ihrer Schülerinnen und Schüler einnehmen? Eine Untersuchung über die Akkuratheit von Beurteilungen physischer Fähigkeitsselbstwahrnehmungen. *Unterrichtswissenschaft*, 46, 215–231. <https://doi.org/10.1007/s42010-017-0006-4>
- Seyda, M., & Langer, A. (2020). Die Orchestrierung des diagnostischen und didaktischen Handelns im Sportunterricht—Eine Frage der Kompetenz! 69, 353–357. <https://doi.org/10.30426/SU-2020-08-3>
- Shavelson, R. J. (1978). Teachers' Estimates of Student 'States of Mind' and Behavior. *Journal of Teacher Education*, 29(5), 37–40. <https://doi.org/10.1177/002248717802900511>
- Spinath, B. (2005). Akkuratheit der Einschätzung von Schülermerkmalen durch Lehrer und das Konstrukt der diagnostischen Kompetenz: Accuracy of teacher judgments on student characteristics and the construct of diagnostic competence. *Zeitschrift Für Pädagogische Psychologie*, 19(1/2), 85–95.
- Stodden, D. F., Goodway, J. D., Langendorfer, S. J., Roberton, M. A., Rudisill, M. E., Garcia, C., & Garcia, L. E. (2008). A Developmental Perspective on the Role of Motor Skill Competence in Physical Activity: An Emergent Relationship. *Quest*, 60(2), 290–306. <https://doi.org/10.1080/00336297.2008.10483582>
- Südkamp, A., Kaiser, J., & Möller, J. (2012). Accuracy of teachers' judgements of students' academic achievement: A meta-analysis. *Journal of Educational Psychology*, 104(3), 743–762. <https://doi.org/10.1037/a0027627>
- Tschannen-Moran, M., & Hoy, A. W. (2001). Teacher efficacy: Capturing an elusive construct. *Teaching and Teacher Education*, 17(7), 783–805. [https://doi.org/10.1016/S0742-051X\(01\)00036-1](https://doi.org/10.1016/S0742-051X(01)00036-1)
- Urban, D., & Mayerl, J. (2014). *Strukturgleichungsmodellierung: Ein Ratgeber für die Praxis. [Structural Equation Modeling. A Best Practice Guide]*. Springer VS. <https://doi.org/10.1007/978-3-658-01919-8>
- Urhahne, D., & Wijnia, L. (2021). A review on the accuracy of teacher judgements. *Educational Research Review*, 32, 100374. <https://doi.org/10.1016/j.edurev.2020.100374>
- 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, 82–98.
- Voss, T., Kunter, M., & Baumert, J. (2011). Assessing teacher candidates' general pedagogical/psychological knowledge: Test construction and validation. *Journal of Educational Psychology*, 103(4), 952–969. <https://doi.org/10.1037/a0025125>
- Weiber, R., & Mülhhaus, D. (2014). *Strukturgleichungsmodellierung*. Springer Gabler. <https://doi.org/10.1007/978-3-642-35012-2>
- Weinert, F. E. (2001). Concept of competence: A conceptual clarification. In D. S. Rychen & L. H. Salganik (Eds), *Defining and selecting key competencies* (pp. 45–65). Hogrefe & Huber.

Appendix

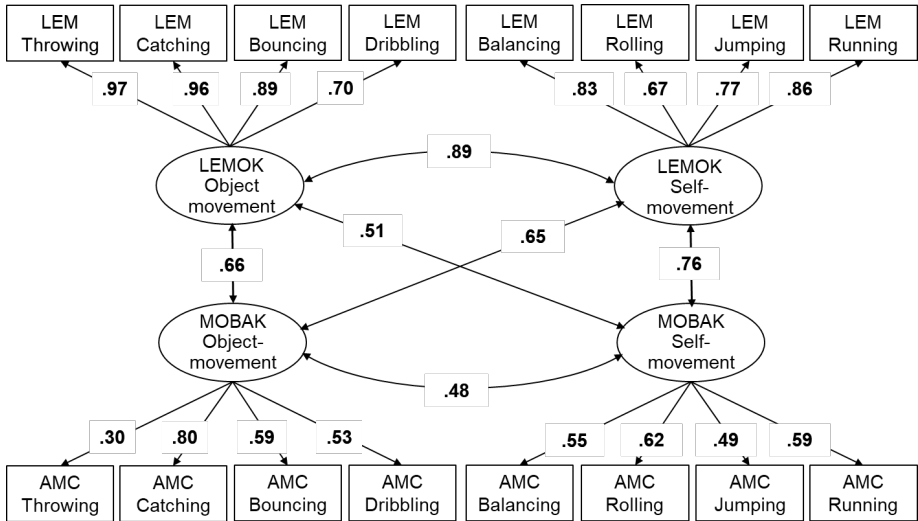


Figure 1. Structural equation model with the factor's pupils' MOBAK-3-4 performance (MOBAK OM and SM) and teachers' judgement (LEMOK OM and SM) for year four of primary school. Bold factors $p < .05$.

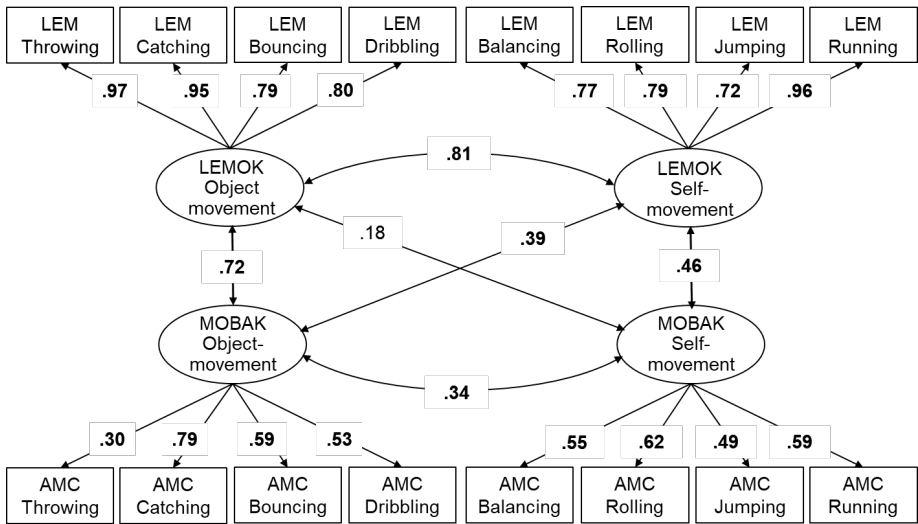


Figure 2. Structural equation model with the factor's pupils' MOBAK-5-6 performance (MOBAK OM and SM) and teachers' judgement (LEMOK OM and SM) for year five of primary school. Bold factors $p < .05$.

Aufgabenanalytische Kompetenzen angehender Sportlehrkräfte als eine Voraussetzung zur kognitiven Aktivierung von Schüler*innen im Sportunterricht

von Philipp Hendricks

Zusammenfassung: Im Studium sollten angehende Sportlehrkräfte aufgabenanalytische Kompetenzen erwerben, mit denen sie kognitiv aktivierende Aufgabenmerkmale (z.B. Operatoren) wahrnehmen, interpretieren und eine Entscheidung zur Ausprägung der kognitiven Anforderungen treffen können. Der Forschungsbeitrag geht im Rahmen einer Mixed-Methods Studie den Fragen nach, wie akkurat Sportstudierende kognitive Anforderungen beurteilen können und welche Informationsverarbeitungsprozesse währenddessen ablaufen. Zunächst beurteilten 323 Studierende die kognitiven Anforderungen von zehn Aufgaben. Anschließend wurden zehn der zuvor befragten Studierenden in einem qualitativen Interview mit integrierter Methode des lauten Denkens befragt. Die Analysen ergaben, dass die Studierenden die Ausprägung kognitiver Anforderungen leicht überschätzten. In den Interviews zeigte sich, dass die »Realist*innen« im Gegensatz zu den »Überschätzer*innen« vermehrt auf ihr Fachwissen zurückgriffen und umfangreichere entscheidungsrelevante Informationen nutzten. Der Beitrag gibt Hinweise darauf, dass die inhaltliche Bedeutung kognitiv aktivierender Aufgabenmerkmale in der Sportlehrkräftebildung verstärkt thematisiert werden sollte.

Schlagwörter: Sportlehrkräfteprofessionalisierung, diagnostische Kompetenzen, aufgabenanalytische Kompetenzen, kognitive Aktivierung, Mixed-Methods

Task analysis skills of pre-service physical education teachers as a prerequisite for the cognitive activation of students in physical education

Abstract: During their studies, pre-service PE teachers should acquire skills that enable them to perceive, interpret and make decisions regarding cognitively activating task characteristics (e.g., operators). This mixed-methods study investigates how accurately pre-service PE teachers can assess cognitive demands and what information-processing processes are involved in the assessment. First, 323 pre-service PE teachers evaluated the cognitive demands of ten tasks. Subsequently, ten of the previously surveyed pre-service PE teachers participated in qualitative interviews incorporating the think-aloud method. The analysis revealed that pre-service PE teachers tended to slightly overestimate the level of cognitive demand. The interviews showed that, unlike the "overestimators," the "realists" more frequently drew on content knowledge and utilized a broader range of decision-relevant information. The findings suggest that assessing cognitive demands should be more explicitly addressed in PE teacher education.

Keywords: physical education teacher education; diagnostic competence; task analysis skills; cognitive activation, mixed-methods research

Aufgabenanalytische Kompetenzen angehender Sportlehrkräfte als eine Voraussetzung zur kognitiven Aktivierung von Schüler*innen im Sportunterricht

1 Einleitung

Ein Ziel von Sportunterricht besteht darin, Schüler*innen in ihrer motorischen, kognitiven und sozialen-affektiven Entwicklung individuell zu unterstützen und sie zu einer urteils- und handlungsfähigen Teilhabe an der Sport- und Bewegungskultur zu befähigen (u.a. Gogoll, 2014). Daraus ergibt sich für Sportlehrkräfte die Herausforderung, nicht nur körperliche-motorische Lernprozesse, sondern auch wissens- und verständnisbezogene kognitive Lernprozesse systematisch anzubahnen (Hapke & Kramer, 2020; Sygusch et al., 2022). In diesem Kontext wurde in den letzten Jahren insbesondere die kognitive Aktivierung als eine Basisdimension der Unterrichtsqualität im Sportunterricht intensiv diskutiert (u.a. Engelhardt et al., 2024; Herrmann & Gerlach, 2020; Wibowo et al., 2024). (Sport-)Unterricht wird dann als kognitiv aktivierend verstanden, »wenn er Lernende zum vertieften Nachdenken und zu einer elaborierten Auseinandersetzung mit dem Unterrichtsgegenstand anregt« (Lipowsky, 2020, S. 92). Für die Auswahl kognitiv aktivierender Aufgaben, müssen Sportlehrkräfte die Aufgaben u.a. hinsichtlich ihrer Merkmale (z.B. das Potential zum Anregen anspruchsvoller Denkprozesse) beurteilen und an das individuelle Kompetenzniveau der Schüler*innen anpassen. Eine wesentliche Voraussetzung dafür besteht darin, dass Sportlehrkräfte unterrichtsrelevante Einschätzungen treffen, wofür sie diagnostische Kompetenzen benötigen.

Diagnostische Kompetenzen stellen einen Teil der professionellen Handlungskompetenzen dar (Krauss et al., 2004) und können als das Vermögen von Sportlehrkräften verstanden werden, die Lern- und Leistungsvoraussetzungen ihrer Schüler*innen sowie Merkmale von Aufgaben angemessen zu beurteilen (Praetorius & Südkamp, 2017). In den letzten Jahren wurden diagnostische Kompetenzen von Sportlehrkräften vor allem im Hinblick auf die Beurteilung von Schüler*innen untersucht (zsf. Langer & Hendricks, i.d.B.). Die Beurteilung von Aufgaben blieb im sportpädagogischen/-didaktischen Diskurs bisher weitestgehend unberücksichtigt (Hendricks, 2024; Leukel et al., 2023), obwohl sie eine Voraussetzung ist, um Aufgaben an die Lernvoraussetzungen der Schüler*innen anzupassen. Angehende Sportlehrkräfte sollten in ihrer Hochschulausbildung aufgabenanalytische Kompetenzen erwerben, mit denen sie die kognitiv aktivierenden Merkmale von Aufgaben beurteilen können. Dies ist eine Voraussetzung für die zielgerichtete und an den Lern- und

Leistungsvoraussetzungen der Schüler*innen orientierten kognitiven Aktivierung¹. Bislang fehlt ein differenziertes Bild darüber, ob angehende Sportlehrkräfte die kognitiv aktivierenden Merkmale in Aufgaben beurteilen können und welche informationsverarbeitenden Prozesse dabei ablaufen. Informationen hierzu sind notwendig, um die Sportlehrkräftebildung an den Voraussetzungen der Studierenden auszurichten. Der Forschungsbeitrag setzt an dieser Stelle an und stellt erste Ergebnisse des laufenden Forschungsprojektes »Aufgabenanalytische Kompetenzen angehender Sportlehrkräfte und deren Bedeutung für die Lernprozesssteuerung« an der TU Dortmund vor.

2 Theoretischer Hintergrund

2.1 Kognitiv aktivierende Aufgaben sind lernförderliche Aufgaben

Im Zuge der Diskussion um die Unterrichtsqualität im Fach Sport (u.a. Herrmann & Gerlach, 2020) resultierten vermehrt Überlegungen zur kognitiven Aktivierung durch Aufgaben (u.a. Engelhardt et al., 2024; Niederkofler & Herrmann, 2024). Die Zuwendung zu kognitiven Inhalten und Denkprozessen ist im Sportunterricht keineswegs neu (u.a. Brodtmann & Trebels, 1979). In jüngerer Zeit liegt den Überlegungen jedoch ein verändertes Verständnis von Lernen im Sportunterricht zugrunde. Lernen wird demnach als »aktiver, selbstgesteuerter, situativer und konstruktiver Prozess« betrachtet (Aschebrock et al., 2010, S. 13). In diesem Sinne gelten Aufgaben als kognitiv aktivierend, wenn sie einen Bezug zum Vorwissen aufweisen, offene Problemstellungen enthalten oder Metakognitionen anregen (Engelhardt et al., 2024; Widmer et al., 2024). Kognitiv aktivierende Aufgaben sollten u.a. durch Anregung zur Elaboration anspruchsvolle Denkprozesse auslösen und so komplex gestellt sein, dass sie nicht durch einfach abrufbares Wissen gelöst werden können. Auf diese Weise soll eine lediglich passive Auseinandersetzung der Schüler*innen mit dem Lerngegenstand sowie die Aneignung von trägem Wissen und blindem Können vermieden werden (u.a. Engelhardt et al., 2023; Fauth & Leuders, 2022; Heemsoth, 2024). In Anlehnung an gängige Lernzieltaxonomien zählen zu anspruchsvollen Denkprozessen bspw. das Analysieren, Entwickeln und Bewerten, wohingegen sich kognitiv anspruchslose Aufgaben durch reines Reproduzieren auszeichnen (u.a. Anderson et al., 2001). Eine Möglichkeit, die kognitiven Anforderun-

1 In der sportpädagogischen/-didaktischen Diskussion werden verschiedene Ansätze der (kognitiven) Aktivierung thematisiert (u.a. ästhetische Aktivierung bei Hartmann et al., 2023; zsf. Heemsoth, 2024). In diesem Forschungsbeitrag liegt der Fokus auf einem wissensorientierten Aktivierungsverständnis, das auf eine vertiefte mentale Auseinandersetzung mit einem sport- und bewegungsbezogenen Lerngegenstand abzielt (u.a. Engelhardt et al., 2024; Sygusch et al., 2022). Die Beurteilung motorischer Aspekte, die für eine gezielte Aktivierung einer bewegungsbezogenen Auseinandersetzung der Schüler*innen mit dem Lerngegenstand relevant sind, wird im Beitrag nicht thematisiert. Sie ist jedoch Teil des diesem Beitrag zugrundeliegenden Forschungsprojekts.

gen in sportunterrichtlichen Aufgaben zu bestimmen, stellt Gogoll (2014) im Modell der sport- und bewegungskulturellen Kompetenz dar (Tab. 1).

Tab. 1: Fünf Ausprägungen kognitiver Anforderungen (angelehnt an Gogoll, 2014)

Stufe	Kognitive Anforderungen	Beispiel
5	Wissen über multivariate Interdependenzen – Hierbei werden komplexe Ursachen-Wirkungs-Zusammenhänge beschrieben und fachwissenschaftlich begründet. Auf dieser Stufe werden sich wechselseitig beeinflussende Merkmale aufeinander bezogen. Es steht das Wissen über die gegenseitigen Abhängigkeiten im Fokus	Wie sollte ein optimales Aufwärmprogramm für einen Weitsprungwettbewerb aufgebaut sein? Beurteilt, inwiefern euer selbst entwickeltes Aufwärmprogramm euch für einen Weitsprungwettbewerb (z.B. die Bundesjugendspiele) optimal erwärmen würde. Diskutiert dabei, wie sich ein schlechtes Aufwärmen auf eure Leistung auswirken könnte.
4	Wissen über Ursache-Wirkungs-Zusammenhänge – Hierbei werden Ursache-Wirkungs-Zusammenhänge dargestellt und fachwissenschaftlich begründet. Auf dieser Stufe werden Abfolgen von Ereignissen und Zuständen aufeinander bezogen. Es steht das Wissen über die Zusammenhänge zwischen Ereignissen und Zuständen im Fokus.	Diskutiert, was dazu geführt hat, dass ihr weit gesprungen seid. Stellt Leitsätze für einen weiten Sprung auf, indem ihr den folgenden Satz beendet: »Um weit zu springen, ...«
3	Verlaufs-/Prozesswissen – Hierbei werden Verläufe und die dabei ablaufenden Prozesse beschrieben (z.B. Wenn-Dann-Beziehungen). Auf dieser Stufe steht das Wissen darüber, was während eines Prozesses passiert, im Fokus.	Beschreibe den Bewegungsablauf, wenn man mit der Hocksprungtechnik in die Weitsprunggrube springt.
2	Faktenwissen – Hierbei werden z.B. Fakten, Begriffe oder Konzepte, die einen fachwissenschaftlichen Bezug aufweisen, genannt oder zum Bearbeiten der Aufgabe genutzt. Auf dieser Stufe wird Faktenwissen reproduziert, ohne dabei Zusammenhänge zu erläutern, Erklärungen oder Argumentationen anzuführen und Vergleiche heranzuziehen.	Nenne die wichtigsten Bewegungsmerkmale der Hocksprungtechnik.
1	Unreflektiertes Erfahrungswissen – Hierzu zählen z.B. Wissensbestände der Schüler*innen, die ohne einen fachwissenschaftlichen Bezug genutzt werden. Einfache Vorher-Nachher-Vergleiche des eigenen sportlichen Handelns fallen in diese Stufe, sofern kein fachwissenschaftlicher Zusammenhang hergestellt wird.	Denkt an eure bisherigen Erfahrungen im Weitsprung. Was ist wichtig, damit man möglichst weit springen kann?

Die in Tab. 1 dargestellten Ausprägungen kognitiver Anforderungen enthalten beschreibbares Wissen über Situations-Verhaltens-Zusammenhänge, welches von Schüler*innen in Bewegungssituationen erlangt wird und für ein verständnisvolles Handeln genutzt werden kann. Im Gegensatz zu anderen sportpädagogischen/-didaktischen Ansätzen zur Beschreibung kognitiver Anforderungen (u.a. Frahm, 2022; Neumann, 2014; Pfitzner, 2013) bieten sie wesentliche Vorteile für empirische Untersuchungen. Die angenommene hierarchische Struktur wurde in anderen Fächern bereits empirisch bestätigt (Bernholt et al., 2009) und in einer sportdidaktischen Untersuchung hinsichtlich der angenommenen inneren Struktur sowie in Bezug auf ihre inhaltliche Validität in einer Expert*innenbefragung geprüft (Hendricks, 2024). Es wird angenommen, dass höhere Stufen auf niedrigeren Stufen aufbauen und durch die Verknüpfung von Elementen der darunterliegenden Stufen entstehen (Bernholt et al., 2009). Aufgaben, die eine Reproduktion von Faktenwissen erfordern, stellen somit eine niedrigere Anforderung dar als Aufgaben, die zu einer fachwissenschaftlichen Begründung mit komplexen Ursache-Wirkungs-Zusammenhängen auffordern. In Form von Operatoren (z.B. nenne, begründe, diskutiere) können sportunterrichtliche Aufgaben hinsichtlich ihrer Ausprägung der kognitiven Anforderungen unterschieden werden. Über Operatoren werden Denk- und Handlungsaufforderungen an die Schüler*innen formuliert. Sie stellen typische schulische Aufgabenstellungen dar, die eine kognitive Tätigkeit ausdrücken. Auf diese Weise ist es möglich, unterschiedlich kognitiv anspruchsvolle Aufgaben zu entwickeln und gezielt im Sportunterricht einzusetzen. Kognitiv anspruchsvolle Aufgaben haben sich vielfach als lernförderlich erwiesen (Widmer et al., 2024). Aufgaben, die Schüler*innen leicht überfordern und sie in ihrer individuellen Zone der nächsten Entwicklung fördern, gelten als besonders lernförderlich (Vygotsky, 1978). Hierfür ist es wichtig, dass Sportlehrkräfte die kognitiven Anforderungen in Aufgaben und die Voraussetzungen aufseiten der Schüler*innen möglichst genau beurteilen können. Für die Beurteilung der Ausprägung der kognitiven Anforderungen benötigen Sportlehrkräfte aufgabenanalytische Kompetenzen.

2.2 Aufgabenanalytische Kompetenzen von (Sport-)Lehrkräften

Aufgabenanalytische Kompetenzen beschreiben ein fachspezifisches kognitives Vermögen von (Sport-)Lehrkräften, das sich auf das aufgabenbezogene unterrichtliche Handeln bezieht. Sie befähigen (Sport-)Lehrkräfte dazu, Aufgaben hinsichtlich verschiedener lernrelevanter Merkmale zu beurteilen (Hendricks, 2024). Eine Voraussetzung für die Beurteilung von Aufgaben ist die Analyse ihrer Merkmale. Bei Beurteilungen wird der analysierten Aufgabe auf Grundlage einer Urteilsdimension (z.B. der kognitiven Anforderungen, siehe Tab. 1) eine Ausprägung zugeschrieben (Betsch et al., 2011). In Anlehnung an die kompetenzorientierte Sportlehrkräftebildungsforschung (Baumgartner, 2022) werden aufgabenanalytische Kompetenzen als Teil der diagnostischen Kompetenzen betrachtet und können als Kompetenzbereich

professioneller Kompetenzen in einem heuristischen Rahmenmodell veranschaulicht werden (Abb. 1). Das KompeSport-Modell (Baumgartner, 2022) integriert verschiedene Auffassungen von Kompetenzen in einem Typologie- und Topologiemodell professioneller Kompetenz(en) von Sportlehrkräften.

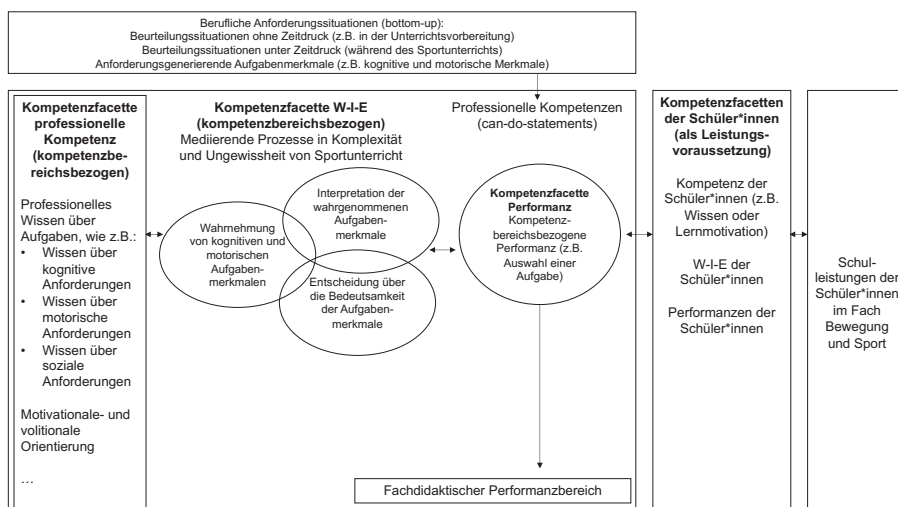


Abb. 1: Aufgabenanalytische Kompetenzen als Kompetenzbereich des KompeSport-Modells (Hendricks, 2025, adaptiert nach Baumgartner, 2022, S. 42; Blömeke et al., 2015)

Ausgangspunkt ist die berufliche Anforderungssituation, Aufgaben zu analysieren. Daraus lassen sich professionelle Kompetenzen ableiten, die in einem Zusammenhang mit dem Lernen der Schüler*innen stehen. Die *Kompetenzfacette professionelle Kompetenz* setzt sich aus latenten Aspekten zusammen (z.B. professionelles Wissen). Professionelles Wissen über Aufgaben bezieht sich u.a. auf ein am Lerngegenstand orientiertes Wissen über Aufgabenmerkmale. Es meint einerseits das Fachwissen (z.B. Kenntnisse über kognitive Aktivierung in sportunterrichtlichen Aufgaben) und andererseits das fachdidaktische Wissen (z.B. Wissen über kognitive Anforderungen in Aufgaben; Krauss et al., 2004). Die *Kompetenzfacette W-I-E* beschreibt die Wahrnehmungs-, Interpretations- und Entscheidungsprozesse, die der Aufgabenanalyse zugrunde liegen. In der Aufgabenanalyse werden kognitive Aufgabenmerkmale, wie bspw. die Operatoren im Aufgabentext wahrgenommen und ihre Bedeutung für die Ausprägung der Anforderungen unter Rückgriff auf das professionelle Wissen sowie auf Informationen über die Schüler*innen interpretiert. Im Entscheidungsprozess wird die Bedeutung entscheidungsrelevanter Informationen abgewogen und eine Entscheidung zur Ausprägung der Anforderungen getroffen. In Anlehnung an duale Prozessmodelle (u.a. Kahneman, 2003) kann davon ausgegangen werden, dass die Wahrnehmungs-, Interpretations- und Ent-

scheidungsprozesse entweder intuitiv oder reflexiv verlaufen (Loibl, 2020; Ostermann et al., 2019). In der Unterrichtsvorbereitung analysieren (Sport-)Lehrkräfte Aufgaben meist reflektiert und ohne größeren Zeitdruck. Hierbei kann eine reflektierte Analyse unter Rückgriff auf das professionelle Wissen erfolgen. Während des Unterrichts müssen Aufgaben hingegen aufgrund der zeitlich-räumlichen Organisationsstruktur (z.B. keine feste Sitzordnung, ständig wechselnde Position der Schüler*innen im Raum) (Neuber, 2020) und dem häufig prozessorientierten Charakter von Bewegungen (Neumaier, 2009) meist unter einem hohen Zeitdruck analysiert werden. Dadurch werden vermutlich weniger, jedoch eher subjektiv bedeutsame Informationen intuitiv verarbeitet (u.a. Bröder & Hilbig, 2017). Die *Kompetenzfacetten Performanz* beschreibt die Könnensleistung in einer beruflichen Anforderungssituation, wie beispielsweise die Auswahl einer Aufgabe im Unterricht. Dem heuristischen Modell folgend kann angenommen werden, dass die Auswahl von Aufgaben einen Einfluss auf die Lernprozesse und damit auch auf die Schulleistungen der Schüler*innen im Fach Sport hat. Das adaptierte KompeSport-Modell (Abb. 1) eröffnet die Möglichkeit, aufgabenanalytische Kompetenzen theoretisch zu modellieren und mit einem Schwerpunkt auf die unterschiedlichen Kompetenzfacetten empirisch zu untersuchen. So ist es bspw. möglich, durch einen Wissenstest das professionelle Wissen über Aufgaben (Kompetenzfacette professionelle Kompetenz) oder die Wahrnehmungs-, Interpretations- und Entscheidungsprozesse während der Aufgabenanalyse (Kompetenzfacette W-I-E) durch die Methode des lauten Denkens (Buber, 2009) zu erfassen.

Für die Auswahl von kognitiv aktivierenden Aufgaben, die an das Lern- und Leistungsniveau der Schüler*innen angepasst sind, müssen die kognitiven Anforderungen in Aufgaben akkurat beurteilt werden. Die Akkuratheit ist ein Maß, wie genau Lehrkräfte bestimmte Merkmale eines Beurteilungsgegenstands beurteilen (z.B. die Anforderungen in Aufgaben). Die Akkuratheit lässt sich über die vereinfachte Komponentenanalyse bestimmen (Schrader & Helmke, 1987). Hierbei können die Niveau-, die Differenzierungs- und die Rangkomponente unterschieden werden. In diesem Forschungsbeitrag wird der Fokus auf die Niveauebene gelegt, da ihr für die Auswahl von Aufgaben, die an das Lern- und Leistungsniveau der Schüler*innen angepasst sind, eine entscheidende Rolle zukommt. Die *Niveauebene* bezieht sich auf die Einschätzung des mittleren Leistungsniveaus der Klasse, einzelner Schüler*innen oder der Anforderungen von Aufgaben. Eine Lehrkraft schätzt bspw. die Anforderungen von Aufgaben ein. Anschließend wird die Einschätzung mit den tatsächlichen Anforderungen² verglichen. Auf diese Weise können Aussagen darüber getroffen werden, ob die Lehrkraft die Ausprägung der Anforderungen angemessen beurteilt, bzw. unter- oder überschätzt (Schrader & Helmke, 1987). In Anlehnung an Vygotsky (1978) kann davon ausgegangen wer-

2 Die »tatsächlichen Anforderungen« können empirisch bspw. über eine Expert*innenbefragung oder über die Lösungshäufigkeit der Aufgaben durch die Schüler*innen ermittelt werden.

den, dass sich eine akkurate Einschätzung und eine leichte Unterschätzung der kognitiven Anforderungen in Aufgaben lernförderlich auf die Schüler*innen auswirken. Eine Überschätzung führt dazu, dass tendenziell eher zu leichte Aufgaben gestellt werden, die kognitiv wenig fordernd sind und daher nicht lernförderlich wirken.

2.3 Empirische Befunde zu aufgabenanalytischen Kompetenzen

Im Gegensatz zu dem Anspruch, Aufgaben mit anspruchsvollen Denkprozessen zu stellen, scheint Unterricht in Deutschland häufig durch einen Fragen-entwickelnden Stil geprägt zu sein, in dem der Lerngegenstand kleinschrittig aufgearbeitet wird. Komplexe, verständnisorientierte Fragen werden deutlich seltener eingesetzt (u.a. Kleinknecht, 2019; Klieme & Schreyer, 2020; Neubrand et al., 2011). Dies könnte ein Hinweis darauf sein, dass Lehrkräfte die kognitiven Anforderungen in Aufgaben nicht angemessen beurteilen können. So zeigen bspw. Mathematiklehrkräfte bei der Beurteilung der Aufgabenschwierigkeit sowohl eine Tendenz zur Unterschätzung (u.a. Hosenfeld et al., 2002) als auch zur Überschätzung (Lintorf et al., 2011; McElvany et al. 2009). Zudem haben sie Schwierigkeiten, anforderungsgenerierende Merkmale von Aufgaben zu benennen und das kognitive Anforderungsniveau zu identifizieren (Scheja, 2019). Hinsichtlich des kognitiven Aktivierungspotentials von Aufgaben ist fachübergreifend zu vermuten, dass Lehrkräfte dies eher überschätzen (Hammer, 2015; Widmer, 2023).

Auch im Fach Sport zeigt sich eine Tendenz, dass Sportlehrkräfte eher Aufgaben stellen, die zu wenig anspruchsvollen Denkprozessen anregen (u.a. Bröcheler & Pfitzner, 2014; Hendricks, 2024; Pfitzner, 2018). Zudem stellen angehende Sportlehrkräfte in Praxisphasen häufiger Aufgaben, die Schüler*innen zur Reproduktion von Wissen auffordern und weniger zum eigenständigen Problemlösen, zu tiefergehenden Begründungen oder Argumentationen anregen (Heemsoth, 2024). Die Frage, inwiefern (angehende) Sportlehrkräfte Aufgaben beurteilen können und welche kognitiven Prozesse dabei ablaufen, wurde bisher nur vereinzelt thematisiert. In zwei Studien untersuchte Hendricks (2024; 2025), inwiefern es Sportlehrkräften (Hendricks, 2024) und Sportstudierenden (Hendricks, 2025) gelingt, in einem Vergleich zwischen zwei Aufgaben, die jeweils kognitiv und motorisch anspruchsvollere Aufgabe auszuwählen. Es zeigte sich, dass es sowohl für Sportlehrkräfte als auch für Sportstudierende herausfordernd ist, die Aufgaben mit einer höheren Ausprägung in den kognitiven und motorischen Anforderungen auszuwählen (Hendricks, 2024; 2025). In einem Vergleich zwischen Sportlehrkräften und Trainer*innen aus dem Turnen untersuchten Leukel et al. (2023) die Urteilsgenauigkeit und die kognitiven Prozesse, die der Urteilsbildung zugrunde liegen. Die Ergebnisse verdeutlichen, dass die Sportlehrkräfte im Gegensatz zu den Trainer*innen eine signifikant niedrigere Urteilsgenauigkeit aufwiesen. Zudem stuften die Sportlehrkräfte den Schweregrad der Bewegungsfehler als weniger schwerwiegend ein.

3 Forschungsfragen und -hypothesen der vorliegenden Untersuchung

Die vorliegenden Befunde könnten darauf hindeuten, dass angehende Sportlehrkräfte die Ausprägung kognitiver Anforderungen in Aufgaben nicht akkurat beurteilen können. Damit sich die Sportlehrkräftebildung am Lern- und Leistungsstand angehender Sportlehrkräfte orientieren kann, scheinen hierzu tiefergehende Erkenntnisse notwendig. Dieser Forderung wird mit der Beantwortung der folgenden Forschungsfragen nachgegangen:

- *Forschungsfrage 1: Inwiefern gelingt es Sportstudierenden, die Ausprägung kognitiver Anforderungen in Aufgaben akkurat zu beurteilen?*

Die Tendenz von angehenden Sportlehrkräften, Aufgaben zu stellen, die kaum zu anspruchsvollen Denkprozessen anregen (Heemsoth, 2024), könnte ein Hinweis darauf sein, dass sie die Ausprägung der kognitiven Anforderungen überschätzen. Dies würde mit Erkenntnissen anderer Fächer übereinstimmen (u.a. Hammer, 2015; Lintorf et al., 2011; McElvany et al. 2009; Widmer, 2023). Für die Beantwortung der ersten Forschungsfrage ergibt sich daraus die folgende Hypothese:

- *Hypothese zur Forschungsfrage 1: Sportstudierende überschätzen die Ausprägung der kognitiven Anforderungen in Aufgaben.*

Erkenntnisse über die Urteilsakkuratheit allein reichen nicht aus, um der oben formulierten Forderung nach einer am Lern- und Leistungsstand der Studierenden orientierten Sportlehrkräftebildung nachzukommen. Es sind darüber hinaus Informationen zu möglichen Unterschieden in den Wahrnehmungs-, Interpretations- und Entscheidungsprozesse erforderlich, um herauszufinden, wodurch eine akkurate Beurteilung bedingt wird. Vor dem Hintergrund der Notwendigkeit, lernförderliche Aufgaben zu stellen (vgl. Kapitel 2) und der gleichzeitigen Tendenz, eher wenig anspruchsvolle Denkprozesse anzuregen, werden Studierende mit einer akkuraten Beurteilung und mit einer Tendenz zur Überschätzung der kognitiven Anforderungen miteinander verglichen:

- *Forschungsfrage 2: Inwiefern unterscheiden sich die Wahrnehmungs-, Interpretations- und Entscheidungsprozesse während der Aufgabenanalyse bei Sportstudierenden mit einer akkuraten Beurteilung im Vergleich zu denjenigen mit einer Tendenz zur Überschätzung?*

4 Methodisches Vorgehen

Das methodische Vorgehen zur empirischen Untersuchung der Forschungsfragen folgt einem sequenziellen Vertiefungsdesign (Burzan, 2016) mit einer quantitativen Teilstudie (Fragestellung 1), auf die eine qualitative Teilstudie (Fragestellung 2) folgt. Beide Teilstudien wurden im Zeitraum von Oktober 2024 bis Februar 2025 durchgeführt. Die qualitative Befragung fand in einem zeitlichen Abstand von ca. zwei Monaten nach der quantitativen Studie statt.

4.1 Stichprobe

An der *quantitativen Teilstudie* nahmen $N = 323$ Sportlehramtsstudierende (55,5% männlich; 43% weiblich; 0,6% divers; 0,9% keine Angabe; $M_{\text{Alter}} = 22,3$ Jahre, $SD = 3,27$; 68,1% Bachelor; 31,9% Master) teil. Die Befragten wurden über universitäre Lehrveranstaltungen an drei universitären Standorten in NRW gewonnen. Sie wurden im Vorfeld über die Ziele und den Ablauf der Befragung aufgeklärt. Die Teilnahme war freiwillig und erfolgte anonym. Die Studierenden konnten sich freiwillig für eine mögliche Teilnahme an der qualitativen Teilstudie bereiterklären.

Für die qualitative Teilstudie konnten von denjenigen Studierenden, die sich für eine Teilnahme am Interview bereiterklärten, jeweils $n = 5$ Studierende mit einer realistischen Einschätzung (Niveauelemente zwischen $-0,1 - 0$) sowie einer Tendenz zur Überschätzung (Niveauelemente zwischen $0,7 - 1,1$) gewonnen werden³. Die »Realist*innen« repräsentieren Studierende mit akkuraten Beurteilungen der Anforderungen in der quantitativen Studie (siehe Beschreibung in Kapitel 4.3). Sie setzten sich aus je einer Bachelor- und Masterstudentin sowie zwei Bachelor- und einem Masterstudenten zusammen ($M_{\text{Anzahl Praktika}} = 1,2$, $SD = 1,79$, $\text{Min} = 0$, $\text{Max} = 4$; Tätigkeit als Vertretungslehrkraft: Ja = 2, Nein = 3). Die »Überschätzer*innen« repräsentieren Studierende mit einer Tendenz zur Überschätzung der Anforderungen in der quantitativen Studie. Sie setzten sich aus einer Bachelorstudentin und drei Masterstudentinnen sowie einem Masterstudenten zusammen ($M_{\text{Anzahl Praktika}} = 2$, $SD = 1,23$, $\text{Min} = 0$, $\text{Max} = 3$; Tätigkeit als Vertretungslehrkraft: Ja = 2, Nein = 3).

4.2 Datenerhebung

Quantitative Teilstudie

Es wurden Aufgaben für eine durchschnittliche lern- und leistungsstarke achte Klasse im Bewegungsfeld *Laufen, Springen, Werfen – Leichtathletik* (Weitsprung) unter der pädagogischen Perspektive *das Leisten erfahren, verstehen und einschätzen* in Form von Textvignetten entwickelt. Die Aufgaben unterscheiden sich hinsichtlich ihrer Ausprägung der kognitiven Anforderungen systematisch voneinander. Die Ausprägung der kognitiven Anforderungen wurde auf Grundlage der in Tab. 1 dargestellten Anforderungen abgeleitet und über Operatoren (z.B. nenne, diskutiere) auf der Sichtstruktur der Textvignetten abgebildet. Auf diese Weise ist eine klare Unterscheidung zwischen den kognitiven Anforderungen der Aufgaben möglich.

3 Es liegen keine Richtlinien zur Interpretation der statistischen Kennwerte der Niveauelemente der diagnostischen Kompetenzen und zu aufgabenanalytischen Kompetenzen vor (siehe u.a. Praetorius et al., 2011; Spinath, 2005). Für die Auswahl der Realist*innen wurden Studierende ausgewählt, die möglichst nah am Idealwert 0 liegen ($NK = -0,1 - 0,1$). Für die Überschätzer*innen wurden Studierende ausgewählt, die möglichst weit vom Idealwert 0 entfernt sind ($NK = 0,7 - 1,1$). Aufgrund der aus dem Forschungsstand abgeleiteten Hypothese, wonach eine Tendenz zur Überschätzung vorliegt, werden in dieser Studie die Realist*innen mit den Überschätzer*innen verglichen.

Abb. 2 zeigt zwei Beispielaufgaben mit unterschiedlich ausgeprägten kognitiven Anforderungen. Bei Aufgabe 1 müssen die Schüler*innen bekannte Bewegungsmerkmale nennen, d.h. Faktenwissen reproduzieren. Aufgabe 2 verlangt von den Schüler*innen Ursache-Wirkungs-Zusammenhänge aufzustellen, indem sie durch eigene Bewegungserfahrungen herausfinden, wie sie möglichst weit springen.

Aufgabe 1 (kognitive Anforderungsstufe 2)	Aufgabe 2 (kognitive Anforderungsstufe 4)
Wer kann mir sagen, wie die Technik aus der letzten Woche heißt? <i>Die Schüler*innen nennen die Schrittsprungtechnik.</i> Genau. Was ist wichtig, um damit weit zu springen? <i>Die Schüler*innen nennen Aspekte wie Schrittstellung; Nachziehen des Sprungbeins; Klappmesserbewegung. Die Lehrkraft schreibt die Begriffe auf ein Whiteboard.</i> Die Technik wollen wir jetzt anwenden, indem jede*r von euch mehrere Sprünge damit macht. Ihr springt abwechselnd mit eurer Partnerin bzw. eurem Partner aus der Absprungzone. Eure Partnerin bzw. euer Partner beobachtet euch und gibt euch Tipps. Wenn ihr noch nicht alle Merkmale auf einmal umsetzen könnt, konzentriert ihr euch auf einzelne Aspekte."	"Ihr sollt herausfinden, wie ihr weit springt. Jede*r von euch macht dafür drei Sprünge, in denen ihr versucht möglichst weit zu springen. Eure Partnerin bzw. euer Partner filmt euch dabei. Analysiert danach eure Videos. Wann seid ihr besonders weit gesprungen, wann nicht? Diskutiert, warum ihr weit/ nicht weit gesprungen seid und stellt Leitsätze für einen weiten Sprung auf. Beendet dafür den folgenden Satz: "Um weit zu springen,...". Führt danach jeweils drei Sprünge aus und versucht eure Leitsätze umzusetzen. Entscheidet selbstständig, ob ihr alle Leitsätze gleichzeitig, oder nacheinander umsetzt."

Abb. 2: Beispielaufgaben

Die Aufgaben wurden anschließend hinsichtlich der in der Aufgabenentwicklung theoretisch angenommenen Ausprägung der kognitiven Anforderungen sowie der inhaltlichen Validität mithilfe einer zweistufigen Expert*innenbefragung (je $n = 6$ Professor*innen, promovierte Sportdidaktiker*innen, Fachleiter*innen) geprüft und überarbeitet. Das Ziel der Befragung bestand darin, einen größtmöglichen Konsens zwischen den Expert*innen und der theoretisch angenommenen Norm zu erreichen. Auf diese Weise sollte ein Vergleichsmaßstab für die Beurteilungen der Studierenden abgeleitet werden. Aufgrund der ordinalen Skalierung der Anforderungsstufen wurde die Übereinstimmung über einen eindeutigen Modalwert und eine Abweichung von diesem um maximal eins operationalisiert. Aufgaben, für die nach der zweiten Expert*innenbefragung keine Übereinstimmung erzielt wurde, wurden im weiteren Verlauf der Studie ausgeschlossen (vgl. Hendricks, 2024). Durch dieses Vorgehen konnte ein Aufgabenpool mit insgesamt 21 Aufgaben entwickelt werden. Für die Befragung der Studierenden wurden zehn Aufgaben so ausgewählt, dass jede Anforderungsstufe zweimal auftrat. Die Studierenden wurden in einem querschnittlichen Design mithilfe des Online-Befragungstools SoSci-Survey befragt. Zum Einstieg wurden sie über die wesentlichen Kontextinformationen informiert (z.B. die fünfstufige Ausprägung der kognitiven Anforderungen (Tab. 1), Klassenstufe, Bewegungsfeld etc.). Anschließend wurden den Studierenden zehn Aufgaben in zufälliger Reihenfolge präsentiert und sie wurden gebeten, die Ausprägung der kognitiven Anforderungen auf der fünfstufigen Skala zu beurteilen.

Qualitative Teilstudie

Es wurden halbstrukturierte leitfadengestützte Einzelinterviews (Misoch, 2015) mit integrierter Methode des lauten Denkens (angelehnt an Buber, 2009; Fischer, 2009) durchgeführt. Das Interview unterteilte sich in vier Phasen, die auf die Anwendung der Methode des lauten Denkens ausgerichtet waren. Die *Einstiegsphase* diente dazu, die Befragten gedanklich in das Thema der Aufgaben im Sportunterricht einzuführen. Hierbei berichteten die Befragten, worin sich Aufgaben im Sportunterricht von ihrem Zweitfach unterscheiden und was Aufgaben im Sportunterricht aus ihrer Sicht anspruchsvoll macht. Darauf aufbauend wurde in der *Erprobungsphase* die Methode des lauten Denkens eingeführt. Hierzu wurde den Studierenden zunächst das Vorgehen der Methode, die fünfstufige Ausprägung der kognitiven Anforderungen sowie wesentliche Kontextinformationen (Klassenstufe etc.) erläutert. Anschließend erprobten die Studierenden das Vorgehen, indem sie die Ausprägung der kognitiven Anforderungen in zwei Probeaufgaben beurteilten und währenddessen ihre Gedanken laut aussprachen. Das Ziel der Methode des lauten Denkens liegt allgemein darin, Denkstrukturen von Befragten durch verbale Auskünfte zu identifizieren (Buber, 2009). Es wird davon ausgegangen, dass durch sprachliche Äußerungen ein Zugang zu den während des Handlungsvollzugs im Fokus der Aufmerksamkeit stehenden Kognitionen eröffnet werden kann (Fischer, 2009)⁴. In der vorliegenden Studie wurde eine unstrukturierte und zeitgleiche Variante des lauten Denkens gewählt (Buber, 2009). Die Analyse der Aufgaben fand reflektiert und ohne Zeitdruck statt. In der *Hauptphase* beurteilten die Studierenden nacheinander die Ausprägung der kognitiven Anforderungen von drei Aufgaben und verbalisierten zeitgleich ihre Gedanken. Der Interviewer hielt sich währenddessen zurück und erinnerte gegebenenfalls daran, dass alle Gedanken laut ausgesprochen werden sollten. Auf diese Weise sollte sichergestellt werden, dass die Befragten während des Interviews konsistent laut gedacht haben. Abschließend erfolgte eine *Ausstiegsreflexion*, in der die Interviewsituation reflektiert wurde. Hierbei ging es um eine Selbsteinschätzung, inwiefern den Befragten die Methode des lauten Denkens schwergefallen war und ob sie mit der Aufgabenanalyse zurechtgekommen waren. Die Interviews wurden über die Videotelefonie-Software Zoom durchgeführt und dauerten im Durchschnitt 32 Minuten (Min = 24:54 Minuten; Max = 42:52 Minuten).

4.3 Datenaufbereitung und -analyse

Quantitative Teilstudie

Zur Beantwortung der ersten Forschungsfrage wurde die Niveauelemente nach Schrader und Helmke (1987) mithilfe der Programme Microsoft Excel und SPSS 27 berechnet. Die Niveauelemente entsprechen der Differenz zwischen dem mittleren

4 Eine kritische Betrachtung dieser Annahme (siehe u.a. Neuweg, 2015) erfolgt in der Diskussion.

Studierendenurteil und dem durch die Expert*innebefragung ermittelten Vergleichsmaßstab. Bei einer akkuraten Einschätzung der Studierenden im Vergleich zu den Expert*innen nimmt die Niveauekomponente den Idealwert 0 an. Werte größer 0 weisen auf eine Tendenz zur Überschätzung und Werte kleiner 0 auf eine Unterschätzung hin. Mittels Ein-Stichproben-t-Test wurde anschließend geprüft, inwiefern der Mittelwert der Beurteilungen der Studierenden vom Idealwert 0 verschieden ist (Praetorius et al., 2011). Für die Interpretation der Ergebnisse wurde ein Signifikanzniveau von $\alpha = 0,05$ festgelegt (Bortz, 2005).

Qualitative Teilstudie

Für die Beantwortung der zweiten Forschungsfrage wurden die Interviews nach der Befragung vollständig schriftlich transkribiert. Die Auswertung erfolgte anhand der inhaltlich strukturierenden qualitativen Inhaltsanalyse (Kuckartz, 2014) mithilfe der Auswertungssoftware MAXQDA 2022. In einem ersten Schritt wurden die transkribierten Interviews vollständig durchgearbeitet und entlang der aus dem adaptierten KompeSport-Modell (Abb. 1) deduktiv abgeleiteten Hauptkategorien *Wahrnehmung*, *Interpretation* und *Entscheidung* kodiert. Daraufhin wurden alle kodierten Textstellen der gleichen Hauptkategorie zusammengestellt und Subkategorien induktiv am Material für jede Hauptkategorie bestimmt (Tab. 2).

Tab. 2: Übersicht Haupt- und Subkategorien

Haupt- und Subkategorie	Erläuterung
Wahrnehmung	Aussagen der Studierenden, ...
Wahrnehmung von Operatoren	... bei denen Operatoren genannt werden, die im Aufgabentext enthalten sind.
Wahrnehmung von kognitiven Aktivitäten	... bei denen kognitive Aktivitäten genannt oder beschrieben werden, die durch die Aufgabenbearbeitung angeregt werden sollen.
Interpretation	Aussagen der Studierenden, ...
Rückgriff auf Fach- und fachdidaktisches Wissen	... bei denen in der Interpretation Fach- und fachdidaktisches Wissen hinzugezogen wird, das nicht aus dem Aufgabentext ersichtlich ist.
Zielgruppenberücksichtigung	... bei denen in der Interpretation der Lern- und Entwicklungsstand oder (vermutetes) Vorwissen der Schüler*innen hinzugezogen wird.
Entscheidung	Aussagen der Studierenden, ...
Einfache Begründung	... in denen einzelne inhaltlich entscheidungsrelevante Informationen genannt werden und darauf aufbauend eine Entscheidung getroffen wird.

Umfangreiche Begründung	... in denen inhaltlich entscheidungsrelevante Informationen und ihre Bedeutung für die Ausprägung der Anforderung diskutiert werden und darauf aufbauend eine Entscheidung getroffen wird.
-------------------------	---

Danach wurden alle Interviews mit dem erweiterten Kategoriensystem erneut durch eine Person kodiert. Nach der Kodierung erfolgte eine fallbezogene thematische Zusammenfassung entlang der Subkategorien. Diese wurde für die Studierenden mit einer akkuraten Beurteilung und einer Tendenz zur Überschätzung getrennt voneinander durchgeführt. Das Ziel lag darin, die Ergebnisse beider Gruppen (Realist*innen und Überschätzer*innen) am Ende der Auswertung zu vergleichen und Gemeinsamkeiten und Unterschiede herauszustellen.

5 Ergebnisse

Quantitative Teilstudie

Die Analyse zur ersten Forschungsfrage, inwiefern es angehenden Sportlehrkräften gelingt, die Ausprägung kognitiver Anforderungen in Aufgaben akkurat zu beurteilen, zeigt, dass die Niveauelemente leicht positiv ausfällt ($M = 0,09$; $SD = 0,46$; $Min = -1,2$; $Max = 1,5$). Die Analyse mittels Ein-Stichproben-t-Test ergab, dass der Mittelwert signifikant vom Idealwert 0 verschieden ist ($t(323) = 3,37$; $p < 0,001$). Die Effektstärke nach Cohen (1992) liegt bei $d = 0,19$, womit ein sehr kleiner Effekt vorliegt. Die Ergebnisse zeigen somit, dass die Studierenden die Ausprägung kognitiver Anforderungen mit einem kleinen Effekt im Mittel leicht überschätzten.

Qualitative Teilstudie

Die Analyse zur zweiten Forschungsfrage, inwiefern sich die Informationsverarbeitungsprozesse während der Aufgabenanalyse bei Studierenden unterscheiden, zeigt für die Wahrnehmung, dass die *Realist*innen* während der Aufgabenanalyse bei jeder Aufgabe die Operatoren (z.B. nennen, begründen, entwickeln) nennen konnten:

*»Also wenn man sich die Operatoren anguckt. Das noch mal diskutiert wird«
(R. 378, Pos. 16).*

Zudem beschrieben sie sowohl eher niedrigere (z.B. Reproduzieren von Fachwissen) als auch höhere kognitive Prozesse (z.B. Problemlösen), die zur Bearbeitung der Aufgabe notwendig sind. In der Interpretation nahmen sie auf ihr eigenes Fachwissen Bezug. Dies wird daran deutlich, dass sie bspw. die Schrittsprungtechnik als Fachsprache bezeichneten oder das Konzept der kognitiven Aktivierung explizit anführten:

»Wenn man sich den ersten Teil anguckt, dann denke ich, dass hier vor allem auch eine kognitive Aktivierung stattgefunden hat, in dem Sinne, dass wie das eben schon besprochen wurde, es eine Wenn-Dann Beziehung gibt. Also was

muss ich tun, um meinen Körper auf speziell den Weitsprung vorzubereiten, auch die Übungen durchzuführen (...)« (R. 378, Pos.16).

Darüber hinaus erläuterten sie Unterschiede in der Erwärmung zwischen dem Weitsprung und anderen Bewegungsfeldern und die daraus resultierende Konsequenz, dass Schüler*innen verschiedene Wissensbestandteile miteinander verknüpfen müssen:

»Und da dafür ist hier auch auf jeden Fall schon ein bisschen Fachwissen gefragt, weil man sich halt nicht generell aufwärmen soll, sondern explizit für den Weitsprung. Da ist es wichtig, auf jeden Fall zu wissen, welche Muskeln ich überhaupt besonders aufwärmen möchte. Also sprich jetzt nicht beim Weitsprung irgendwie stark, wie zum Beispiel beim Turnen die Handgelenke warm zu machen. Und das sollten die Schüler auf jeden Fall mit einbeziehen« (R. 166, Pos. 31).

Zudem wurde in der Interpretation ein Zusammenhang zwischen der Anforderung des Reflektierens als kognitive Aktivität und den Voraussetzungen der Schüler*innen hergestellt. Hierbei wurde einerseits auf die Notwendigkeit eingegangen, dass die Schüler*innen Vorwissen benötigen und andererseits wurde ein Bezug zu ihrem Alter hergestellt. Im Entscheidungsprozess finden sich Aussagen, die sich sowohl den einfachen (3x) als auch den umfangreichen (16x) Begründungen zuordnen lassen. Bei einfachen Begründungen wurden Operatoren oder kognitive Aktivitäten als entscheidungsrelevante Informationen genannt, um darüber die Einschätzung zur Ausprägung der kognitiven Anforderung zu begründen. Im Gegensatz dazu wurden in umfangreichen Begründungen verschiedene inhaltliche entscheidungsrelevante Informationen und ihre Bedeutung für die Ausprägung der Anforderung berücksichtigt:

»Anschließend geht es um die Anwendung (.) noch mal bezogen auf das Kognitive würde ich nicht unbedingt sagen, dass hier eine Erhöhung der Kognition stattfindet. Allein deswegen, weil kein weiteres Wissen abgefragt wird und beispielsweise jetzt keine Bewegung weiterentwickelt werden muss, die in keinen Wirkungszusammenhang gebracht werden muss« (R. 378, Pos. 18).

Die Analyse für die Überschätzer*innen zeigt für die Wahrnehmung, dass sie die Operatoren der Aufgabe (z.B. nennen, diskutieren, begründen) ebenfalls nennen konnten und die zur Bearbeitung der Aufgabe notwendigen kognitive Prozesse beschrieben. In der Interpretation wurde lediglich von einer Person explizit auf Fach- und fachdidaktisches Wissen Bezug genommen. Die Zielgruppe wurde in der Interpretation nur vereinzelt berücksichtigt, indem zur Aufgabenbearbeitung notwendiges Vorwissen angeführt wurde (3x). Im Entscheidungsprozess finden sich auch bei den Überschätzer*innen Aussagen, die sich sowohl den einfachen (13x) als auch den umfangreichen Begründungen (4x) zuordnen lassen.

»Ja. Also es ist Erfahrungswissen, weil die ja irgendwie was wissen müssen. Und dann haben sie so Vorher-Nachher Vergleiche. Welche Stufe ist das dann? Eins, zwei?« (Ü. 448, Pos. 29).

Dieses Beispiel verdeutlicht exemplarisch, dass die Überschätzer*innen vermehrt einzelne entscheidungsrelevante Informationen als Begründung für ihre Entscheidung herangezogen haben, ohne einen expliziten inhaltlichen Bezug zum Weitsprung herzustellen. Umfangreiche Begründungen, bei denen verschiedene inhaltlich entscheidungsrelevante Informationen und ihre Bedeutung für die Ausprägung der Anforderung berücksichtigt werden, fanden selten statt.

6 Diskussion und Ausblick

Die Studie bietet Einblicke darin, ob angehende Sportlehrkräfte die Ausprägung kognitiver Anforderungen akkurat beurteilen können und inwiefern sich informationsverarbeitende Prozesse bei Studierenden mit unterschiedlich akkuraten Beurteilungen voneinander unterscheiden.

Die Ergebnisse der quantitativen Teilstudie weisen darauf hin, dass die Studierenden die Ausprägung der kognitiven Anforderungen leicht überschätzten. Die für die erste Forschungsfrage formulierte Hypothese konnte somit bestätigt werden. Die Ergebnisse stimmen zudem mit Studien aus der Mathematik (u.a. Hammer, 2015) oder dem Deutschunterricht (u.a. Widmer, 2023) überein und könnten ein Hinweis darauf sein, dass angehende Sportlehrkräfte tendenziell Aufgaben mit einer geringen kognitiven Komplexität auswählen (Heemsoth, 2024). Dies kann für die Gestaltung eines lernförderlichen Sportunterrichts problematisch sein, da eine am Lern- und Leistungsstand der Schüler*innen angepasste Auswahl der Aufgaben sowie eine sukzessive Erhöhung der kognitiven Anforderungen sich positiv auf deren Lernleistungen auswirken können (Wank et al., 2019). Für die Interpretation muss allerdings die sehr kleine Effektstärke ($d = 0,19$) und der mögliche Extremwert der Überschätzung von maximal zwei berücksichtigt werden. Auch wenn die Effektstärke nah an einem kleinen Effekt ($d = 0,2$; Cohen, 1992) liegt, sollte die im Mittel leichte Tendenz zur Überschätzung der Ausprägung kognitiver Anforderungen ($M = 0,09$) nicht überbewertet werden. Es scheint eher so, dass angehende Sportlehrkräfte die Ausprägung der kognitiven Anforderungen im Mittel akkurat einschätzten. Das Ergebnis wird vermutlich dadurch bedingt, dass die Operatoren auf der Sichtstruktur des Aufgabentextes liegen. Kognitiv wenig anspruchsvolle Aufgaben (bspw. Faktenwissen nennen) und anspruchsvolle Aufgaben (bspw. einen Sachverhalt diskutieren) sind dadurch vermutlich leicht zu identifizieren. Diese Annahme wird durch die qualitative Teilstudie bestärkt. In beiden Gruppen zeigte sich, dass die Studierenden die Operatoren sowie die mit der Aufgabe aufseiten der Schüler*innen angestrebten kognitiven Aktivitäten wahrnehmen konnten. Hierbei unterschieden sich die beiden Gruppen nicht voneinander.

In der Interpretation und im Entscheidungsprozess konnten Unterschiede zwischen den Realist*innen und Überschätzer*innen identifiziert werden. In der Interpretation griffen die Realist*innen vermehrt explizit auf ihr Fach- und fachdidaktisches

Wissen zurück. So wurde bspw. die kognitive Aktivierung als theoretisches Konstrukt herangezogen. Bei den Überschätzer*innen nahm lediglich eine Person explizit auf ihr Fachwissen Bezug, indem die unterschiedlichen Phasen des Weitsprungs angeführt wurden. Ein weiterer Unterschied in der Interpretation liegt im Zielgruppenbezug. Die Realist*innen bezogen in der Interpretation häufiger das notwendige Vorwissen der Schüler*innen mit ein und es wurde zumindest von einer Person der Entwicklungsstand von Schüler*innen einer achten Klasse berücksichtigt. Insgesamt wurde die Zielgruppe in beiden Gruppen jedoch recht selten angeführt, was aus der kleinen Stichprobe resultieren könnte. Auch wenn die Mehrzahl der befragten Studierenden durch Praktika oder Vertretungstätigkeiten bereits Praxiserfahrung sammeln konnten, ist fraglich, ob dies ausreicht, um eine »durchschnittlich lern- und leistungsstarke« achte Klasse beurteilen zu können. Dies könnte ein weiterer Grund sein, warum die Zielgruppe in der Interpretation selten berücksichtigt wurde.

Im Entscheidungsprozess zeigte sich, dass die Realist*innen im Vergleich zu den Überschätzer*innen umfangreichere entscheidungsrelevante Informationen nutzten und deren inhaltliche Bedeutung für die Ausprägung der Anforderung stärker diskutierten. Im Gegensatz dazu nutzten die Überschätzer*innen vermehrt Informationen auf der Sichtstruktur des Aufgabentextes, ohne deren inhaltliche Bedeutung zu diskutieren. Obwohl kognitive Aktivierung in der Unterrichtsqualitätsforschung als fachübergreifendes Konstrukt diskutiert wird (zsf. Wemmer-Roth et al., 2024), sollten fachliche Lerngegenstände als Bezugspunkt zur unterrichtlichen Umsetzung dienen (zsf. Keller et al. 2024). Für die sportunterrichtliche Umsetzung kognitiv anspruchsvoller Aufgaben, mit denen der Lerngegenstand am Lern- und Leistungsniveau der Schüler*innen orientiert werden kann, gilt es daher, zentrale Konzepte oder Prinzipien eines sport- und bewegungsbezogenen Inhalts (in dieser Studie Laufen, Springen, Werfen - Weitsprung) bei der kognitiven Auseinandersetzung zu berücksichtigen (Engelhardt et al., 2024). Andernfalls verbleibt die Auseinandersetzung mit dem Lerngegenstand vermutlich auf einer oberflächlichen Ebene, ohne anspruchsvolle lerngegenstandsbezogene Denkprozesse auszulösen. Für die Beurteilung der Ausprägung kognitiver Anforderungen könnte daraus folgen, dass (angehende) Sportlehrkräfte sport- und bewegungsbezogene inhaltliche Aspekte in der Interpretation und Entscheidungsfindung berücksichtigen müssen. Die Ergebnisse deuten an, dass die Realist*innen dies stärker umsetzten, was ein Grund für die akkurateren Beurteilungen sein könnte. Die Nutzung von fachspezifischen inhaltlichen Aspekten scheint für die akkurate Beurteilung der Ausprägung kognitiver Anforderungen in Aufgaben somit relevant zu sein. Ergebnisse von Studien aus der Mathematikdidaktik bestärken diese Annahme und legen nahe, dass sich die Vermittlung von fachdidaktischem Wissen über anforderungsgenerierende Merkmale in Aufgaben positiv auf die Beurteilung auswirken kann (Rieu et al., 2020; Scheja, 2019). Für die universitäre Sportlehrkräftebildung könnte es demnach einen Mehrwert haben, die Bedeutung der Ausprägung kognitiver Anforderungen in universitären Lehrveranstaltungen stärker zu thematisieren. Ob dies durch die Vermitt-

lung von aufgabenbezogenem Wissen (u.a. Rieu et al., 2020), durch professionelle Reflexionen mit einem expliziten Selbstbezug hinsichtlich der eigenen Erfahrungen im Stellen von Aufgaben (vgl. Ausführungen zum Reflexionsbegriff bei Lenske & Lohse-Bossenz, 2023) oder durch eine Kombination aus beidem erfolgen kann, ist eine offene Frage. Ungeklärt ist zudem, ob aufgabenanalytische Kompetenzen in theoretischen oder fachpraktischen Lehrveranstaltungen adressiert werden sollten.

Die Ergebnisse müssen vor dem Hintergrund der eingesetzten Methode kritisch interpretiert werden. Die Methode des lauten Denkens basiert auf der Annahme, dass während der Aufgabenanalyse explizierbare kognitive Prozesse ablaufen (Buber, 2009; Fischer, 2009). Folgt man dieser Annahme, so konnten mit der Studie Unterschiede hinsichtlich der Nutzung von Fach- und fachdidaktischem Wissen in der Interpretation und Entscheidung zwischen beiden Gruppen aufgezeigt werden. Dem entgegen könnte argumentiert werden, dass die Überschätzer*innen ihr fach- und fachdidaktisches Wissen nicht verbalisierten, aber dennoch implizit nutzten (u.a. Neuweg, 2015). Professionelles Wissen muss in beruflichen Anforderungssituationen nicht bewusst explizierbar sein. Dies wäre bspw. der Fall, wenn die Informationsverarbeitung in Situationen mit Zeitdruck intuitiv erfolgt. Da mit der Interviewsituation jedoch eine Situation ohne Zeitdruck simuliert wurde, kann davon ausgegangen werden, dass die Informationsverarbeitung reflektiert, d.h. unter Rückgriff auf explizierbares Wissen stattfand. Einschränkend muss berücksichtigt werden, dass in dieser Studie kein Wissenstest zur Erhebung des Fach- oder fachdidaktischen Wissens eingesetzt wurde. Dem KompeSport-Modell folgend (Abb. 1; Baumgartner, 2022), hat das professionelle Wissen über Aufgaben einen Einfluss auf die Informationsverarbeitungsprozesse. Die Unterschiede zwischen den Realist*innen und den Überschätzer*innen könnten daher auch aus einem unterschiedlich ausgeprägten professionellen Wissen über Aufgaben resultieren.

In diesem Zusammenhang stellt sich zudem die Frage, ob die Befragten während des Interviews ihre Gedanken konsistent laut ausgesprochen oder stattdessen phasenweise ihr Vorgehen bei der Aufgabenanalyse erklärt haben. Letzteres würde bedeuten, dass keine Informationsverarbeitungsprozesse der Aufgabenanalyse erfasst wurden (Buber, 2009). Um sicherzustellen, dass die Befragten ihre Gedanken konsistent laut ausgesprochen haben, wurden sie von dem Interviewer situativ darauf hingewiesen, ihre Gedanken laut auszusprechen (z.B. wenn die Befragten aufhörten, ihre Gedanken laut auszusprechen oder begannen, ihr gedankliches Vorgehen zu erklären). Auf diese Weise sollte der Prozess des lauten Denkens wieder angeregt werden (siehe Vorgehen bei Buber, 2009).

Aufgrund der kleinen Stichprobe gibt die qualitative Teilstudie nur einen ersten Einblick in mögliche Unterschiede zwischen den Realist*innen und Überschätzer*innen in einer Disziplin des ausgewählten Bewegungsfelds Laufen, Springen, Werfen - Leichtathletik. Inwiefern die Ergebnisse auch auf andere Bewegungsfelder

oder leichtathletische Disziplinen übertragen werden können, ist bislang ungeklärt. Zudem stehen Analysen zum Zusammenhang zwischen der Studienerfahrung und der Akkuratheit der Beurteilungen ebenso noch aus, wie die Auswertung der Beurteilung motorischer Anforderungen. Mit dieser Studie wurde somit nur ein Teilausschnitt der Beurteilung kognitiv aktivierender Aufgaben betrachtet.

Die für das Forschungsvorhaben vorgenommene analytische Trennung zwischen kognitiven und motorischen Anforderungen muss mit Blick auf eine motorisch-kognitive Aktivierung (Herrmann & Gerlach, 2020) kritisch hinterfragt werden. Unter einer ganzheitlichen Perspektive sollte die Analyse von kognitiven Anforderungen nicht losgelöst von motorischen Anforderungen erfolgen, da wissens- und verständnisbezogene kognitive Lernprozesse zusammen mit körperlichen-motorischen Lernprozessen angebahnt werden. In weiteren Analysen sollte daher auch die Beurteilung der Ausprägung der motorischen Anforderungen und ihr Zusammenhang zur Beurteilung kognitiver Anforderungen untersucht werden. Auf diese Weise kann das Zusammenspiel zwischen der Beurteilung kognitiver und motorischer Anforderungen für die gezielte Aktivierung einer bewegungsbezogenen Auseinandersetzung der Schüler*innen mit dem Lerngegenstand stärker berücksichtigt werden.

Eine weitere Frage liegt zudem darin, inwiefern (angehende) Sportlehrkräfte mit unterschiedlich akkuraten Beurteilungen im Unterricht kognitiv aktivierende Aufgaben stellen, die dem Lern- und Leistungsniveau der Schüler*innen angemessen sind und inwiefern dies mit der Ausprägung des Fach- und fachdidaktischen Wissens zusammenhängt. Die vorliegenden Ergebnisse geben diesbezüglich Anlass für eine tiefergehende Auseinandersetzung mit den aufgabenanalytischen Kompetenzen angehender Sportlehrkräfte.

Literatur

- Anderson, L. W., & Krathwohl, D. R. (2001). *A Taxonomy for Learning, Teaching, and Assessing. A Revision of Bloom's Taxonomy of Educational Objectives*. Pearson Education.
- Aschebrock, H., Edler-Koeller, M., & Maass, P. (2010). Lernaufgaben im Fach Sport - Wege einer kompetenzorientierten Unterrichtsentwicklung. *Lernhilfen für den Sportunterricht*, 59(3), 13-16.
- Baumgartner, M. (2022). Professionelle Kompetenz(en) von Sportlehrkräften – Begriffe, Traditionen, Modellierungen und Perspektiven. In R. Sygusch, J. Hapke, S. Liebl, & C. Töpfer (Hrsg.), *Kompetenzorientierung im Sport. Grundlagen, Modellentwurf und Anwendungsbeispiele* (S. 35-50). Hofmann.
- Bernholt, S., Parchmann, I., & Commons, M.L. (2009). Kompetenzmodellierung zwischen Forschung und Unterrichtspraxis. *Zeitschrift für Didaktik der Naturwissenschaften*, 15, 219-245.
- Betsch, T., Funke, J., & Plessner, H. (2011). *Denken – Urteilen Entscheiden Problemlösen*. Springer.
- Blömeke, S., Gustafsson, J. E., & Shavelson, R. J. (2015). Beyond Dichotomies Competence Viewed as a Continuum. *Zeitschrift für Psychologie*, 223(1), 3–13.
- Bortz, J. (2005). *Statistik für Human- und Sozialwissenschaftler*. Springer.

- Bröcheler, R., & Pfitzner, M. (2014). Sportlehrkräfte und ihre Aufgaben: Eine qualitative Studie zu den Intentionen im Sportunterricht eingesetzter Aufgaben. In M. Pfitzner (Hrsg.), *Aufgabenkultur im Sportunterricht. Konzepte und Befunde zur Methodendiskussion für eine neue Lernkultur* (S. 295-322). Springer VS.
- Buber, R. (2009). Denke-Laut-Protokolle. In R. Buber, & H.H. Holzmüller (Hrsg.), *Qualitative Marktforschung* (S. 555- 568). Gabler.
- Burzan, N. (2016). *Methodenplurale Forschung - Chancen und Probleme von Mixed Methods*. Beltz Juventa.
- Cohen, J. (1992). A Power Primer. *Psychological Bulletin*, 112(1), 155-159.
- Engelhardt, S., Hapke, J., & Töpfer, C. (2023). Kognitive Aktivierung im Sportunterricht: ein Scoping Review zur Umsetzung methodisch-didaktischer Merkmale und zugrundeliegender konzeptioneller Grundannahmen. *Unterrichtswissenschaft*, 52(3), 293-324.
- Engelhardt, S., Hapke, J., & Töpfer, C. (2024). Steht nicht überall drauf, wo sie auch drin ist!« – Kognitive Aktivierung in internationalen sportdidaktischen Konzepten. In J. Wibowo, C. Krieger, E. Gerlach, & F. Bükers (Hrsg.), *Kognitive Aktivierung im Sportunterricht (3. vollständig überarbeitete Neuauflage (13-32)*. Universität Hamburg.
- Fauth, B. & Leuders, T. (2022). *Kognitive Aktivierung im Unterricht: Wirksamer Unterricht Band 2*. Institut für Bildungsanalyse Baden-Württemberg.
- Fischer, B. (2009). *Die pädagogische Perspektive "Das Leisten erfahren, verstehen und einschätzen" im Sportunterricht der gymnasialen Oberstufe in NRW. Eine Untersuchung zur unterrichtlichen Realisierung von curricularen Ansprüchen*. Dissertation, Technische Universität Dortmund.
- Frahm, H. (2022). Entwicklung eines fachbezogenen Kategoriensystems zur Analyse von Aktivierungspotenzialen von Aufgaben im Sportunterricht. In R. Sygusch, J. Hapke, S. Liebl, & C. Töpfer (Hrsg.), *Kompetenzorientierung im Sport. Grundlagen, Modellentwurf und Anwendungsbeispiele* (S. 87-100). Hofmann.
- Gogoll, A. (2014). Das Modell der sport- und bewegungskulturellen Kompetenz und seine Implikation für die Aufgabenkultur im Sportunterricht. In M. Pfitzner (Hrsg.), *Aufgabenkultur im Sportunterricht. Konzepte und Befunde zur Methodendiskussion für eine neue Lernkultur* (S. 93-110). Springer VS.
- Hammer, S. (2015). *Professionelle Kompetenz von Mathematiklehrkräften im Umgang mit Aufgaben in der Unterrichtsplanung Theoretische Grundlegung und empirische Untersuchung*. Dissertation, Ludwig-Maximilians-Universität München.
- Hapke, J., & Cramer, C. (2020). Professionalität und Professionalisierung in der Sportlehrerbildung. Potenzial des Ansatzes der Meta-Reflexivität. *Zeitschrift für sportpädagogische Forschung*, 8(2), 39-58.
- Hartmann, M., Stabick, O., & Bähr, I. (2023). Fruchtbare Momente durch ästhetische Aktivierung? Perspektiven von Sportlehrkräften. In E. Balz, & T. Bindel (Hrsg.), *Bildungszugänge im Sport: Grundlagen und Offerten* (S. 85-95). Springer VS.
- Heemsoth, T. (2024). Unterrichtliches Handeln angehender Sportlehrkräfte zur Aktivierung von Lernenden im Sportunterricht. *Zeitschrift für Erziehungswissenschaft*, 72, 1613-1638.
- Hendricks, C.P. (2024). *Aufgabenanalysefähigkeit als Teil der professionellen Handlungskompetenz von Sportlehrkräften – Theoretische Grundlegung und empirische Zugänge*. Springer VS.
- Hendricks, P. (2025). Aufgabenanalysefähigkeit von angehenden Sportlehrkräften – Eine quantitative Studie zu einem bisher vernachlässigten Thema in der Sportlehrkräftebildungsforschung. *Zeitschrift für sportpädagogische Forschung*, 13(2), 113-134.
- Herrmann, C., & Gerlach, E. (2020). Unterrichtsqualität im Fach Sport - Ein Überblicksbeitrag zum Forschungsstand in Theorie und Empirie. *Unterrichtswissenschaft*, 45(3), 361-384.
- Hosenfeld, I., Helmke, A., & Schrader, F.-W. (2002). Diagnostische Kompetenz. Unterrichts- und lernrelevante Schülermerkmale und deren Einschätzung durch Lehrkräfte in der Unterrichtsstudie SALVE. In M. Prenzel & J. Doll (Hrsg.), *Bildungsqualität von Schule: Schulische und außerschulische Bedingungen mathematischer, naturwissenschaftlicher und überfachlicher Kompetenzen* (S. 65-82). Beltz.
- Kahneman, D. (2003). A perspective on judgment and choice: Mapping bounded rationality. *American Psychologist*, 58(9), 697-720.

- Keller, S. D., Steffensky, M., Winkler, I., Lindmeier, A., Bertram, Ch., Herrmann, Ch., Schreyer, P., Fauth, B., & Praetorius, A.-K. (2024). Kognitive Aktivierung in den Fachdidaktiken: Ein fachübergreifendes fachspezifisches und lerngegenstandsorientiertes Konstrukt. In A.-K. Praetorius, W. Wemmer-Rogh, P. Schreyer, & M. Brinkmann (Hrsg.), *Kognitive Aktivierung unter der Lupe. Bestandsaufnahme und Möglichkeiten der Weiterentwicklung eines prominenten Konstrukts* (S. 233-247). Waxmann.
- Kleinknecht, M. (2019). Aufgaben und Aufgabenkultur. *Zeitschrift für Grundschulforschung*, 12(7), 1-14.
- Klieme, E. & Schreyer, P. (2020). Unterrichtsgestaltung und Unterrichtsqualität. In J. Grünkorn, E. Klieme, A.-K. Praetorius, & P. Schreyer (Hrsg.), *Mathematikunterricht im internationalen Vergleich: Ergebnisse aus der TALIS-Videostudie Deutschland* (S. 13-30). Leibniz-Institut für Bildungsforschung und Bildungsinformation.
- Krauss, S., Kunter, M., Brunner, M., Baumert, J., Blum, W., Neubrand, M., Jordan, A., & Löwen, K. (2004). COACTIV: Professionswissen von Lehrkräften, kognitiv aktivierenden Mathematikunterricht und die Entwicklung von mathematischer Kompetenz. In J. Doll & M. Prenzel (Hrsg.), *Bildungsqualität von Schule. Lehrerprofessionalisierung, Unterrichtsentwicklung und Schülerförderung als Strategien der Qualitätsverbesserung* (S. 31-53). Waxmann.
- Kuckartz, U. (2014). *Qualitative Inhaltsanalyse. Methoden, Praxis, Computerunterstützung*. Beltz Juventa.
- Langer, A. & Hendricks, P. (under review). Diagnostische Kompetenzen aus sportpädagogischer Perspektive in Theorie und Empirie. *Zeitschrift für sportpädagogische Forschung* (i.d.B.).
- Lenske, G., & Lohse-Bossenz, H. (2023). Stichwort: Professionelle Reflexionen im Pädagogischen Kontext. *Zeitschrift für Erziehungswissenschaft*, 26, 1133-1164.
- Leukel, C., Leuders, T., Bessi, F., & Loibl, K. (2023). Unveiling cognitive aspects and accuracy of diagnostic judgments in physical education teachers assessing different tasks in gymnastics. *Front. Educ.*, 8(1162499), 1-14.
- Lintorf, K., McElvany, N., Rjosk, C., Schroeder, S., Baumert, J., Schnotz, W., Horz, H., & Ullrich, M. (2011). Zuverlässigkeit von diagnostischen Lehrerurteilen – Reliabilität verschiedener Urteilsmaße bei der Einschätzung von Aufgabenschwierigkeiten. *Unterrichtswissenschaft*, 39(2), 102-120.
- Lipowsky, F. (2020). Unterricht. In E. Wild, & J. Möller (Hrsg.), *Pädagogische Psychologie*. Springer.
- 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.
- McElvany, N., Schroeder, S., Hachfeld, A., Baumert, J., Richter, T., Schnotz, W., Horz, H., & Ullrich, M. (2009). Diagnostische Fähigkeiten von Lehrkräften bei der Einschätzung von Schülerleistungen und Aufgabenschwierigkeiten bei Lernmedien mit instruktionalen Bildern. *Zeitschrift für Pädagogische Psychologie*, 23(34), 223-235.
- Misoch, S. (2015). *Qualitative Interviews*. de Gruyter Oldenbourg.
- Neuber, N. (2020). *Fachdidaktische Konzepte Sport. Zielgruppen und Voraussetzungen*. Springer VS.
- Neubrand, M., Jordan, A., Krauss, S., Blum, W., & Löwen, K. (2011). Aufgaben im COACTIV-Projekt: Einblicke in das Potenzial für kognitive Aktivierung im Mathematikunterricht. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss, & M. Neubrand (Hrsg.), *Professionelle Kompetenzen von Lehrkräften: Ergebnisse des Forschungsprogramms COACTIV* (S. 115-132). Waxmann.
- Neumaier, A. (2009). *Koordinatives Anforderungsprofil und Koordinationstraining* (3. Überarbeitete Auflage). Sportverlag Strauß.
- Neumann, P. (2014). Aufgabenanalyse im Sportunterricht – eine fachdidaktische Annäherung. In M. Pfitzner (Hrsg.), *Aufgabenkultur im Sportunterricht. Konzepte und Befunde zur Methodendiskussion für eine neue Lernkultur* (S. 185-204). Springer VS.
- Neuweg, G. H. (2015). Der Tacit Knowing View – Konturen eines Forschungsprogramms. In G. H. Neuweg (Hrsg.), *Das Schweigen der Könner. Gesammelte Schriften zum impliziten Wissen* (S. 113-132). Waxmann.
- Niederkofler, B. & Herrmann, Ch. (2024). Kognitive Aktivierung und Handlungskontrolle in Aufgabensituationen im Sportunterricht. In J. Wibowo, C. Krieger, E. Gerlach, & F.

- Bükers (Hrsg.), *Kognitive Aktivierung im Sportunterricht (3. vollständig überarbeitete Neuauflage (96-123))*. Universität Hamburg.
- Ostermann, A., Leuders, T., & Philipp, K. (2019). Fachbezogene diagnostische Kompetenzen von Lehrkräften – Von Verfahren der Erfassung zu kognitiven Modellen zur Erklärung. In T. Leuders, M. Nückles, S. Mikelskis-Seifert, & K. Philipp (Hrsg.), *Pädagogische Professionalität in Mathematik und Naturwissenschaft* (S. 93-116). Springer.
- Pfitzner, M. (2013). Fachdidaktische Aufgabenanalyse in Sport. In M. Kleinknecht (Hrsg.), *Lern- und Leistungsaufgaben im Unterricht. Fächerübergreifende Kriterien zur Auswahl und Analyse* (S. 176-192). Klinkhardt.
- Pfitzner, M. (2018). *Lernaufgaben im kompetenzförderlichen Sportunterricht: Theoretische Grundlagen und empirische Befunde*. Springer VS.
- Praetorius, A.-K., & Südkamp, A. (2017). Eine Einführung in das Thema der diagnostischen Kompetenz von Lehrkräften. In A. Südkamp & A.-K. Praetorius (Hrsg.), *Diagnostische Kompetenz von Lehrkräften* (S. 13-18). Waxmann.
- Praetorius, A.-K., Karst, K., Dickhäuser, O., & Lipowsky, F. (2011). Wie gut schätzen Lehrer die Fähigkeitsselbstkonzepte ihrer Schüler ein? Zur diagnostischen Kompetenz von Lehrkräften. *Psychologie in Erziehung und Unterricht*, 58(2), 81-91.
- Rieu, A., Loibl, K., Leuders, K., & Herppich, S. (2020). Diagnostisches urteilen als informationsverarbeitender Prozess – Wie nutzen Lehrkräfte ihr Wissen bei der Identifizierung und Gewichtung von Anforderungen in Aufgaben? *Unterrichtswissenschaft*, 48(4), 503-529.
- Scheja, B. (2019). *Analysen zum kognitiven Anspruch von Mathematikaufgaben – Befunde aus zentralen Prüfungen und Lehrerfortbildungen*. Dissertation, Universität zu Köln.
- Schrader, F.-W., & Helmke, A. (1987). Diagnostische Kompetenz von Lehrern. Komponenten und Wirkung. *Empirische Pädagogik*, 1, 27-52.
- Spinath, B. (2005). Akkuratheit der Einschätzung von Schülermerkmalen durch Lehrer und das Konstrukt der diagnostischen Kompetenz. *Zeitschrift für Pädagogische Psychologie*, 19(1/2), 85-95.
- Sygyusch, R., Hapke, J., Liebl, S., & Töpfer, C. (2022). *Kompetenzorientierung im Sport: Grundlagen, Modellentwurf und Anwendungsbeispiele*. Hofmann.
- Vygotsky, L. S. (1978). *Mind in Society: The Development of Higher Psychological Processes*. Cambridge: Harvard University Press.
- Wank, Y., Chen, A., Schweighardt, R., Zhang, T., Wells, S., & Ennis, C. (2019). The nature of learning tasks and knowledge acquisition: The role of cognitive engagement in physical education. *Eur Phy Educ Rev.*, 25(2), 293-310.
- Wemmer-Roth, W., Praetorius, A.-K., Schreyer, P., & Herbert, B. (2024). Konzeptualisierung und theoretische Fundierung von kognitiver Aktivierung: Ein kritischer Literaturüberblick. In A.-K. Praetorius, W. Wemmer-Roth, P. Schreyer, & M. Brinkmann (Hrsg.), *Kognitive Aktivierung unter der Lupe. Bestandsaufnahme und Möglichkeiten der Weiterentwicklung eines prominenten Konstrukts* (S. 15-51). Waxmann.
- Wibowo, J., Krieger, C., Gerlach, E., & Bükers, F. (2024). *Kognitive Aktivierung im Sportunterricht (3. vollständig überarbeitete Neuauflage)*. Universität Hamburg.
- Widmer, A.-K. (2023). Können Grundschullehrkräfte das kognitive Aktivierungspotential von Rechtschreibaufgaben einschätzen? Jahrestagung der DGfE-Kommission Grundschulforschung und Pädagogik der Primarstufe, Universität Siegen.
- Widmer, A.-K., Hess, M., & Lipowsky, F. (2024). Merkmale kognitiv aktivierender Aufgabenstellungen im Unterricht: Eine Illustration anhand von Beispielaufgaben. In A.-K. Praetorius, W. Wemmer-Roth, P. Schreyer, & M. Brinkmann (Hrsg.), *Kognitive Aktivierung unter der Lupe. Bestandsaufnahme und Möglichkeiten der Weiterentwicklung eines prominenten Konstrukts* (S. 153-167). Waxmann.

Digital Transformation of a Diagnostic Procedure for Swimming Skill Analysis: Teachers' Evaluation of a Low-Fidelity App Prototype (Click Dummy) of the 'Owl's Eye'

von Inga Fokken, Ilka Staub, Michel Brinkschulte, Edgar Sauerbier, Nele Schlapkohl & Tobias Vogt

Abstract: Physical education teachers are challenged with accurately analysing their learners' swimming skills in order to design appropriate swimming lessons. To support them in this task, the 'Owl's Eye' diagnostic procedure was developed and evaluated. The 'Owl's Eye' systematically evaluates basic aquatic skills and easily performs an initial classification of learners' skills in their first swimming lessons. Following the successful testing of the paper-based version, the 'Owl's Eye' was further developed in digital form as part of a project funded by the Federal Ministry of Research, Technology and Space (BMFTR). The aim of this study is to investigate users' perspectives on the app prototype. To this end, interactions with a click dummy were captured using 'think-aloud' protocols and subsequently analysed by means of a qualitative analysis. The findings show that the app prototype was predominantly evaluated positively in terms of its design, usefulness, and effectiveness in dealing with limited teaching time. At the same time, suggestions for improvement were made, such as clearer navigation, an introductory video, and the ability to use functions offline. Overall, the findings of this study suggest that the digital transformation of the 'Owl's Eye' procedure further facilitates teachers' diagnostic work in school swimming lessons.

Keywords: app development, digital transformation, motor skill assessment, physical education, school swimming lessons, teacher education

Digitale Transformation eines Diagnoseverfahrens zur Analyse des schwimmerischen Könnens – Evaluation eines Low-Fidelity-App-Prototyps (Click Dummy) des »Eulenblicks« durch Lehrkräfte

Zusammenfassung: Sportlehrkräfte stehen vor der Herausforderung, das schwimmerische Können ihrer Lernenden genau zu analysieren, um einen passgenauen Schwimmunterricht zu gestalten. Um sie dabei zu unterstützen, wurde das Diagnoseverfahren »Eulenblick« entwickelt und evaluiert. Mit dem »Eulenblick« kann systematisch das schwimmerische Können der Lernenden analysiert werden, er ermöglicht eine einfache Klassifizierung der Lernenden in ihren ersten Schwimmstunden. Nach der erfolgreichen Evaluation der papierbasierten Version wurde der »Eulenblick« im Rahmen eines vom Bundesministerium für Forschung, Technologie und Raumfahrt (BMFTR) geförderten VIP+ Validierungsprogramms in digitaler Form weiterentwickelt. Ziel der Studie ist, die Perspektiven der Nutzenden auf den App-Prototypen zu untersuchen. Dazu wurde der Umgang mit einem Click Dummy mithilfe von »Think-Aloud«-Protokollen erhoben und anschließend anhand einer qualitativen Inhaltsanalyse ausgewertet. Die Ergebnisse zeigen, dass die App u.a. hinsichtlich ihres Designs, ihrer Nützlichkeit und ihrer Effektivität im Umgang mit begrenzter Unterrichtszeit überwiegend positiv bewertet wurde. Gleichzeitig wurden Verbesserungsvorschläge gemacht, wie eine klarere Navigation, ein Einführungsvideo und die Möglichkeit, Funktionen offline zu nutzen. Insgesamt deuten die Ergebnisse dieser Studie darauf hin, dass die digitale Transformation des »Eulenblick«-Verfahrens das diagnostische Handeln der Lehrkräfte im Schwimmunterricht weiter vereinfacht.

Schlagwörter: App-Entwicklung, Ausbildung von Sportlehrkräften, Bewertung sportmotorischer Fertigkeiten, Digitale Transformation, schulischer Schwimmunterricht

Digital Transformation of a Diagnostic Procedure for Swimming Skill Analysis: Teachers' Evaluation of a Low-Fidelity App Prototype (Click Dummy) of the 'Owl's Eye'

1 Introduction

It is well-accepted that the objective of (school) swimming lessons is to enhance learners' swimming competence in order to prevent drownings and to foster a lifelong engagement in all aquatic activities (Mekkaoui et al., 2025). The conditions for implementing appropriate swimming lessons are challenging in many ways. In addition to often very limited teaching time, the diversity of learners' swimming skills is one of the main challenges. While some primary school children can already swim before they start swimming lessons at school, others – even at secondary school – may have limited experience of swimming, be unable to swim, or even be afraid of the water (DLRG, 2022). In order to provide the best possible support for all learners, especially those who cannot yet swim, it is crucial to ensure that swimming lessons are adapted to the learner's individual swimming skill level and provide opportunities for learning growth (Mekkaoui et al., 2022). This requires a high level of didactic and methodological competence on the part of the physical education (PE) teachers, but also an accurate analysis of the learners' swimming skills at the beginning of the swimming lessons.

Typically, learners are simply classified as either swimmers or non-swimmers (Fokken et al., 2023, see Appendix 1). But for the purpose of adaptive lesson structuring, it is insufficient to merely ascertain whether or not the learners can swim (e.g., by determining their ability to swim an assigned distance in deep water or a certain quality of stroke technique). In order to learn to swim properly, children must master different phases. Firstly, they must become familiar with the water. Once this initial habituation phase is complete, they must learn basic aquatic skills. This includes being able to fully submerge, breath rhythmically, float on the front and back and glide with reduced drag (Bissig & Gröbli, 2004; Wilke & Daniel, 2007). Once this phase is reached, children can learn elementary water locomotion and, based on this, the well-known swimming stroke techniques (Rheker, 2011). As these skills form the foundation of comprehensive aquatic education (Costa et al., 2012; Langendorfer & Bruya, 1995), they require particular attention during the diagnostic process.

There are some validated instruments available for assessing children's actual aquatic skills, such as the SMACC¹ (Moreno Murcia et al., 2020), the AAST² (Mertens

1 Scale to Measure Aquatic Competence in Children

2 Actual Aquatic Skills Test

et al., 2021), and the SCAS³ (Sundan et al., 2023). However, as these instruments were primarily developed for research purposes, they are only partially suitable for everyday use in school swimming lessons due to their methodological requirements.

To address this issue and explore ways to support PE teachers in their diagnostic work, the 'Owl's Eye' procedure for analysing swimming skills was developed, based on the 'Assessment of Basic Aquatic Skills' (ABAS, Vogt & Staub, 2020). The ABAS was designed to be easy to apply and reliable, and has been validated among coaches, teachers, students and parents with varying levels of experience in teaching swimming. However, the ABAS comprises a total of 19 test tasks. Although the ABAS does not require a complicated experimental setup, it is still too time-consuming to be practical for use in schools. Nevertheless, it is precisely in schools that the 'Owl's Eye' is intended to be used.

To this end, the number of ABAS test tasks has been reduced and restructured. This enables an initial classification of learners' skills to be carried out using only five core tasks (e.g., Which learners are able to fully submerge, or float on their front and back?). These five core tasks are complemented by a combined task test to verify the children's ability to swim in deep water. In addition to the small number of tasks required for an initial analysis of swimming skills, what makes the 'Owl's Eye' procedure special is that it is directly linked to a didactic framework for comprehensive swimming education (Staub & Fokken, 2020). This allows teachers to immediately classify children into one of five colour-coded phases and develop support measures. This could encourage more teachers to use the 'Owl's Eye' as a diagnostic tool in their swimming lessons.

As well as ensuring that motor competence assessment tools are valid and reliable for the population in which they are used, it is also important that they are feasible for the intended users, such as teachers (Lander et al., 2015). Consequently, the 'Owl's Eye' was subjected to an evaluation of its feasibility and teachers' subjective value at an early stage in the development phase (Fokken et al., unpublished findings currently under review, see Appendix 2). A total of 72 teachers ($N = 14$ in a qualitative pre-study and $N = 58$ in a quantitative questionnaire) therefore tested the 'Owl's Eye' (paper-pencil-version) in practice and then evaluated the procedure. For this purpose, teachers were provided with a manual and prepared documentation sheets for recording learners' swimming skills. The procedure was generally well received. The majority of teachers evaluated it as useful and valuable for analysing learners' swimming skills. For instance, 83.3%⁴ of respondents agreed with the statement that the 'Owl's Eye' is a reliable method of gathering diagnostic information for planning appropriate swimming lessons, selecting 'correct' (64.8%) or 'rather correct' (18.5%). As high-quality data is fundamental to determining

3 Swimming Competence Assessment Scale

4 All percentages reported here refer to the quantitative study involving 58 participants.

learners' current skills and support needs (Jäger, 2009; Leuders et al., 2020; van Ophuysen et al., 2013), the improvement in data collection perceived when using the 'Owl's Eye' is particularly significant. Furthermore, 75.4% ('correct': 45.1%; 'rather correct': 29.8%) agreed that the 'Owl's Eye' helped them to better structure the swimming skill analysis. Most respondents also stated that using the procedure saved them work in preparing the swimming skill analysis ('correct': 46.4%; 'rather correct': 26.8%). Regarding their own perceived competence, 72.7% of respondents ('correct': 36.6%; 'rather correct': 36.6%) said that using the 'Owl's Eye' made them feel more confident in analysing their learners' swimming skills. Overall, the surveyed teachers believe that the 'Owl's Eye' procedure improves the collection of data and, consequently, the analysis of swimming skills. However, the study also revealed that teachers clearly preferred a digital version of the 'Owl's Eye', as this would minimise the use of paper-based materials and provide a more comprehensive and accessible overview.

The transformation of the 'Owl's Eye' procedure into an app represents an innovation for physical education in school. Although a wide range of apps and browser-based tools has meanwhile been developed to support teachers and learners in physical education, these tools typically serve other purposes. Many apps are designed to divide teams, document scores, measure time, or quantitatively record physical performance (e.g., *SportZens*). In addition, there are numerous apps that provide games and exercises for physical education and use artificial intelligence to suggest lesson plans; however, these suggestions are usually not linked to the actual skill level of the class (e.g., *LessonPrep*).

Furthermore, some apps present ideal-typical movement patterns of various sports skills and include slow-motion and freeze-frame functions to support learners' self-directed learning phases. The app 'AssesslinkPE' is designed for physical education skill testing and uses a 3D motion capture sensor to instantly compare a performed movement with biomechanically optimal execution. However, the app currently includes only nine skills (e.g., volleyball forearm pass, tennis forehand, vertical jump and landing, overhand throw) and is therefore of limited applicability. Moreover, it is not clear from the information provided on the associated website to what extent the app allows for functional deviations from ideal-typical movement patterns. The 'Video Delay' app records the learners' movements and then plays them back multiple times with a delay that can be set as desired. This enables learners to observe their own movements and identify any errors. Teachers can also use the video recording to discuss the execution of the movement with the learner immediately afterwards. As the video is not saved, the school's data protection requirements are met.

Outside the domain of physical education, a range of predominantly digital diagnostic tools has been developed in recent years to support teachers in identifying

learning progress and individual support needs. For example, the 'Ernst Klett Verlag', a German educational publishing house, offers test booklets for reading and writing across different age groups. Based on the test results, teachers can plan appropriate support measures.

The digital transformation of the positively evaluated 'Owl's Eye' procedure builds on this approach for swimming education. Test results are recorded digitally, and the analysis can be displayed for individual learners, subgroups, or the entire class. Based on these results, content for subsequent swimming lessons is derived immediately and in a targeted manner.

2 The challenge of analysing motor/swimming skills

In the context of educational diagnostics, analysis refers to a predominantly qualitative approach in which the information collected is examined in detail in order to identify underlying structures, connections or causes and to derive consequences for teaching or the individual support of the learners (Schilling & Leiss, 2022). Therefore, teachers' competencies should include the ability to use appropriate strategies and methods for gathering and processing information regarding to learners' current (motor) skills, as well as the ability to analyse and correctly interpret the resulting data (Nitko, 2004). A model of diagnostic quality by van Ophuysen and Behrmann (2015) divides the diagnostic procedure into two process components (information gathering and information processing) and two outcome components (data and judgment). In this model, not only the final judgment, but also the diagnostic process, is considered to indicate the diagnostics' quality. However, research has shown that teachers often lack diagnostic competence in teaching and learning in general, as well as in physical education specifically (Brühwiler, 2017; Lounsbery & Coker, 2008; Niederkofler et al., 2018; Seyda & Langer, 2020; Ward et al., 2021; Wilson et al., 2020). Consequently, they often do not succeed when analysing their learners' (motor) skills. Some studies suggest that much information is gathered incidentally and not systematically documented, which can compromise the quality of data for later decision-making purposes (Edelenbos & Kubanek-German, 2004). These observations find equal application in the context of swimming lessons.

A study of Fokken et al. (2023, see Appendix 1) has shown that teachers' approaches to accurately analyse learners' swimming skills seem not to be optimal in most cases. The procedures used are almost exclusively informal or semiformal, with the resulting analysis mostly based on the teachers' subjective impressions. In school pedagogy, diagnostic approaches are classified as informal, semi-formal and formal (Heck & Scheuer, 2019). Informal diagnostics are characterised as unsystematic and subjective (e.g., spontaneous observations), while semiformal diagnostics are structured but lack standardisation. Formal diagnostics follow test-theoretical, standardised procedures. Formal methods serve to complement or validate less formal

approaches (Behrmann & Kaiser, 2017). Therefore, it has been pointed out for some time that teachers should use semi-formal and formal diagnostic procedures more frequently in order to reduce assessment bias (Behrmann & Kaiser, 2017; Seyda & Langer, 2020). At the same time, the use of motor skill tests as a formal diagnostic method in physical education is the subject of controversial discussions (Funke-Wienecke, 2007; Gerlach et al., 2014; Schweihofen, 2013).

In the study mentioned above of Fokken et al. (2023), only 4.9% of the participants stated that they use motor skill tests to analyse their learners' swimming skills (Fokken et al., 2023), although more than half of the teachers (60.6%) believe motor skill tests to be helpful in analysing motor skills in physical education. However, 40.1% of teachers consider motor skill tests difficult to carry out from an organisational point of view. This concern is not unfounded. For instance, the manual of the MOBAK test, which assesses basic motor skills, specifies that a test administrator should allow approximately 35 minutes to conduct the test with three to four children. Physical education teachers conducting the test with a whole class are advised to spread the tasks over several lessons (Herrmann & Seelig, 2020). Although the analysis of learners' motor skills is the first step toward targeted instructional strategies, it can be assumed that there are many limitations to the field application of evidence-based motor skill tests in school settings (Barnett et al., 2020; Hulteen et al., 2020) and that they are often not suitable for school use. Well-documented barriers using formal assessment tools include limited time, an overcrowded curriculum, and teachers' lack of competence, expertise or training (Griffiths et al., 2018; Klingberg et al., 2018; Lander et al., 2017; Lander et al., 2015). In order to support teachers in their diagnostic work, the barriers mentioned above must be taken into account when developing new tools, and teachers' needs must be incorporated into the development process.

3 Digital support for analysing learners' swimming skills

To date, there is no established digital tool specifically designed for the analysis of swimming skills in an educational setting. The development of such a tool is therefore a first attempt to address this issue. Sousa-Sá et al. (2023) state, that digital solutions for assessing motor competence could help to overcome existing barriers to motor competence assessment and might facilitate its integration into teaching practice. The transformation of the 'Owl's Eye' into a digital application therefore aims to increase accessibility and usability, encouraging wider use among teachers. This work is supported by the BMFTR through a three-year VIP+ project (July 2023 – June 2026), in which the digital transformation proceeds multiple iterative steps accompanied by regular validation. The present study represents an early element of the validation process and was guided by the following main research question:

How do teachers evaluate the usability of the low-fidelity click dummy of the 'Owl's Eye' app, used to simulate digital support in analysing learners' swimming skills, and what added value do they perceive in using the app for their diagnostic work?

4 Materials and methods of the Click Dummy Study

4.1 Study design

A low-fidelity prototype of the 'Owl's Eye' application in the form of an interactive click dummy was developed to visualise the design and navigation structure of the app. The 'Owl's Eye' app is designed to provide digital support for the structured observation and analysis of swimming skills in school settings. It allows teachers to create and manage class lists and maintain attendance records for their swimming groups. The core function of the app is to guide teachers step by step through the diagnostic procedure known as the 'Owl's Eye'. Results can be documented directly within the app. The app then automatically classifies learners according to their skill level based on these entries. It also generates recommendations for suitable games and practice activities derived from the diagnostic outcomes.

The click dummy was created using *Figma*, a collaborative interface design tool that enables app designers to create, share, and iterate user interface designs in real time. It simulated essential user interactions, enabling participants to navigate through the prototype, but did not include any functional back-end processes. In this version, users could click through the app's screens and experience the intended flow and interface design, but no real data processing or system functions were implemented. This approach enabled early usability testing and the collection of feedback regarding layout clarity, intuitive navigation, and task flow before technical implementation.

The study followed a qualitative exploratory design to investigate participants' experiences. Using the think-aloud method (Cooke, 2010), participants were asked to verbalise their thoughts while completing predefined tasks within the prototype. The think-aloud method was chosen because it provides direct, real-time insights into teachers' cognitive processes while interacting with the click dummy of the 'Owl's Eye' app prototype. By verbalising their thoughts while navigating, participants reveal how they understand the interface, make decisions, and solve problems, enabling researchers to immediately identify potential sources of confusion or error. This real-time data is particularly valuable for identifying usability issues that might remain undetected in retro perspective interviews. Moreover, the method captures context-specific feedback, as participants interact with the prototype as it would be used in practice.

4.2 Participants

The click dummy's user interface was evaluated by 11 PE teachers, including five pre-service and six in-service teachers (63.6% female, 36.4% male; $M = 33.3$ years, $SD = 11.14$). All participants (P1-P11) were part of the 'Owl's Eye' project network and therefore already familiar with the former paper-based version of the procedure. Their general teaching experience ranged from 2 months to 28 years, with their experience in teaching swimming lessons distributed similarly. Most participants were employed at primary schools, while some taught at secondary school types, including comprehensive schools and grammar schools. All participants provided written informed consent before taking part in the study, in accordance with the ethical standards set out in the Declaration of Helsinki and with approval from the researchers' university Ethics Committee (approval number 173/2021).

The required sample size for usability testing depends on the number of test cycles needed to adequately optimise a prototype and on the number of participants per cycle required to achieve data saturation. Previous studies have shown that meaningful improvements to a prototype can usually be achieved within two to three test cycles, with data saturation often already reached with five to seven participants per cycle (Molich, 2010). Since the present study only tested the click dummy and further validation studies are planned for subsequent prototype versions, a single test round with 11 participants was conducted and considered sufficient to gain initial insights.

4.3 Data collection

Data were collected through individual think-aloud sessions, a qualitative method commonly used in usability studies (Cooke, 2010; Fan et al., 2020). All sessions were conducted online in March 2025 and lasted $18:55\text{min} \pm 4:15\text{min}$. The 11 sessions were facilitated by two researchers, ensuring consistency in instructions and procedures across participants. At the beginning of each session, participants were presented with a brief user story to illustrate a real-life swimming lesson scenario. Then they interacted with the click dummy, including entering information about learners' attendance and swimming skill levels; however, all entries and the corresponding analysis results were pre-set.

A small selection of screenshots from the click dummy can be seen in Fig.1⁵. The first screenshot⁶ displays the start screen, from which the personal login process is carried out using a password. The second screenshot shows the attendance list, where learners' attendance for swimming lessons can be documented by simple

5 The app will initially only be available in German. An English version is planned for the medium term.

6 Screenshots are numbered consecutively from left to right across rows.

clicking. Using the 'attention' icon, teacher receive a direct indication that there is specific information to be considered for '*Anna B.* ', such as details regarding a medical condition or disability. The third screenshot presents the selection options related to the app's core functions. Teachers can choose whether they wish to view the current learning status of their class or conduct the analysis. The fourth screenshot illustrates, using again the example of '*Anna B.*', the five core tasks that must be successfully completed as part of the 'Owl's Eye' procedure. Teachers can record task completion or non-completion with a single click. The fifth screen provides teachers with an overview, following completion of the 'Owl's Eye', showing how many learners have been assigned to each phase of learn to swim. This is easily identifiable through colour coding. The sixth screenshot shows an excerpt from the implementation of the combined task test with two learners from the class.

Participants were encouraged to express their thoughts, comments, and questions as comprehensively and unfiltered as possible while using the app. The test administrator did not interact with the participants, except to answer questions about the process or any technical issues. Participants' comments were captured via Webex (version 44.9.1.30809, Cisco Systems) and transcribed verbatim, providing a detailed record for descriptive analysis.

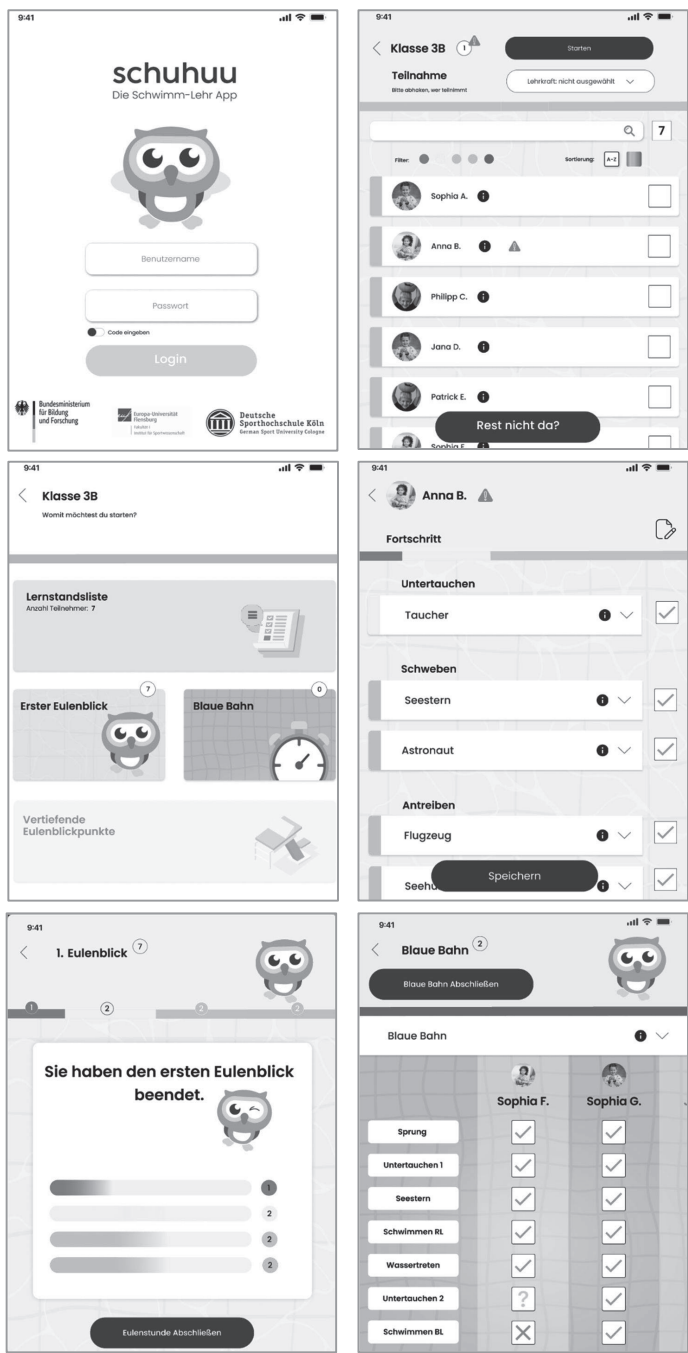


Figure 1: Screenshots Click Dummy

4.4 Data analysis

The think-aloud data were analysed inductively following the principles of qualitative content analysis according to Kuckartz (2018) to gain a deeper understanding of teachers' perceptions of the 'Owl's Eye' app prototype and its' potential for use in school swimming lessons. The analysis was conducted using MAXQDA software, which supported the systematic coding and organisation of the data. As the study did not rely on structured interview questions, the analysis followed an iterative and exploratory process. This approach resulted in personalised statements, with participants focusing on different aspects of the app.

Two independent researchers conducted the coding process, collaboratively developing the category system. Ultimately, the participants' statements were subsumed under four overarching categories that captured the main aspects of their evaluation (see Table 1):

- the app's orientation and structure (a),
- layout and colour schemes (b),
- functionality of the app in terms of conducting the analysis (c),
- typography and editorial text (d).

Individual statements that could not be assigned to a category but still seem worth mentioning have been compiled under (e) 'other insights'.

5 User feedback: Findings from the Click Dummy Study

For presentation, the findings were organised in a table following the structure of Sousa-Sá et al. (2024), who investigated teachers' perceptions of a digital app for assessing motor competence assessment. The table contains the categories mentioned above, which are structured under the headings 'Summary of Feedback', 'Teacher Quotes' and 'Changes to be taken'. The findings of the inductive content analysis are summarised under the heading 'Summary of Feedback'. Within the section 'Teacher Quotes' some examples of direct quotes from the teachers (P1-P11) can be found. Under the heading 'Changes to be taken', suggestions for changes and possible additions have been derived from the findings of the 'think-aloud' sessions.

Table 1: Click Dummy Testing Findings

(Inductive) Categories	Summary of Feedback	Teacher Quotes (P1-P11)	Changes to be taken
(a) Orientation & Structure	- Easy registration via login button - App is clearly laid out - Overall nice and simple - Confusing orientation at the beginning - App is reduced to the essentials - Good overview - Professional impression - Lots of nice features (videos, pictures)	'So, I find it clear at first glance.' (P3) 'When I first opened the app, I wasn't quite sure where to start.' (P7) 'You have to familiarise yourself with it a little first...' (P2) 'Now I have to find my way around a bit.' (P4)	- Insert a short tutorial explaining how to navigate the site. - Insert orientation aids, e.g., icons/ buttons.
(b) Layout & Colour Schemes	- Very subtle colours - Beautiful colour scheme - More colour would be nice - A slightly more neutral background would be better - Appealing owl emoji/symbol, very cute - Pleasantly surprised	'I also really like the colour scheme here, because it's subtle but fits the theme.' (P5) 'However, I think a few more colours could be added.' (P7) 'The use of colour is very helpful in all areas for identifying which child is at which learning level.' (P11)	- Conduct further surveys to verify trends in both directions and adjust the colour scheme accordingly if necessary.
(c) Functionality of the app (in terms of conducting the analysis)	- Automatic sorting of children according to analysis is helpful - Colour coding shows who is at what stage and what still needs to be checked - Attendance is easy to record - Analysis with guidance useful - Problems with the order due to lack of overview - Confusion because two tasks are listed one below the other	'Okay, now everyone will be processed in order. I think that's actually quite good.' (P7) 'It's cool that I can see right away which group the children are assigned to.' (P5) 'Okay, that's easy, I can see all the children right away and just go through them all.' (P8) 'It's good that it's arranged like this. So, it's very easy to tick everything off.' (P9) 'That means the child hasn't done it yet and it's not ticked off here that they've done it. But I can still	- Insert an overview or explanatory video at the beginning explaining how to use the app.

<i>(Inductive) Categories</i>	<i>Summary of Feedback</i>	<i>Teacher Quotes (P1-P11)</i>	<i>Changes to be taken</i>
	- Clear and systematic because you can proceed according to the criteria - Separate areas for the different parts of the diagnosis	<i>see the things that are missing. Now I'm confused.'</i> (P1) <i>'What I like is that when you click on the individual exercises, you get an evaluation straight away.'</i> (P4)	
<i>(d) Typography & Editorial Text</i>	- Meaning of check marks and crosses unclear - Meaning of icons not always clear - Videos and photos very helpful in confirming how things should look like - Support provided by illustrations - Change in position of the 'Save' button unfortunate - Sometimes not clear that you have to scroll to see more tasks or children	<i>'What are the crosses for?' (P9)</i> <i>'Apparently, this is where the diver's instructions are stored. That's a cool idea.'</i> (P6) <i>'Yes, I think it's really good that it's shown again so that you can quickly get an overview.'</i> (P9) <i>'I'm not quite sure what the exclamation mark is for.'</i> (P3) <i>'Because, well, I have to say that I have a question mark myself when it comes to the question mark.'</i> (P8) <i>'I'm missing the legend.'</i> (P4)	- Add of a help section. This should include the meaning of symbols and icons.
<i>(e) Other Insides</i>	- Desire to be able to make notes/comments about the children - Concerns that using iPads in the swimming pool is not practical - Question about how the app works without an internet connection - Desire for several teachers to be able to access the app	<i>'It would be cool if teachers could add their own comments.'</i> (P8) <i>'But as I said, I don't think running around the edge of the pool with an iPad is very practical.'</i> (P9) <i>'So, without Wi-Fi... How does it work?' (P1)</i> <i>'I'm a big fan of the idea of digitising the Owl's Eye.'</i> (P3)	- Checking the technical opportunities to add a comment function - Consideration of potential network issues during the development process

The following summary provides an overview of the key outcomes from the think-aloud sessions. It highlights the central feedback that derived from participants' comments on the click dummy, as detailed in Table 1.

When they opened the app for the first time, most participants commented positively on its appearance (P3, P5-P11). They described the design as appealing (P1, P5, P11) and clear (P1, P3, P5-11). After a brief orientation on the start screen, the teachers proceeded in different ways when interacting with the click dummy: Some looked at the different areas of the app first (P1, P3, P6, P7, P11), while others immediately started making entries in the attendance list, for example (P2, P4, P5, P8-P10). This resulted in some initial comments about minor orientation issues within the app (P2, P4). Some participants mentioned the colour scheme positively (P3, P7, P11), particularly the automatic assignment of children based on the key tasks of the 'Owl's Eye'. At the same time, there were mixed comments about the use of colours, with one wanting more (P7) and another wanting less (P10). The automated assignment of children after completion of the 'Owl's Eye', which forms the core of the app, was widely regarded as a valuable feature (P2, P4-P9, P11). The guided analysis was rated as useful by several participants (P1, P3, P8, P10, P11). However, some participants found it difficult to understand the order of operations (P4, P6, P7, P9). Participants also expressed a need for guidance on menu navigation and the system, for example in the form of an introductory video, particularly for those with little experience of the procedure (P6-P8). The option to review individual test tasks in text or video form was highlighted as a positive feature (P1, P3, P5-P7, P10). There was widespread confusion regarding the symbols and icons used, and there was a general call for a 'legend' (P1, P3, P4-P6, P8, P9, P11). Furthermore, it was mentioned several times that participants lacked the option of adding individual notes to learners' profiles (P8, P11). Additionally, there were concerns about the use of online devices in swimming lessons, as it might be difficult to establish an internet connection and ensure that the devices did not get wet (P1, P4, P9).

6 Discussion

The present study aimed to explore how teachers evaluate the usability of the 'Owl's Eye' low-fidelity prototype, and the added value they perceive in using this digital tool to support diagnostic work during swimming lessons. The findings offer valuable insights into the acceptance of the prototype in terms of usability, as well as its perceived contribution to teachers' diagnostic work. This is particularly relevant in light of recent work on pedagogical diagnostics in swimming, as well as the need for applicable procedures to be used in schools (Fokken et al., 2023). In addition, digital technologies are becoming an increasingly important part of teachers' daily work (Sousa-Sá et al., 2024). It is therefore important to provide them with diagnostic tools also in the form of digital solutions.

Overall, the teachers' positive evaluation of the clarity, structure and visual design of the app prototype indicates a high level of perceived usability. This findings align with research suggesting that intuitive layouts and clear workflow, along with the replacement of paperwork through app-based solutions, are crucial prerequisites for the successful adoption of digital tools in educational contexts (Sousa-Sá et al., 2024).

At the same time, usability issues identified in the think-aloud protocols, such as occasional confusion about the order of operations and unclear symbolism, indicate that high overall acceptance does not preclude the need for further optimisation. In line with previous research on assessment tools in physical education, even minor usability barriers may hinder sustained use in time-constrained settings (Barnett et al., 2020; Griffiths et al., 2018). Accordingly, teachers' feedback highlights the need for additional scaffolding elements (e.g., tutorials or legends) to support work routines, even though the 'Owl's Eye' prototype has already received a positive usability evaluation that underscores its practical benefits.

The findings gain particular relevance when situated within the broader research context of diagnostic competence in physical education, beyond usability. Numerous studies have demonstrated, that PE teachers often struggle to analyse learners' motor skills in a structured and valid manner (Seyda & Langer, 2020; Ward et al., 2021; Wilson et al., 2020). While an accurate analysis of the learners' current skills is recognised as a prerequisite for adaptive teaching, teachers rarely use standardised motor tests because they are perceived as impractical in a school context (Fokken et al., 2023). At the same time, teachers' perspectives differ regarding how the results of motor tests should be interpreted and used (Gerlach et al., 2014; Schweihofen, 2013).

Against this backdrop, the positive evaluation of the 'Owl's Eye' app' structured diagnostic workflow by teachers is particularly significant. The automatic classification of learners and clear allocation to the five phases of learn-to-swim were seen as main advantages and help teachers gain understanding of their learners' current swimming skills (Fokken et al., unpublished findings currently under review, see Appendix 2). This finding addresses the lack of practicable diagnostic tools that link the skill analysis with instructional decision-making in physical education (Knudson, 2013).

Although there are many apps that support PE teachers by providing exercises, games or documentation features (e.g., *SportZens*, *LessonPrep*), these tools rarely connect instructional suggestions to the analysis of the learners' skills. Similarly, video-based applications, such as *'Video Delay'*, primarily enhance visual feedback, leaving the interpretation of learners' performance largely on teachers' subjective judgement and professional experience (Vogt & Staub, 2020). In contrast, the 'Owl's Eye' provides a structured framework that guides teachers through a stan-

standardised diagnostic procedure based on defined observation criteria and core tasks. In this function as a digital decision-support system that links diagnostic results directly to teaching practice, extending well beyond the purpose of mere video feedback. From a research perspective, this represents a significant development, as the app not only supports the analysis of swimming skills but also facilitates the translation of diagnostic information into pedagogical action. Although this linkage had already been theoretically established through the didactic framework for comprehensive swimming education (Staub & Fokken, 2020), it previously relied on teachers to actively perform this translation themselves by drawing on the framework. The 'Owl's Eye' app now supports and partially automates this process in everyday teaching practice and teachers' appreciation of this feature suggests that the app could help to reduce uncertainty in instructional planning.

The findings also suggest potential long-term implications for teachers' diagnostic competence. Teachers indicated that the guided structure and automated process of the 'Owl's Eye' app could encourage the continued use. This is particularly relevant, as repeated engagement with the app can be assumed to enhance teachers' ability to link observed swimming skills with learning phases and instructional goals. While the present study cannot empirically confirm competence development, the perceived support suggests that the app may function as a learning tool for teachers, gradually strengthening their diagnostic competence through repeated use.

7 Limitations

While the findings of the present study provide valuable insights into the usability and potential benefits of the digital 'Owl's Eye' prototype, several limitations must be considered.

First, the study was based on a low-fidelity app prototype (click dummy) that did not yet include interactive data processing, automated feedback, or real-time instructional recommendations. Consequently, the findings have limited ecological validity because teachers' evaluations were based on anticipated rather than fully enacted use. Therefore, future research should investigate the performance of a fully functional version of the app in authentic swimming lessons, examining its effects on diagnostic accuracy, instructional decision-making and teaching quality.

Second, the sample size was relatively small ($N = 11$) and consisted exclusively of teachers who were already familiar with the paper-based version of the 'Owl's Eye'. This familiarity may have positively influenced their perceptions of usability and usefulness. Further studies should include a broader and more heterogeneous sample, including teachers without prior experience with the 'Owl's Eye'.

Thirdly, although 'think-aloud' protocols proved effective in gaining detailed insights into usability issues and perceived added value, they cannot fully capture

teachers' behavior and challenges in actual lesson situations. Consequently, some of the practical constraints of swimming lessons may be underrepresented. Complementary methods, such as classroom observations, would enable a more thorough evaluation of how the app is actually used.

Finally, external factors such as varying access to online devices or infrastructure (e.g., internet access in indoor pools) were not experimentally controlled but could significantly influence real-world implementation.

8 Conclusion

Overall, the findings of this study indicate that digitally transforming the 'Owl's Eye' procedure represents a promising way to support diagnostic practices in school swimming lessons. From the perspective of educational diagnostics, effective diagnostic processes require teachers to systematically gather and process information about learners' current skills in order to derive appropriate (swimming) lessons (Nitko, 2004; van Ophuysen & Behrmann, 2015). Against the background of research showing that teachers often experience difficulties with diagnostic competence in teaching in general and in physical education in particular, the app-based implementation of the procedure appears to provide meaningful support in structuring these processes. Teachers perceived the low-fidelity app prototype as a usable and pedagogically meaningful tool for analysing learners' swimming skills and for translating diagnostic information into instructional guidance. When considered alongside existing research, the findings suggest that the app fills a gap in physical education by providing a structured and practical diagnostic tool that meaningfully connect the analysis with instruction in swimming education. By supporting teachers in systematically collecting and interpreting learners' swimming skills data, the 'Owl's Eye' app has the potential to contribute to more adaptive, inclusive and effective swimming lessons. At the same time, the study emphasises that the long-term use of such a tool depends on resolving identified usability barriers, such as the technical infrastructure of swimming pools and situational constraints. Further research is needed to examine the effectiveness of a fully functional version in real teaching conditions. If successful, this approach could also be used as a model for diagnostic support tools in other areas of physical education and motor skill development.

References

- Barnett, L. M., Stodden, D. F., Hulteen, R. M., & Sacko, R. S. (2020). Motor competence assessment. In T. Brusseau, S. Fairclough, & D. Lubans (Eds.), *The Routledge Handbook of Youth Physical Activity* (1 ed.). Routledge.
- Behrmann, L., & Kaiser, J. (2017). Das Modell pädagogischer Diagnostik nach Ingenkamp und Lissmann. In A. Südkamp & A.-K. Praetorius (Eds.), *Diagnostische Kompetenz von Lehrkräften* (pp. 59-63). Waxmann.

- Bissig, M., & Gröbli, C. (2004). *SchwimmWelt: Schwimmen lernen - Schwimmtechnik optimieren*. SchulVerlag.
- Brühwiler, C. (2017). Diagnostische und didaktische Kompetenz als Kern adaptiver Lehrkompetenz. In A. Südkamp & A.-K. Praetorius (Eds.), *Diagnostische Kompetenz von Lehrkräften* (pp. 123-134). Waxmann.
- Cooke, L. (2010). Assessing concurrent think-aloud protocol as a usability test method: A technical communication approach. *IEEE Transactions on Professional Communication*, 53(3), 202-215. <https://doi.org/10.1109/TPC.2010.2052859>
- Costa, M. J., Bragada, J. A., Marinho, D. A., Silva, A. J., & Barbosa, T. M. (2012). Longitudinal interventions in elite swimming: a systematic review based on energetics, biomechanics, and performance. *Journal of Strength and Conditioning Research / National Strength & Conditioning Association*, 26(7), 2006-2016. <http://www.ncbi.nlm.nih.gov/pubmed/22531620>
- DLRG. (2022). Schwimmfähigkeit der Bevölkerung 2022. *DLRGPresse*.
- Edelenbos, P., & Kubanek-German, A. (2004). Teacher assessment: The concept of 'diagnostic competence'. *Language Testing*, 21(3), 259-283. <https://doi.org/10.1191/0265532204lt284oa>
- Fan, M., Shi, S., & Truong, K. N. (2020). Practices and Challenges of Using Think-Aloud Protocols in Industry: An International Survey. *Journal of Usability Studies*, 15(2), 85-102.
- Fokken, I., Staub, I., & Vogt, T. (2022). Lernstandsdiagnostik im Schwimmunterricht - Lerngruppen analysieren und Unterricht passgenau gestalten. *sportunterricht*, 70.
- Fokken, I., Staub, I., & Vogt, T. (2023). Physical education teachers' swimming skill analysis in 6- to 12-year-old children: Findings from an online survey. *Journal of Teaching in Physical Education*, 43(2), 333-342. <https://doi.org/10.1123/jtpe.2022-0297>
- Funke-Wienecke, J. (2007). Bewegungsdiagnose praktisch. *sportpädagogik*, 31(1), 4-9.
- Gerlach, E., Leyener, S., & Herrmann, C. (2014). »Denn wir wissen nicht, was wir messen?«. Zur Frage der Output-Diagnostik im Sportunterricht mit Hilfe von motorischen Tests. *sportunterricht*, 63(7), 194-200.
- Griffiths, A., Toovey, R., Morgan, P. E., & Spittle, A. J. (2018). Psychometric properties of gross motor assessment tools for children: a systematic review. *BMJ Open*, 8(10). <https://doi.org/10.1136/bmjopen-2018-021734>
- Heck, S., & Scheuer, C. (2019). Diagnostic Competence of Physical Education Teachers. In M. Peters (Ed.), *Encyclopedia of Teacher Education*. Springer. https://doi.org/http://dx.doi.org/10.1007/978-981-13-1179-6_339-1
- Herrmann, C., & Seelig, H. (2020). *MOBAK-5-6: Motorische Basiskompetenzen in der 5. und 6. Klasse. Testmanual*
- Hulteen, R. M., Barnett, L. M., True, L., Lander, N. J., del Pozo Cruz, B., & Lonsdale, C. (2020). Validity and reliability evidence for motor competence assessments in children and adolescents: A systematic review. *Journal of sports sciences*, 38(15), 1717-1798.
- Jäger, R. S. (2009). Lehrerprofessionalität - Bedingungen, Genese, Wirkungen und ihre Messung. In O. Zlatkin-Troitschanskaia, K. Beck, D. Sembill, R. Nikolaus, & R. Melder (Eds.), *Lehrerprofessionalität - Bedingungen, Genese, Wirkungen und ihre Messung* (pp. 105-116). Beltz.
- Klingberg, B., Schranz, N., Barnett, L. M., Booth, V., & Ferrar, K. (2018). The feasibility of a fundamental movement skill assessment for pre-school aged children. *Journal of sports sciences*, 37(4), 378-386. <https://doi.org/10.1080/02640414.2018.1504603>
- Knudson, D. V. (2013). *Qualitative diagnosis of human movement: Improving performance in sport and exercise*. Human Kinetics.
- Kuckartz, U. (2018). *Qualitative Inhaltsanalyse. Methoden, Praxis, Computerunterstützung* (4. Aufl. ed.). Beltz.
- Lander, N. J., Hanna, L., Brown, H., Telford, A., Morgan, P. J., Salmon, J., & Barnett, L. M. (2017). Physical education teachers' perspectives and experiences when teaching FMS to early adolescent girls. *Journal of Teaching in Physical Education*, 36(1), 113-118. <https://doi.org/10.1123/jtpe.2015-0201>
- Lander, N. J., Morgan, P. J., Salmon, J., & Barnett, L. M. (2015). Teachers' perceptions of a fundamental movement skill (FMS) assessment battery in a school setting. *Measurement*

- in *Physical Education and Exercise Science*, 20(1), 50-62. <https://doi.org/https://doi.org/10.1080/1091367X.2015.1095758>
- Langendorfer, S. J., & Bruya, L. D. (1995). *Aquatic readiness - developing water competence in young children*. Human Kinetics.
- Leuders, T., Loibl, K., & Dörfler, T. (2020). Diagnostische Urteile von Lehrkräften erklären - Ein Rahmenmodell für kognitive Modellierungen und deren experimentelle Förderung. *Unterrichtswissenschaft*, 48, 493-502.
- Lounsbury, M., & Coker, C. (2008). Developing skill-analysis competency in physical education. *Quest*, 60(2), 255-267. <https://doi.org/https://doi.org/10.1080/00336297.2008.10483580>
- Mekkaoui, L., Schnitzler, C., Sidney, M., Gandrieau, J., Camporelli, F., & Potdevin, F. (2022). Building the foundation of aquatic literacy in 4–6 years-old children: A systematic review of good pedagogical practices for children and parents. *International Journal of Environmental Research and Public Health*, 19(10), 6180.
- Mekkaoui, L., Vogt, T., Camporelli, F., Costa, A. M., de Martelaer, K., Garrido, N. D., Ljodal, I., Olstad, B. H., Rejman, M., & Rudnik, D. (2025). Improving water safety and long-term engagement in aquatic physical activity: The aquatic literacy framework. *Journal of Teaching in Physical Education*.
- Mertens, L., De Martelaer, K., Saakslähti, A., & D'Hondt, E. (2021). The Inter-Rater and Intra-Rater Reliability of the Actual Aquatic Skills Test (AASST) for Assessing Young Children's Motor Competence in the Water. *Int. J. Environ. Res. Public Health*, 19(1). <https://doi.org/10.3390/ijerph19010446>
- Molich, R. (2010). A critique of »how to specify the participant group size for usability studies: A practitioner's guide« by Macefield. *Journal of Usability Studies*, 5(3), 124-128. <https://doi.org/https://dl.acm.org/doi/abs/10.5555/2835434.2835439>
- Moreno Murcia, J. A., de Paula Borges, L., & Huescar Hernandez, E. (2020). Design and Validation of the Scale to Measure Aquatic Competence in Children (SMACC). *Int J Environ Res Public Health*, 17(17). <https://doi.org/10.3390/ijerph17176188>
- Niederkofler, B., Herrmann, C., & Amesberger, G. (2018). Diagnosekompetenz von Sportlehrkräften - Semiformelle Diagnose von motorischen Basiskompetenzen. *Zeitschrift für sportpädagogische Forschung*, 2018(2), 72-96.
- Nitko, A. J. (2004). *Educational assessment of students* (4 ed.). Pearson. <https://doi.org/https://doi.org/10.25656/01:11487>
- Rheker, U. (2011). *Alle ins Wasser: Spiel und Spaß für Anfänger* (3. Aufl. ed., Vol. 1). Meyer & Meyer Verlag.
- Schilling, L., & Leiss, D. (2022). Förderung von Diagnose- und Interventionskompetenzen mithilfe von Praxisbezügen - Konzeption eines Seminars für die erste Phase der Lehrkräfteausbildung. In T. Ehmke, S. Fischer-Schöneborn, K. Reusser, D. Leiss, T. Schmidt, & S. Weinhold (Eds.), *Innovationen in Theorie-Praxis-Netzwerken - Beiträge zur Weiterentwicklung der Lehrkräftebildung* (Vol. 1, pp. 159-183). Dominik Leiss, Torben Schmidt, Swantje Weinhold.
- Schweihofen, C. (2013). Motorische Diagnostik als pädagogische Diagnostik? *sportunterricht*, 62(3), 85-89.
- Seyda, M., & Langer, A. (2020). Die Orchestrierung des diagnostischen und didaktischen Handels im Sportunterricht. Eine Frage der Kompetenz! *sportunterricht*, 69(8), 353-357.
- Sousa-Sá, E., Lander, N., Alqumsan, A. A., Alsanwy, S., Nahavandi, D., Toomey, N., Mohamed, S., Lewis, S., & Barnett, L. M. (2024). Physical Education Teachers' Perceptions of a Motor Competence Assessment Digital App. *Journal of Teaching in Physical Education*, 43(2), 276-291. <https://doi.org/10.1123/jtpe.2022-0222>
- Sousa-Sá, E., Lander, N. J., Alqumsan, A. A., Alsany, S., Nahavandi, D., Toomey, N., Mohamed, S., Lewis, S., & Barnett, L. M. (2023). Physical education teachers' perceptions of a motor competence assessment digital app. *Journal of Teaching in Physical Education*, 43(2), 276-291. <https://doi.org/https://doi.org/10.1123/jtpe.2022-0222>
- Staub, I., & Fokken, I. (2020). Vermittlungsinhalte einer umfassenden schwimmerischen Grundbildung. In T. Vogt (Ed.), *Vermittlungskompetenz in Sport, Spiel und Bewegung* (pp. 129–148). Meyer & Meyer.

- Sundan, J., Haga, M., & Lorås, H. (2023). Development and content validation of the swimming competence assessment scale (SCAS): A modified Delphi study. *Perceptual and motor Skills*, 130(4), 1762-1780.
- 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.
- van Ophuysen, S., Lintorf, K., & Harazd, B. (2013). Zur Qualität professioneller pädagogischer Diagnostik im Schulalltag. Forschungsbefunde und -desiderate. In K. Schwippert, M. Bonsen, & N. Berkeneyer (Eds.), *Schul- und Bildungsforschung. Diskussionen, Befunde und Perspektiven. Festschrift für Wilfried Bos*. Waxmann.
- Vogt, T., & Staub, I. (2020). Assessment of basic aquatic skills in children: inter-rater reliability of coaches, teachers, students and parents. *Journal of Physical Education and Sport*, 20(2), 577-583. <https://doi.org/0.7752/jpes.2020.02085>
- Ward, P., Shiri, A., Fatih, D., Peter, I., Insook, K., & and Li, W. (2021). Skill analysis for teachers: considerations for physical education teacher education. *Journal of Physical Education, Recreation & Dance*, 92(2), 15-21. <https://doi.org/10.1080/07303084.2020.1853635>
- Wilke, K., & Daniel, K. (2007). *Schwimmen: Lernen - Üben - Trainieren [Swimming: Learn - Practice - Train]* (6. ed.). Limpert.
- Wilson, W. J., Brian, A., & Kelly, L. E. (2020). The effects of online motor skill assessment training on assessment competence of physical educators. *Journal of Motor Learning and Development*, 9(1), 1-13. <https://doi.org/https://doi.org/10.1123/jmld.2020-0011>

Appendix 1: Summary Study 1

How do teachers analyse learners' swimming skills?

Purpose	Existing research evidence indicates that PE teachers face challenges in analysing their learners' motor skills in order to facilitate a purposeful learning process (Ward et al., 2021). Lounsbury and Coker (2008) note that the research that addressed the question of how competent PE teachers are in analysing the motor skills of their learners is disheartening. However, as there are no studies that explicitly refer to swimming lessons, this study aimed to find out how teachers analyse their learners' swimming skills.
Methodology	A quantitative online survey was conducted among 551 German primary and lower secondary school teachers (81.3% female, 45.7 ± 9.6 years old). A pre-study with ten PE teachers (80.0% female, 45.7 ± 18.6 years old) was done in the form of qualitative interviews in order to generate statements for the subsequent quantitative online survey. The online questionnaire comprised nine item-batteries with a total of 73 statements to comment on using a 5-point scale. The data was collected and analysed in four categories: (1) Diagnostic information gathering in swimming lessons, (2) Teachers' experiences with diagnostics in physical education, (3) Teachers' attitude towards the use of motor tests and (4) Teachers' subjective competence evaluation.
Findings	It was found that teachers used different strategies to analyse learners' swimming skills, ranging from simple observation of movement to structured but still informal testing procedures. Almost all teachers observed their learners while playing and moving freely in the water (92.9%) or while swimming without specific instruction (85.2%). Approximately half of the study participants (50.3%) do not use set criteria in the diagnostic process. For the majority of teachers, it is particularly important to divide the learner group into swimmers and non-swimmers. A large

proportion of teachers (69.9%) reported using their own monitoring sheet or checklist to record the data collected. Only about 5% of the study participants reported using standardised motor tests to determine their learners' swimming skills. In addition, it was found that only few teachers were able to draw on knowledge from their academic studies or traineeship for their diagnostic work. 54.7% of the teachers expressed a desire for a scientifically based tool to assist them in analysing their learners' swimming skills (Fokken et al., 2023).

Appendix 2: Summary Study 2

How do teachers evaluate a diagnostic procedure for swimming skill analysis?

Purpose The findings of Study I confirmed the authors' observations and suspicions that teachers lack a tool with which to analyse their learners' swimming skills. Existing tests do not seem to be suitable for use in (German) schools, so many teachers rely on their own designs or make unstructured observations. Therefore, the authors have developed a user-friendly diagnostic procedure for swimming skill analysis, which is called the 'Owl's Eye'. The procedure is based on the Assessment of Basic Aquatic Skills (ABAS, Vogt & Staub, 2020) and uses tasks from the ABAS task battery that have been tested for inter-rater reliability. The special quality of this procedure is that it is linked to a didactic framework for a comprehensive swimming education (Staub & Fokken, 2020). This enables teachers to use the framework to directly evaluate the results and determine the next learning objectives. The 'Owl's Eye' is structured in two steps. Initially, teachers classify their learners by conducting a few mandatory tasks to get a first insight into their learning group. The central objective of the classification is to determine whether the learners can already fully submerge and float on front and back. The initial classification is supplemented by a combined task in deep water that includes jumping into the water, submerging twice, treading water, floating on the back, and swimming on front and back for at least ten meters. In a second step, additional tasks may be administered to gain more detailed information. The two-step procedure permits the efficient analysis of the students' swimming skills, even in larger learning groups in up to 25 and more (Fokken et al., 2022). This study aims to have teachers evaluate the feasibility and subjective value of the procedure and to determine the extent to which the procedure is suitable for supporting and improving diagnostics.

Methodology German PE teachers were asked to implement the 'Owl's Eye' into at least one of their swimming lessons at the beginning of the school year for diagnostic purposes and to evaluate it afterwards. As was the case in the first study, a pre-study involving 14 PE teachers (64.3% female, 43.2 ± 11.7 years old) was conducted in the form of qualitative interviews in order to generate statements for the subsequent quantitative survey. 58 German teachers of primary and lower secondary school (81.0% female, 45.4 ± 10.1 years old) participated in the main study and completed the online questionnaire after implementation. All participants were provided with the required information for conducting the diagnostic procedure (e. g. explanations of the two-step procedure via texts and FAQ videos, suggestions for a child-friendly implementation, documentation sheets paper-pencil). After implementation, teachers were invited to indicate which of the materials and information provided they had used for preparation and during implementation. The options for their responses were

categorised as 'yes', 'partly' or 'no'. Teachers were then asked to provide a rating for the materials on a scale of 1 to 5, with 1 indicating that the material was 'not at all helpful/useful' and 5 indicating that the material was 'very helpful/useful'. In addition, a battery of 56 statements was given, which required a response on a five-point scale. The scale ranged from 1 'not correct' to 5 'correct'. The statements were used to request the subjective evaluation regarding the feasibility and value of the Owl's Eye in general and to gain insight into teachers' personal assessment of its use.

Findings

The evaluation of the survey has demonstrated that the 'Owl's Eye' procedure is, in general, well-received. The findings to date suggest that the procedure has the potential to serve as a valuable tool for improving the accuracy and effectiveness of swimming skill analysis, and to become useful for analysing swimming skills in practice. The procedure was evaluated by the majority of teachers (89.3%) as a valuable diagnostic tool for structuring the diagnostic process. Furthermore, teachers noted that the procedure enabled them to more systematically analyse their observations and to reach more reliable conclusions. About three quarters (73.2%) of teachers agreed 'correct' or 'rather correct' with the statement, that using the Owl's Eye materials saved them work in preparing the swimming skill analysis (Fokken et al., unpublished findings currently under review). In this context, a number of respondents expressed the need for a digital record of test results to be available. The findings suggested that the Owl's Eye should be further developed and that a transformation into an app should be pursued.

Diagnostische Kompetenzen von Sportlehramtsstudierenden durch authentische Situationen professionell entwickeln

von Anneke Langer & Nele Schlapkohl

Zusammenfassung: Die Förderung diagnostischer Kompetenzen ist Teil der professionellen Entwicklung von Studierenden und sollte möglichst authentisch und berufsnahe gestaltet sein. Es wurde ein Seminar »Bewegen im Wasser« konzipiert und durchgeführt, das die Umsetzung von Peer Assessments (formatives Assessment) und die Durchführung standardisierter Methoden für die Beurteilung schwimmerischer Grundfertigkeiten von Kindern (summatives Assessment) integriert. Ziel des Beitrags ist es, erlebte Potenziale und Herausforderungen aus Sicht von Sportlehramtsstudierenden für ihre professionelle Entwicklung diagnostischer Kompetenzen durch die Teilnahme an diesem Seminar zu identifizieren. Die Ergebnisse zeigen u. a., dass im Peer Assessment der Einsatz von Videoanalysen als Potenzial für die Verbesserung der eigenen Wahrnehmung erfahren wird. Als herausfordernd erleben die Studierenden, ein präzises Feedback an erfahrene Schwimmer:innen zu formulieren. Im summativen Assessment empfinden sie u. a. als Potenzial, die schwimmerischen Grundfertigkeiten von Kindern einzustufen. Als Herausforderung erleben sie bei der Einstufung der Kinder, Diskrepanzen in den Diagnosen zwischen Studierenden. Die Ergebnisse geben Hinweise, wie diagnostische Kompetenzen von Sportlehramtsstudierenden gefördert werden können.

Schlagwörter: Diagnostische Kompetenzen, Hochschulausbildung von Sportlehrkräften, Lehrerprofessionalisierung, Peer Assessment, Summatives Assessment

Professional development of diagnostic competencies from pre-service PE teachers through authentic situations

Abstract: The promotion of diagnostic competencies is part of students' professional development and should be authentic and work-related. A seminar on movement in water was designed and conducted which integrated the implementation of peer assessment (formative assessment) and the use of standardized methods for assessing children's basic swimming skills (summative assessment). The objective of this article is identifying the potential and challenges experienced by pre-service PE teachers in terms of their professional development of diagnostic competencies through participation in this seminar. The results show, among other things, that they experience in peer assessment the use of video analysis as a potential to improve one's own perception. They are challenged by formulating precise feedback for experienced swimmer. In summative assessment, they see potential in classify children's basic swimming skills. They perceive a challenge due to discrepancies in their classifications of children between pre-service PE teachers. The results indicate how diagnostic competencies can be promoted in physical education teacher education.

Keywords: diagnostic competencies, physical education teacher education, professionalization of teachers, peer assessment, summative assessment

Diagnostische Kompetenzen von Sportlehramtsstudierenden durch authentische Situationen professionell entwickeln

1 Einführung

Die Gestaltung eines adaptiven und damit lernförderlichen Sportunterrichts für heterogene Lerngruppen stellt eine berufliche Anforderung an Sportlehrkräften dar. Für die Bewältigung dieser Anforderung benötigen Sportlehrkräfte professionelle Kompetenzen, wie z. B. diagnostische und didaktische Kompetenzen (Baumgartner, 2022a; Schrader, 2009). Der Prozess, in dem Informationen über heterogene Lerngruppen gesammelt und verarbeitet werden, um bildungsbezogene Entscheidungen treffen zu können, wird als pädagogische Diagnose bezeichnet. Bei der Diagnose sind fachspezifische Besonderheiten zu berücksichtigen (Latzko & Gottlebe, 2022). Diese erfolgt im Rahmen der vorliegenden Studie im Kontext des Bewegungsfelds »Bewegen im Wasser«. Sportlehrkräfte können beispielsweise durch eine Ergebnis- oder Leistungsdiagnostik (summatives Assessment) Einblicke in die Ergebnisse der Lernprozesse ihrer Schüler:innen erhalten und auf dieser Grundlage Entscheidungen z. B. über die Bewegungsdemonstration der Kraultechnik treffen. Das summative Assessment liefert jedoch kaum umfassende Informationen, die für die Unterstützung des weiteren Lernens von Bedeutung sind (Black & Wiliam, 2010). Dies ist darauf zurückzuführen, dass es zu einem definierten Zeitpunkt durchgeführt wird, der meist am Ende einer Lernsequenz liegt (Danthony et al., 2019). Prozessdiagnosen (formatives Assessment) stellen eine stärker integrierte und kontinuierliche Bewertungsmethode dar. Sie zielen darauf ab, den Lernprozess der Schüler:innen zu unterstützen und ermöglichen gleichzeitig den Sportlehrkräften Einblicke in die Wirksamkeit ihres Unterrichts (Wiliam, 2011). Das formative Assessment in Anlehnung an William (2011) ist dadurch gekennzeichnet, dass Lernende aktiv in den Beurteilungsprozess einbezogen werden. Die Durchführung kann beispielsweise durch Selbsteinschätzung oder Peer Assessment erfolgen, d. h. Lernende beobachten und beurteilen sich gegenseitig bei gezielten Antriebsversuchen in widerstandssarmer Wasserlage. Weitere Kennzeichen sind die Klärung von Lernabsichten und die Erteilung von Feedback, die die Lernenden in ihrer Entwicklung unterstützen (William, 2011). Formatives Assessment beinhaltet auch, dass die Lernentwicklung systematisch und mehrfach durch standardisierte und messbare Methoden beurteilt wird (Schildkamp et al., 2020). Sowohl summative als auch formative Assessments haben in der Sportunterrichtspraxis wichtige Funktionen, um u. a. das Lernen von Schüler:innen beurteilen zu können (Kibble, 2017). Allerdings weisen bisherige Befunde zur Beurteilungspraxis von Sportlehrkräften darauf hin, dass sie selten evidenzbasierte Beurteilungsmethoden (summatives Assessment) nutzen, sondern eigene Beurteilungsmethoden entwickeln (Snow et al., 2025). Letztere sind anfällig für Beurteilungsverzerrungen (Schittkowski et al., 2024). Zudem zeigen Studien, dass Sportlehrkräfte selten formatives Assessment durchführen (Tolgfors et al., 2021;

Borghouts et al., 2017, Pastore & Andrade, 2019). Als Gründe hierfür werden genannt, dass Sportlehrkräften Kenntnisse, Erfahrungen und Zeit zur Umsetzung formativen Assessments fehlen und sie formative Beurteilungen von motorischen Kompetenzen als umständlich wahrnehmen (Langer, 2023; Slingerland et al., 2024). Das summative und formative Assessment setzt bei den Sportlehrkräften Wissen, Fähigkeiten und Einstellungen voraus, um einen kohärenten und angemessenen Ansatz für Diagnosen im Schulkontext implementieren zu können. Sportlehrkräfte sollten daher adäquat gefördert werden, um summative und formative Assessments im Sportunterricht durchführen zu können. Im Rahmen der Sportlehrkräfteprofessionalisierung sollte dies bereits in der Hochschulausbildung stattfinden (Grossmann, 2021; Schlapkohl & Langer, 2024). Hierfür sind theoretische und praktische Lerngelegenheiten für Studierenden zu schaffen, die möglichst realitätsnah an die berufliche Praxis von Lehrkräften anknüpfen (Chernikova et al., 2020; Ward et al., 2022). Daran setzt die vorliegende Studie an, in der untersucht wird, welches Potenzial und welche Herausforderungen Sportlehramtsstudierende für ihre professionelle Entwicklung erleben, wenn sie im Seminar »Bewegen im Wasser« erstens ein Peer Assessment (formatives Assessment) und zweitens ein summatives Assessment bei Schüler:innen umsetzen. Die daraus generierten Erkenntnisse dienen dazu, einen Transfer des vermittelten Wissens über summatives und formatives Assessment in eine authentische Anwendung innerhalb der Sportlehrkräftebildung zu ermöglichen (Baumgartner, 2022a; Chernikova et al., 2020) und Informationen zu erhalten, wie diagnostische Kompetenzen von Sportlehramtsstudierenden entwickelt werden können.

2 Theoretischer Hintergrund

2.1 (Authentische) Diagnosesituationen in der Sportlehrkräftebildung im Bewegungsfeld »Bewegen im Wasser«

Diagnosesituationen in der Sportlehrkräftebildung im Bewegungsfeld »Bewegen im Wasser« können in formatives und summatives Assessment unterschieden werden. Die Adressierung dieser Diagnosesituationen in der Sportlehrkräftebildung im Bewegungsfeld »Bewegen im Wasser« hat das Potenzial für Studierende, eine »handlungs- und entscheidungsdruckentlastete Auseinandersetzung« (Kramer, 2020, S. 284) mit dem späteren beruflichen Handlungsfeld und mit sich selbst bereitzustellen. Unter dem Begriff Potenzial wird verstanden, dass Sportlehramtsstudierende eine Möglichkeit zur Verbesserung in der Umsetzung von Diagnosesituationen durch eine Orientierung an Nützlichkeit und Verwertbarkeit für die berufliche pädagogische Praxis erhalten (Kramer, 2020).

Mit Blick auf das formative Assessment stellt das Peer Assessment eine Schlüsselstrategie dar, welches bereits vielfältig in der Sportlehrkräftebildung eingesetzt wird (Eather et al., 2017; Lamb et al., 2013). Durch das Peer Assessment kann der Lernprozess sowohl für die beobachtende Person als auch für die beobachtete Person

verbessert werden (Eather et al., 2017; Lamb et al., 2013). Für Sportlehramtsstudierende bietet es eine effektive Möglichkeit, die eigenen Handlungskompetenzen zu erweitern. Dafür sind beobachtbare bzw. messbare Lernziele zu identifizieren (Schildkamp et al., 2020). Beispielsweise sollen Studierende lernen, Gleichaltrige bei der Einnahme einer widerstandsfähigen Wasserlage zu beobachten, zu beurteilen und Rückmeldungen zu geben. Unterstützend bekommen sie u. a. auch zur eigenen Stabilisierung der Wasserlage eine Bewegungsrückmeldung durch den Einsatz von Videoaufnahmen. Als angehende Sportlehrkräfte sollen sie dabei auch lernen, ein effektives Feedback hinsichtlich des motorischen Lernens zu geben. Das bedeutet, einen (positiven) spezifischen und präzisen Informationsgehalt mitzuteilen, durch den Lernende »die Dissonanz zwischen Ist- und Sollzustand hinsichtlich eines gesetzten Zieles verringern bzw. korrigieren können« (Baumgartner, 2017, S. 249).

Mit Blick auf ein summatives Assessment in der Sportlehrkräftebildung im Bewegungsfeld »Bewegen im Wasser« sollten Sportlehramtsstudierende u. a. lernen, wie sie die Schwimmfähigkeit von Schüler:innen beurteilen können. Verfahren zur Überprüfung der Schwimmfähigkeit im schulischen Schwimmunterricht knüpfen häufig an das Messen von Strecken, Zeiten und Techniken an (Varveri et al., 2016). Authentische summative Diagnosesituationen sind somit häufig dadurch gekennzeichnet, dass Schüler:innen anhand einer bestimmten Schwimmstrecke beurteilt werden. Bei der Beurteilung des schwimmerischen Könnens und des aktuellen Lernstands sollten jedoch weitere Facetten berücksichtigt werden, wie beispielsweise die Überprüfung der schwimmerischen Grundfertigkeit (u. a. das Untertauchen, die Atemkontrolle oder das Gleiten). Dafür können evidenzbasierte Diagnosemethoden im Bewegungsfeld »Bewegen im Wasser« eingesetzt werden. Für deren Einsatz bedarf es einer Schulung sowie Erprobungsmöglichkeiten. Beispielsweise misst der Test Assessment of basic aquatic skills anhand von 19 Bewegungsaufgaben die schwimmerischen Grundfertigkeiten (ABAS; Vogt & Staub, 2020). Auch das daraus entstandene, gekürzte und validierte Beobachtungsverfahren »die Eule & das Schwimmen« (Fokken et al., 2022) orientiert sich an den schwimmerischen Grundfertigkeiten von Schüler:innen und wurde für den schulischen Schwimmunterricht entwickelt.

Durch die Durchführung des formativen und summativen Assessments im Rahmen der Sportlehrkräftebildung im Bewegungsfeld »Bewegen im Wasser« können Studierende (fachdidaktisches) Wissen und Können entwickeln, mit dem sie bestärkt werden, ihre diagnostischen Kompetenzen auch in ihrer späteren Unterrichtspraxis anzuwenden (Tolgfors et al., 2021).

2.2 Professionelle Entwicklung diagnostischer Kompetenzen

Unter professioneller Entwicklung wird ein Prozess verstanden, durch den eine Person »die für effektive professionelle Praxis notwendigen Kenntnisse und Fähigkeiten

erwirbt oder verbessert« (Hoyle, 1991, S. 135), um in einer spezifischen beruflichen Situation professionell handeln zu können (z. B. guten, lernförderlichen Schwimmunterricht zu gestalten). Dafür benötigt die Person fundiertes Wissen. Auf dieses sogenannte Professionswissen kann die Person in konkreten Situationen systematisch zurückgreifen (Hoyle, 1991). In Anlehnung an die kompetenzorientierte Sportlehrkräftebildungsforschung (Baumgartner, 2022a) ist das Professionswissen kompetenzbereichsbezogen.

Das Kompe-Sport-Modell (Baumgartner, 2022a) integriert unterschiedliche Auffassungen von Kompetenzen in einem Typologie- und Topologiemodell professioneller Kompetenz(en) von Sportlehrkräften. Mit professionellen Kompetenzen (pluralisierte Form) werden die Summe der kriteriengeleiteten can-do-statements (Terhart, 2007) beschrieben, die Sportlehrkräfte zur Erfüllung von beruflichen realen Anforderungssituationen benötigen (Baumgartner, 2017). Unter anderem diagnostizieren Sportlehrkräfte Bewegungen im Wasser, um die Entwicklung der Kompetenzfacetten von Schüler:innen zu fördern (z. B. die Performanz von Schüler:innen zu fördern, wie die Nutzung gezielter Antriebsformen über eine längere Strecke) (Abb. 1). »Professionelle Kompetenzen stellen eine – durchaus nicht die einzige – Richtschnur der Aus- und Weiterbildung von Sportlehrkräften dar« (Baumgartner, 2022b, S. 37).

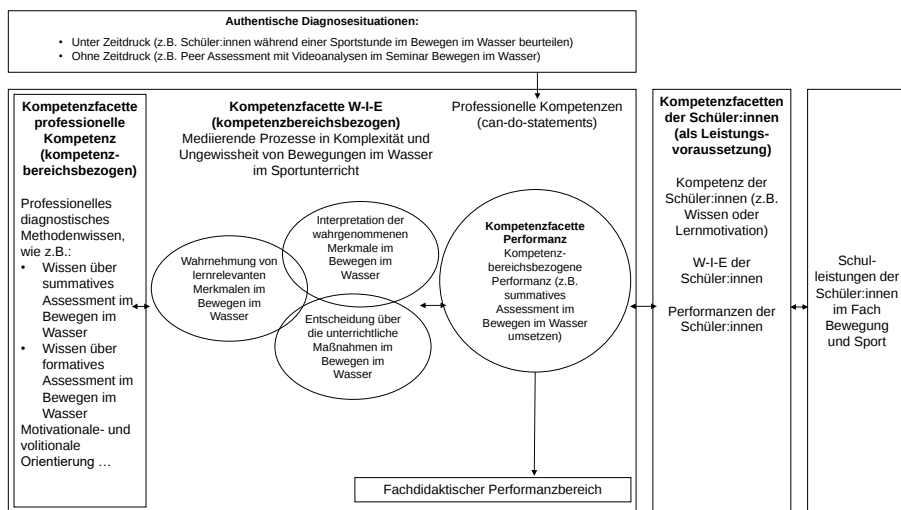


Abb. 1 Diagnostische Kompetenzen als Kompetenzbereich des Kompe-Sport-Modells (adaptiert nach Baumgartner, 2022a, S. 42; Blömeke et al., 2015).

Die *Kompetenzfacette professionelle Kompetenz* setzt sich aus latenten Aspekten zusammen (z. B. professionelles Wissen, motivationale und volitionale Orientierung; Baumert & Kunter, 2006). Professionelles Wissen über diagnostische Kompetenzen bezieht sich u. a. auf diagnostisches Methodenwissen über summative und formative

Beurteilungsformen für Bewegungen im Wasser. Motivationale und volitionale Orientierung umfasst die Absicht von Sportlehrkräften (Terhart, 2007), diagnostische Methoden auch im Schwimmunterricht einsetzen zu wollen. Die *Kompetenzfacette W-I-E* umfasst die Wahrnehmungs-, Interpretations- und Entscheidungsprozesse, die der Diagnose von Bewegungen im Wasser zugrunde liegen. Bei Bewegungen im Wasser werden von einer Sportlehrkraft bspw. Kopf- und Körperhaltungen von Schüler:innen wahrgenommen und ihre Bedeutung für den Vortrieb unter Rückbezug des professionellen Wissens sowie mit Bezug zu Informationen über die Schüler:innen interpretiert. Auf dieser Basis können Sportlehrkräfte Entscheidungen hinsichtlich unterrichtlicher Maßnahmen treffen, wie z. B. die Durchführung von Kontrastübungen. Bei diesen mediiierenden diagnostischen Prozessen von Bewegungen im Wasser können Beobachtungs- und Beurteilungsverzerrungen auftreten (z. B. Reihungseffekt¹ oder Überbewertung von quantifizierbaren Ergebnissen²) (Schittkowski et al., 2024). Die Reduktion dieser Verzerrungen ist wichtig, um akkurate und auch faire Beurteilungen und Bewertungen vornehmen zu können. Eine Reduktion ist u. a. möglich, indem Sportlehrkräfte Wissen über Beurteilungsverzerrungen besitzen und ihre eigenen Diagnosen dahingehend reflektieren (Langer, 2023).

Die *Kompetenzfacette Performanz* beschreibt das beobachtbare Verhalten einer Sportlehrkraft in einer beruflichen Anforderungssituation (Baumgartner, 2022b), beispielsweise, dass ein summatives oder formatives Assessment für das Bewegen im Wasser umgesetzt wird. Dafür können unterschiedliche methodische Zugangsweisen gewählt werden, die von einem Einsatz standardisierter und methodisch kontrollierter diagnostischer Verfahren (z. B. dem Fertigkeitstest ABAS von Vogt & Staub, 2020) bis hin zu subjektiven Beobachtungen von Bewegungen im Wasser und intuitiven Einschätzungen reichen. Letztere sind anfälliger für Beurteilungsfehler, aber schnell in der Schulpraxis umsetzbar (Langer, 2023).

Dem heuristischen Modell folgend kann angenommen werden, dass eine Verbesserung der Performanz einer Sportlehrkraft durch eine professionelle Entwicklung aller drei Kompetenzfacetten und deren Zusammenspiel erreicht werden kann (Baumgartner, 2022a): (1) die qualitative Verbesserung von Aspekten der professionellen Kompetenz (z. B. diagnostisches Methodenwissen über Bewegen im Wasser), (2) die Verbesserung der situationsspezifischen Kompetenzfacette W-I-E (z. B. kriteriengeleitete Beobachtungen von Bewegungen im Wasser) und (3) durch den Einsatz diagnostischer Methoden (z. B. summative und formativer Assessments) in der eigenen Unterrichtspraxis im Bewegen im Wasser (Performanz).

- 1 Der Reihungseffekt bezieht »sich auf die Verzerrung der Bewertung, die durch die Reihenfolge beeinflusst wird, in der Informationen präsentiert werden. Er tritt auf, wenn die Position einer Person in einer Reihe von Bewertungen die Beurteilung dieser Person beeinflusst« (Schittkowski et al., 2024, S. 557).
- 2 Die Überbewertung von quantifizierbaren Ergebnissen: Durch leicht messbare Ergebnisse (wie Schwimmzeiten) »könnten qualitative Aspekte wie Teamarbeit, Anstrengung und technische Fähigkeiten unterbewertet werden« (Schittkowski et al., 2024, S. 557).

3 Forschungsstand

Mit Blick auf empirische Befunde zur Förderung diagnostischer Kompetenzen in der Hochschulausbildung im Fach Sport (i. S. der Umsetzung formativer und summativer Assessments) zeigt sich, dass in den letzten Jahren vermehrt Studien zur Förderung von Sportlehramtsstudierenden in der Umsetzung formativer Assessments sowohl im Rahmen von Seminaren als auch im Sportunterricht (z. B. im Rahmen von Schulpraktika) durchgeführt wurden (Herrero-González et al., 2023). Hingegen gibt es kaum Studien, die die Förderung von Sportlehramtsstudierenden in der Umsetzung summativer Assessments untersuchen.

Einen Grund für die Umsetzung von Fördermaßnahmen zum formativen Assessment in der Hochschulausbildung stellen Erkenntnisse empirischer Studien dar. Sie zeigen, dass vielen Sportlehramtsstudierenden Kenntnisse und Fähigkeiten fehlen, um formatives Assessment erfolgreich in ihre Unterrichtspraxis (z. B. während Schulpraktika) zu integrieren (Herrero-González et al., 2023; Pastore & Andrade, 2019; Schildkamp et al., 2020). Die Ergebnisse von Herrero-González et al. (2023) und Starck et al. (2023) haben gezeigt, dass Erfahrungen mit formativen Assessments im Rahmen der Hochschulausbildung Sportlehramtsstudierenden geholfen haben, ähnliche diagnostische Methoden während ihrer Unterrichtspraxis in der Schule anzuwenden. Eine Umsetzung ist vor allem dann wahrscheinlich, wenn Studierende den Einsatz formativer Assessments als Teil ihrer professionellen Entwicklung ansehen. Andererseits äußern Sportlehramtsstudierende in einer Studie von Tolgfors et al. (2021) zur Umsetzung formativen Assessments in Schulpraktika, dass sie einige Widerstände und Schwierigkeiten erlebt haben. Beispielsweise können die bisherige Beurteilungskultur im Sportunterricht (vorrangig Benotungen als Teil summativen Assessments), aber auch fehlende Kenntnisse über Schüler:innen den Einsatz formativen Assessments in Schulpraktika einschränken (Tolgfors et al., 2021).

Mit Blick auf die Umsetzung und Integration des Peer Assessments in der Sportlehrkräftebildung existieren nur wenige Studien. Eather et al. (2017) haben die wahrgenommenen Lehrkompetenzen von Sportlehramtsstudierenden während eines vierwöchigen Schulpraktikums mit einer mixed-method-Untersuchung erhoben. Die Ergebnisse zeigen, dass die Sportlehramtsstudierenden die unterschiedlichen Perspektiven bei der Beurteilung von ihren Kommilitonen schätzten und respektierten. Zudem kann das Geben von Feedback im Rahmen des Peer Assessments die wahrgenommene Lehrkompetenzen von Sportlehramtsstudierenden verbessern.

Mit Blick auf die Förderung von Sportlehramtsstudierenden in der Umsetzung summativer Assessments sind keine Studien bekannt. Für die Umsetzung hingegen hat eine qualitative Studie von Snow et al. (2025) ergeben, dass Sportlehrkräfte selten evidenzbasierte Bewertungsinstrumente verwendeten, sondern oft ihre eigenen entwickelten. Zu den wahrgenommenen Hindernissen gehörten u. a. unzureichende Unterstützung an den Schulen. Zudem scheinen Sportlehrkräfte nicht aus-

reichend auf die Diagnose motorischer Kompetenzen vorbereitet zu sein. Zu den fördernden Faktoren, evidenzbasierte Bewertungsinstrumente einzusetzen, zählen u. a. unterstützende Materialien und Ressourcen an der Schule. Mit Blick auf das Bewegungsfeld »Bewegen im Wasser« scheint es bislang kaum Studien zur Entwicklung diagnostischer Kompetenzen von Sportlehramtsstudierenden zu geben. Allerdings deuten die wenigen empirischen Befunde zum diagnostischen Vorgehen von Sportlehrkräften in diesem Bewegungsfeld auf Unsicherheiten bei der Umsetzung summativer Assessments hin (Fokken et al., 2023). Ein Grund scheint darin zu liegen, dass eine qualitativ anspruchsvolle Analyse des aktuellen Könnens der Kinder erforderlich ist, die beispielsweise über eine Einteilung in Schwimmer:innen und Nichtschwimmer:innen hinaus geht (Fokken et al., 2023).

4 Forschungsfragen der vorliegenden Untersuchung

Die bisherigen empirischen Befunde könnten darauf hindeuten, dass wenig über die Sichtweisen von Sportlehramtsstudierenden über formative und summative Assessments im Bewegungsfeld »Bewegen im Wasser« bekannt ist und sie wenig Kenntnisse und Erfahrungen über formative und summative Assessments haben. Deren Sichtweisen über formative und summative Assessments, scheinen jedoch eine Voraussetzung für die Anwendung in der Schulpraxis zu sein (Herrero-González et al., 2023; Snow et al., 2025). Eine Umsetzung ist u. a. wahrscheinlich, wenn sie formativen und summativen Assessments einen potenziellen Nutzen zuschreiben (z. B. um Beurteilungsverzerrungen zu verringern) (Borko et al., 2010). Daher ist es relevant, Kenntnisse zu erhalten, inwiefern Sportlehramtsstudierende formative und summative Assessments als Teil ihrer professionellen Entwicklung sehen, nachdem sie dieses selbst durchgeführt haben. Dieser Forderung wird mit der Beantwortung der folgenden Forschungsfragen nachgegangen:

Welche Potenziale und Herausforderungen erleben Sportlehramtsstudierende in einem Seminar im Bewegungsfeld »Bewegen im Wasser« ...

1. ... für ihre professionelle Entwicklung der Kompetenzfacette W-I-E durch Peer Assessment?
2. ... für ihre professionelle Entwicklung der Kompetenzfacette Performanz durch summatives Assessment von Schüler:innen?

5 Methodisches Vorgehen

Das methodische Vorgehen zur Untersuchung der Forschungsfragen umfasst eine qualitative Erhebung der Sichtweisen von Sportlehramtsstudierenden zur professionellen Entwicklung ihrer diagnostischen Kompetenzen durch ein Seminarkonzept im Bewegungsfeld »Bewegen im Wasser« mit authentischen Situationen. Das Seminar wurde im Frühjahrssemester 2024 im Master of Education Studiengang an einer

Hochschule in Schleswig-Holstein durchgeführt. Im direkten Anschluss fand die qualitative Befragung der teilnehmenden Sportlehramtsstudierenden statt.

5.1 Seminarkonzept mit authentischen Diagnosesituationen

Mit dem Seminarkonzept wird das Ziel verfolgt, diagnostische Kompetenzen von Sportlehramtsstudierenden (i. S. der Umsetzung summativer und formativer Assessments) durch authentische Situationen im Bewegungsfeld »Bewegen im Wasser« professionell zu entwickeln. Dafür bietet die Unterteilung in die drei Kompetenzfacetten des Kompe-Sport-Modells (Baumgartner, 2022b) verschiedene Ansatzpunkte, die mit unterschiedlichen, gezielten Vorgehensweisen im Seminar umgesetzt werden:

(1) Auf Ebene der *Kompetenzfacette professionelle Kompetenz* durch die Vermittlung von diagnostischem Methodenwissen zur Umsetzung formativen und summativen Assessments. Dieses Wissen wird in Theoriesitzungen am Anfang des Seminars vermittelt.

(2) Auf Ebene der *Kompetenzfacette W-I-E* durch die Erprobung eines Peer Assessment (formatives Assessment). Dafür wird im Seminar der Fünfer-Schritt des professionellen Diagnostizierens³ (Latzko & Gottlebe, 2022) umgesetzt (s. Tabelle 1). Der Fünfer-Schritt wurde für die Lehrkräftebildung mit dem Ziel entwickelt, eine »Orientierungsgrundlage für die inhaltliche Präzisierung und konkrete Umsetzung« (Latzko & Gottlebe, 2022, S. 49) curricularer Ausbildungsziele zu bieten. Mit der Durchführung des Fünfer-Schritts im Seminar »Bewegen im Wasser« können Sportlehramtsstudierenden ihre diagnostische Urteilsbildung strukturieren und pädagogische Entscheidungen in der Handlungspraxis begründet umsetzen. Der Fünfer-Schritt nach Latzko und Gottlebe (2022) lässt sich für die Entwicklung der *Kompetenzfacetten W-I-E* nutzen, da sowohl Latzko und Gottlebe (2022) als auch Baumgartner (2022a) sich auf dasselbe Kompetenzverständnis nach Weinert (2001) beziehen⁴. Demnach dient das Professionswissen, i. S. des diagnostischen Methodenwissens »zur kritischen Auseinandersetzung mit den methodischen Zugängen« (Latzko & Gottlebe, 2022, S. 48) und bildet eine Voraussetzung für die Wahrnehmung und Interpretation von Bewegungen im Wasser sowie eine Grundlage für

3 Der Fünfer-Schritt des professionellen Diagnostizierens (Latzko & Gottlebe, 2022) besteht aus (1) Theoretische Konzepte zur Kennzeichnung des Diagnoseanlasses aktivieren, (2) theoriebezogene Hypothesen für das Zustandekommen des Problems ableiten, (3) adäquate diagnostische Methoden und entsprechende Instrumente zur Hypothesenprüfung auswählen und einsetzen, (4) alle gesammelten und erhobenen Daten zu einem diagnostischen Urteil integrieren und (5) eine der Diagnose entsprechende Förderung durchführen, d. h., eine Anpassung des unterrichtlichen Settings vornehmen zu können.

4 Das Kompe-Sport-Modell ist eine Synthese zwischen dem Kompetenzmodell von Baumert und Kunter (2013) und Blömeke et al. (2015). Baumert und Kunter (2013) definieren Kompetenzen auch in Anlehnung an Weinert (2001) und ordnen Diagnostische Kompetenzen als eine Facette des Fachdidaktischen Wissens zu.

didaktische Entscheidungen. Darauf aufbauend beinhaltet das Seminarkonzept, dass die Sportlehramtsstudierende die Schwimmtechnik eines Peers aufnehmen und das Video kriteriengeleitet diagnostizieren. Die Videoanalyse ermöglicht den Studierenden meist ohne größeren Zeitdruck und unter Rückbezug ihres professionellen Wissens, Bewegungen im Wasser wahrnehmen und interpretieren zu können. Ohne den Einsatz von Videos, sind Bewegungen im Wasser unter (hohem) Zeitdruck zu analysieren. Dies kann eine Diagnose erschweren (u. a. Feth, 2014) und dazu führen, dass Fehler entstehen können (u. a. Schittkowski et al., 2024). Im Anschluss an die Videoanalyse führen sie eine der Diagnose entsprechende Förderung durch, d. h. sie wählen passende Übungen aus, um die Antriebswirkung der Schwimmtechnik von Kommiliton:in zu verbessern, führen diese gemeinsam durch und geben sich gegenseitig Rückmeldungen zum Lernprozess.

(3) Auf Ebene der *Kompetenzfacette Performanz* durch die Umsetzung eines summarischen Assessments für Schüler:innen. Dafür greift das Seminarkonzept den Zyklus⁵ von Helmke (2017) auf, da dieses Verfahren die Erfassung und Verbesserung der Diagnosefähigkeit (i. S. der Performanz) intendiert. Das Potenzial des Verfahrens liegt u. a. »im Anstoßen von Diskursen« (Helmke, 2017, S. 141) über diagnostische Ergebnisse, vor allem wenn sie von der Erwartung abweichen.

Ein solcher Diskurs wird im Rahmen des Seminars »Bewegen im Wasser« als gewinnbringend für die professionelle Entwicklung erachtet, um mit beruflichen Anforderungen umgehen zu lernen (Baumert & Kunter, 2013). Eine Verknüpfung der beiden Ansätze von Baumgartner (2022a) und Helmke (2017) ist durch deren vergleichbaren Kompetenzverständnisse in Anlehnung an Weinert (2001) bzw. Wahl et al., (2006) möglich. Demnach sind Kompetenzen vorhandene oder erlernbare, kognitive Fähigkeiten und Fertigkeiten, die in unterschiedlichen Situationen erfolgreich und verantwortungsvoll eingesetzt werden können, um berufliche Anforderungen zu begegnen (Weinert, 2001). Im Rahmen des Seminars sollen die Studierenden alle Stufen des Zyklus nach Helmke (2017) umsetzen (s. Tabelle 1). Dafür führen sie an einem Kennenlerntag mit der Schulklasse eine Lernstandsdiagnostik mit der gesamten Lerngruppe und anschließend mit einzelnen Schüler:innen im Seminar »Bewegen im Wasser« durch. Die Erkenntnisse sollen die Studierenden nutzen, um einen Schwimmunterricht für die Schulklasse zu planen, durchzuführen und zu reflektieren (für eine ausführliche Beschreibung siehe Schlapkohl & Langer, 2024). Dabei wird das Modell der Schwimmerischen Grundbildungen von Staub und Fokken (2020) berücksichtigt, das den Schwimmunterricht didaktisch in vier farbcoodierte Phasen strukturiert (rot, gelb, grün und blau), die schrittweise aufeinander

5 Der Zyklus nach Helmke (2017) umfasst A) Auswahl eines Schülermerkmals oder eines Satzes von Aufgaben, B) Erhebung der tatsächlichen Schülerleistung bzw. des Merkmals, C) Ihre (der Lehrkraft, Anm. der Autor:innen) persönliche Prognose, D) Vergleich zwischen Schätzung und empirischem Befund, E) Analyse von Diskrepanzen.

aufbauen. Für jede Lernphase werden Lernziele formuliert⁶, die den Lernprozess strukturieren und zugleich eine individuelle Weiterentwicklung in die nächste Phase ermöglichen sollen. Welches Kind sich in welcher Phase befindet, wird mit Hilfe der Lernstandsdiagnostik ermittelt (das genaue Vorgehen wird in Fokken et al., 2022 beschrieben).

5.2 Datenerhebung und -auswertung

Es wurden elf Masterlehramtsstudierende des Faches Sport (*MAlter* = 26 Jahre, *SD* = 3.74; fünf weiblich; sechs männlich; 2. Master-Fachsemester) leitfadengestützt interviewt, nachdem sie am Seminar »Bewegen im Wasser« teilgenommen haben. Der Leitfaden umfasste einerseits Fragen zur Sichtweise bei der Umsetzung des Peer Assessments, z. B. »Erzähl doch mal, wie war das für Dich eine:n Kommiliton:in im Bewegen im Wasser zu diagnostizieren?« sowie welche Potenziale und welche Herausforderungen dabei erlebt wurden. Auch wurde danach gefragt, wie mit den Herausforderungen umgegangen wurde und wie es empfunden wurde, Kommiliton:in Rückmeldungen zu geben und der Diagnose entsprechende Förderungen durchzuführen. Andererseits beinhaltete der Leitfaden Fragen zur Sichtweise bei der Umsetzung summativer Assessments mit Schüler:innen, z. B. »Wie war es für Dich, eine:n Schüler:in zu diagnostizieren?« Es wurde u. a. gefragt, welche Potenziale und welche Herausforderungen bei der Erhebung der schwimmerischen Grundfertigkeiten von Schüler:innen durch einen (standardisierten) Test erlebt wurden.

Die Interviews dauerten durchschnittlich 32:47 min. Die Äußerungen der Studierenden wurden aufgezeichnet, transkribiert und mittels strukturierender qualitativer Inhaltsanalyse mit dem Softwareprogramm MAXQDA24 kategorial ausgewertet (Kuckartz, 2018). In der Auswertung wurden alle Äußerungen codiert, in denen erlebte Potenziale und Herausforderungen in authentischen Diagnosesituationen beschrieben wurden. Teile des Materials wurden von einer unabhängigen zweiten Person auf der untersten Subkategorieebene co-codiert. Die nach den ersten sieben zufällig ausgewerteten Interviews als Mittelwert über alle Subkategorien erfasste Inter-coderreliabilität ergab 89 %. Aufgrund der hohen Übereinstimmung wurde die Codierung im weiteren Verlauf durch eine Einzelperson fortgeführt.

5.3 Kategorienkonstruktion

Für die Analyse wurde zunächst, ausgehend von den deduktiv abgeleiteten Hauptkategorien, ein in sich schlüssiges Kategoriensystem induktiv und nah am Datenmaterial weiterentwickelt. Dieses dient als Ausgangspunkt für den anschließenden

6 Lernziele des Modells der Schwimmerischen Grundbildung (Fokken et al., 2022; Staub & Fokken, 2020): Phase Rot »Ich fühle mich im und am Wasser wohl«; Phase Gelb »Das Wasser trägt mich«; Phase Grün »Ich komme voran«; Phase Blau »Ich kann am und im Wasser souverän handeln«.

Auswertungsprozess. Zur Validierung der Kategoriensysteme und Codierleitfäden wurden während des gesamten Auswertungsprozesses im Sinne eines zirkulären Verfahrens und im kollegialen Austausch die Inhalte kontinuierlich reflektiert und gegebenenfalls angepasst.

Tabelle 1: Haupt- und Subkategorien der Inhaltsanalyse (die induktiven Subkategorien sind kursiv gekennzeichnet).

Hauptkategorien der erlebten Potenziale und Herausforderungen von authentischen Diagnosesitu- ationen	Subkategorien
Peer Assessment entlang des Fünfer-Schritts (Latzko & Gottlebe, 2022)	Theoretische Konzepte der schwimmerischen Grundbil- dung und Schwimmtechniken erarbeiten (Schritt 1)
	Hypothesen ableiten (Schritt 2)
	Instrumente auswählen und anwenden: Videoaufnahmen jeweils einer Schwimmtechnik (Schritt 3 & 4)
	Der Diagnose entsprechende Förderung durchführen (Schritt 5)
Summatives Assessment von Schüler:innen entlang des Zyklus nach Helmke (2017)	Auswahl von Merkmalen schwimmerischer Grundfertig- keiten (Zyklusstufen a)
	Einschätzung der Leistung durch die Studierenden (Zyklusstufen b)
	Erhebung der Leistung durch einen (standardisierten) Test (Zyklusstufen c)
	Vergleich zwischen persönlicher Einschätzung und Test (Zyklusstufen d)
	Evaluation und Diskussion der Diskrepanz (Zyklusstufen e)
	<i>Anfängliche Unsicherheit und wachsende Sicherheit im Diagnostizieren</i>
	<i>Zukunftsgerichtete Aussagen zum diagnostischen Han- deln als Sportlehrkraft</i>

Um die Forschungsfragen zu beantworten, werden die Ergebnisse entlang der Hauptkategorien der erlebten Potenziale und Herausforderungen analysiert, die Sportlehramtsstudierende für die professionelle Entwicklung ihrer diagnostischen Kompetenzen in einem Seminar mit authentischen formativen und summativen Diagnosesituationen erlebt haben (Tab. 1). Die erlebten Potenziale beziehen sich dabei auf Möglichkeiten (Kramer, 2020) im Seminar »Bewegen im Wasser«, die zu einer wahrgenommenen Verbesserung ihrer diagnostischen Kompetenzen geführt

haben. Mit erlebten Herausforderungen wurden Grenzen für eine diesbezügliche wahrgenommene Verbesserung erfasst.

In der Hauptkategorie Peer Assessment wurden alle Aussagen zu den erlebten Potenzialen bzw. erlebten Herausforderungen eingeordnet, in denen die Studierenden sich gegenseitig in einer Schwimmtechnik diagnostiziert haben. Dazu zählt auch, wenn sie diese mit didaktischen Maßnahmen verknüpft haben. Im Sinne der deduktiven Kategorienbildung wurden in Anlehnung an die Fünfer-Schritt-Methode nach Latzko und Gottlebe (2022) die Subkategorien gebildet. In der Subkategorie »Theoretische Konzepte erarbeiten« (Schritt 1) wurden Aussagen zu Theorie-sitzungen im Seminar zugeordnet, in denen sowohl theoretische Konzepte der Schwimmerischen Grundbildung nach Staub und Fokken (2020) sowie Schwimm-techniken erarbeitet und Hypothesen generiert wurden (Schritt 2). Während des zirkulären Anpassungsprozesses des Codiersystems erfolgte eine Zusammenführung aller Codes von Schritt 3 und 4 in eine gemeinsame Subkategorie. Die Subkategorie »Instrumente auswählen und anwenden« (Schritt 3 und 4) umfasst Aussagen zu den erlebten Potenzialen bzw. erlebten Herausforderungen beim gegenseitigen Filmen in einer selbstgewählten Schwimmtechnik, bei dem anschließenden Beobachten und Beurteilen sowie bei der Integration von Daten zu einem diagnostischen Urteil. Die Subkategorie »Der Diagnose entsprechende Förderung durchführen« (Schritt 5) umfasst Aussagen der Studierende über wahrgenommene Potenziale bzw. Herausforderungen, wenn sie auf Basis der Diagnose Rückmeldungen geben oder Übungen auswählen und umsetzen.

In die Hauptkategorie Diagnostizieren von Schüler:innen entlang der Zyklusstufen nach Helmke (2017) wurden alle Äußerungen zu den erlebten Potenzialen bzw. erlebten Herausforderungen eingeordnet, in denen die Studierenden beschreiben, wie sie Schüler:innen während des Seminars diagnostiziert haben. Diese diagnostischen Handlungen wurden im Sinn der »a-priori-Kategorienbildung« (Kuckartz, 2018, S. 64) den verschiedenen Zyklusstufen nach Helmke (2017) deduktiv zugeordnet. Aussagen von Studierenden, die sich auf Merkmale schwimmerischen Grundfertigkeiten (z. B. Atmen/Atemkontrolle, (Unter-)Tauchen, Gleiten) beziehen, wurden der Zyklusstufe a »Auswahl von Merkmalen« zugeordnet. Die Subkategorie »Einschätzung der Leistung durch die Lehrkraft« (Zyklusstufe b) umfasst Aussagen der Studierenden, in denen sie im Kennenlernspiel mit der Schulklasse eine erste Einschätzung der schwimmerischen Grundfertigkeiten der Schulkinder vorgenommen haben. Die Subkategorie »Erhebung der tatsächlichen Leistung durch einen (standardisierten) Test« (Zyklusstufe c) beinhaltet Aussagen der Studierenden, in denen sie berichten, die schwimmerischen Grundfertigkeiten der Schulkinder mit Hilfe der standardisierten Tests erfasst zu haben (auch während der Kennenlernstunde). Der Subkategorie »Vergleich zwischen der persönlichen Einschätzung und dem Test« (Zyklusstufe d) wurden Vergleiche zwischen persönlichen Einschätzungen und Beurteilungen mit Testergebnissen zugeordnet. Hingegen bezieht sich die Subkategorie

»Evaluation und Diskussion der Diskrepanz« (Zyklusstufe e) auf Aussagen, wie die Studierenden diagnostisches Wissen zur adaptiven Gestaltung des Unterrichts und zur Planung gezielter Übungen nutzen. Wenn die Studierenden anfängliche Bedenken beim summativen Assessment von Schüler:innen äußern und durch das Seminar eine wachsende Sicherheit entwickeln, werden sie im Sinne der induktiven Kategorieneinbildung (Kuckartz, 2018) der Subkategorie »Anfängliche Unsicherheit und wachsende Sicherheit im Diagnostizieren« zugeordnet. Induktiv wurde ebenfalls die Subkategorie »Zukunftsgerichtete Aussagen zum diagnostischen Handeln als Sportlehrkraft« gebildet (s. Tabelle 1). Diesbezüglich wurden Aussagen der Studierenden eingeordnet, ob sie als zukünftige Sportlehrkraft summative Assessments in der Schule umsetzen werden und in welcher Form.

Bei der Analyse wurde jeder Fall zunächst für sich entlang der Kategorien ausgewertet, auf inhaltlich Relevantes reduziert und in einer thematischen Fallzusammenfassung zusammengestellt. Um generalisierende Aussagen treffen zu können, wurden die einzelnen Fälle miteinander verglichen und fallübergreifende Aussagen der Studierenden herausgearbeitet (Kuckartz, 2018). Zitate aus den Transkripten werden den jeweiligen Studierenden durch deren Initialen in Klammern zugeordnet (z. B. MS).

6. Ergebnisse

Ergebnisse zur ersten Forschungsfrage

Die Ergebnisse der ersten Forschungsfrage beziehen sich auf die erlebten Potenziale und Herausforderungen zur professionellen Entwicklung der Kompetenzfacette W-I-E im Peer Assessment. Bei der Codierung zeigt sich, dass sich zu allen fünf Schritten im Fünfer-Schritt des professionellen Diagnostizierens (Latzko & Gottlebe, 2022) Aussagen der Studierenden zuordnen lassen.

Theoretische Konzepte der schwimmerischen Grundbildung und Schwimmtechniken erarbeiten (Schritt 1)

Als Potenzial erleben die Studierenden die theoretische Hinführung zu den Schwimmtechniken zu Beginn des Seminars, um ein Verständnis für diese zu entwickeln. Als hilfreich haben sie dabei auch die Vermittlung von »Idealbewegungen« (MJ, Pos. 24) zu den Schwimmtechniken wahrgenommen, die sie später für das Peer Assessment nutzen konnten. Zudem erleben sie die theoretische Einführung in das Modell der Schwimmerischen Grundbildung als Potenzial für ihr diagnostisches Methodenwissen, da das Modell »handhabbar« (FT, Pos. 87) für das Bewegungsfeld »Bewegen im Wasser« ist. Herausfordernd erleben sie, dass durch das Seminar zwar eine »Grundlage« (SJ, Pos. 61) zu den theoretischen Konzepten geschaffen wurde und sie gelernt haben, welche diagnostische »Instrumente« (SJ, Pos. 61) für das

Bewegen im Wasser genutzt werden können, aber sie sich für die spätere Schulpraxis selber »intensiv damit auseinandersetzen« (SJ) müssen.

Hypothesen ableiten (Schritt 2)

Für das Ableiten von Hypothesen für ihr Peer Assessment haben die Studierenden Informationen auf einer Lernplattform als eine Unterstützungsmaßnahme empfunden. Auf der Lernplattform sind beispielsweise »verschiedene Kategorien« (MS, Pos. 21) zu den Techniken aufgelistet (wie Antriebsmöglichkeiten zu Beinen und Armen oder zur Atemkontrolle). Zudem wurden den Studierenden »Lehrvideos« (VK, Pos. 25) bzw. »Technikleitbilder« (FW, Pos. 41) von Schwimmtechniken zur Verfügung gestellt. Als herausfordernd wurde erlebt, Hypothesen für gute Schwimmer:innen abzuleiten.

Instrumente auswählen und anwenden: Videoaufnahmen jeweils einer Schwimmtechnik (Schritt 3 & 4)

Die Studierenden erkannten ein besonderes Potenzial darin, sich gegenseitig in einer selbst gewählten Schwimmtechnik zu filmen und sich auf diese Weise zu beurteilen. Dies hat zur Verbesserung ihrer Wahrnehmung für Bewegungen im Wasser beigetragen. Als Gründe haben sie die Möglichkeiten genannt, die Bewegungen im Video in reduzierter Geschwindigkeit ansehen sowie mehrfach beobachten und dadurch »viel besser diagnostizier(en)« (SB, Pos. 19) zu können. Dadurch waren »relativ viele typische Fehler« (MS, Pos. 27) erkennbar, »weil wenn man sich das einmal nur anschauen würde, würde man gar nicht alles wahrnehmen« (JK, Pos. 31). Die Studierenden haben dabei die Lehrvideos aus der Lernplattform »im Hinterkopf gehabt« (FW, Pos. 39) und genutzt, um den »Sollzustand (mit dem) Istzustand« (SJ, Pos. 35) der ausgeführten Schwimmtechnik vergleichen und interpretieren zu können. Die Videoanalyse wurde als Potenzial für »alle« (SB, Pos. 21) erlebt, da auch das eigene Beobachten im Video »viel Mehrwert« (SB, Pos. 21) für den Lernenden gebracht hat. Allerdings war dabei für einige Studierende herausfordernd, dass sie das Video sehr oft angesehen haben und »jede einzelne Bewegung« (SJ, Pos. 35) hinterfragt haben. Ohne Videoaufnahme haben die Studierenden die Schnelligkeit der Bewegung als *herausfordernd erlebt*. Dadurch war die »Bewegungsabfolge« (NS, Pos. 69) schwierig zu beobachten, zu beurteilen und zu interpretieren. Studierende haben das Peer Assessment als wenig nützlich empfunden, wenn der oder die zu diagnostizierende Mitstudierende ein:e gute:r Schwimmer:in ist. Es entstanden dadurch Schwierigkeiten bei der Wahrnehmung von weiteren Verbesserungsmöglichkeiten.

Der Diagnose entsprechende Förderung durchführen (Schritt 5)

Bei der Entscheidung für Übungen, mit denen sie ihre:n Kommiliton:in fördern können, haben sich viele der Studierende an den Inhalten der Seminarsitzungen

orientiert, die die Dozentin durchgeführt hat. Diese Erfahrungen erlebten die Studierenden als Potenzial für ihre Entscheidungsprozesse, da sie auf einen »guten Werkzeugkoffer« (MJ, Pos. 4) zurückgreifen können. Darüber hinaus wurde aber auch das »Schwimmwissen aus dem eigenen Sport, dem Triathlon« (FT, Pos. 15) für die Auswahl von Übungen zur Förderung der Schwimmtechnik ihrer Kommiliton:in genutzt. Die Studierenden haben besonderen Wert daraufgelegt, dass die Übungen für ihre Mitstudierenden einen erkennbaren Nutzen haben und zu einem Lernfortschritt beitragen. Als »gewinnbringend« (MJ, Pos. 14) wurde dabei auch das gemeinsame Schwimmen der Übungen erlebt und ermöglichte, dass die Studierenden durch das Peer Assessment die eigene Schwimmtechnik verbessern konnten. Während es vor allem schwimmerfahrenen Studierenden leichter fiel, »typische Fehlerbilder« (MS, Pos. 17) wahrzunehmen und daraufhin Übungen auszuwählen äußerten weniger schwimmerfahrende Studierende es als *herausfordernd* zu entscheiden, »was würde dir auch weiterhelfen?« (MJ, Pos. 36). Auch wurde als *Herausforderung* empfunden, sich bei der Rückmeldung als Teil der Förderung auf das »Wesentliche« (MJ, Pos. 42) zu beschränken, sodass sie manchmal den Mitstudierenden zu viele Rückmeldungen auf einmal gegeben haben. Außerdem äußern einige Studierende, dass Nachfragen zu dem gegebenen Feedback zu Unsicherheiten führten, »wie die Bewegung jetzt wirklich richtig geht« (SJ, Pos. 49).

Ergebnisse zur zweiten Forschungsfrage

Die Ergebnisse zur zweiten Forschungsfrage beziehen sich auf erlebte Potenziale und Herausforderungen von Studierenden für die professionelle Entwicklung ihrer Kompetenzfacette Performanz durch summatives Assessment von Schüler:innen. Bei der Codierung zeigt sich, dass die Studierenden zu zwei der fünf Stufen des Zyklus nach Helmke (2017) keine Aussagen machten. Eine Einschätzung der Eingangsleistung der Schüler:innen durch die Studierenden (Zyklusstufe b) sowie ein Vergleich zwischen persönlicher Einschätzung und Testergebnis (Zyklusstufe d) scheint von den Studierenden nicht durchgeführt zu werden. Bei der Umsetzung der anderen drei Zyklusstufen haben die Studierenden folgende Potenziale und Herausforderungen erlebt:

Auswahl von Merkmalen schwimmerischer Grundfertigkeiten (Zyklusstufen a)

Im Seminar wird das Modell der Schwimmerischen Grundbildung als Grundlage für das summatives Assessment eingeführt und von den Studierenden umgesetzt. Einigen Studierenden ist das diagnostische Vorgehen als »Stufenmodell« (SJ, Pos. 25), anderen als »Diagnosemodell« (FT, Pos. 67) in Erinnerung geblieben. Sie erleben als Potenzial durch die im Modell aufgeführten »sieben Grundfertigkeiten« (FT, Pos. 25), eine Orientierung für ein summatives Assessment im Seminar »Bewegen im Wasser« zu haben. Als herausfordernd äußern sie bei der Auswahl von Merkmalen nicht wie gewohnt einzelne Körperelemente, wie den »Armzug« oder »Beinschlag«

(VK, Pos. 11) zu berücksichtigen, sondern auf die schwimmerischen Grundfertigkeiten zu achten, d. h. »wie (sich) das Kind im Wasser verhält« (VK, Pos. 11).

Erhebung der Leistung durch einen (standardisierten) Test (Zyklusstufen c)

Als besonderes Potenzial haben die Studierenden erlebt, dass sie durch das Modell ihre diagnostische Kompetenz (Performanz) verbessern konnten. Sie haben gelernt das Modell »mit den vier Ebenen« (FT, Pos. 74) in einer »tatsächliche(n) Schulklasse« (MS, Pos. 101) anzuwenden: *»dadurch, dass ich diesen Bogen in der Hand hatte vom Modell war vieles (...) recht einfach. Er hat eine Orientierung gegeben, weil der aus meiner Sicht wirklich die Sachen abdeckt, die zeigen, wie sich das Kind im Wasser fühlt«* (FT, Pos. 23).

Als herausfordernd erleben sie die Ungewissheit, ob sie »jedes Kind perfekt« (FT; Pos. 17) einordnen können und sie haben befürchtet, die Schüler:innen zu überschätzen. Bei der Erhebung der Leistung durch das Modell der Schwimmerischen Grundbildungen haben die Studierenden es als herausfordernd erlebt, dass die einzelnen Stufen in der Theorie zwar »stark abgrenzbar« (MR, Pos. 23), aber in der Praxis vor allem die »höhere(n) Stufen« (MR, Pos. 25) schwer unterscheidbar sind. Die Studierenden haben zudem als Herausforderung geäußert, dass sie mehr Übung in der Anwendung des Modells der Schwimmerischen Grundbildungen brauchen, sodass sie die Kinder »schneller und besser« (JH, Pos. 23) in die Stufen einordnen können.

Evaluation und Diskussion der Diskrepanz (Zyklusstufen e)

Ein Teil der Studierenden haben als Potenzial erlebt, sich über ihre Diagnosen der Schüler:innen auszutauschen, wodurch sie sich »sicher« (FW, Pos. 17) in ihren Urteilen fühlen. Zudem gewinnen die Studierenden dadurch einen »ganz gute(n) Überblick über die Gruppe« (SJ, Pos. 53), was ihnen bei der Stundenplanung geholfen hat. Im Gegensatz dazu, haben andere Studierende geäußert, dass ihnen ein Austausch und eine Diskussion mit Kommilitonen über die Diagnosen der Kinder im Seminar gefehlt hat. Für das Seminar wünschen sie sich »mehr in den Austausch (zu) kommen oder verschiedene Positionen (zu vergleichen), wie ein Kind eingeordnet wird« (FT, Pos. 71). Vereinzelt haben Studierende als Herausforderung erlebt, dass Absprachen über die diagnostizierten Schüler:innen mit anderen Studierenden unzureichend und »nicht präzise genug« (FT, Pos. 27) waren. Als herausfordernd nahmen sie zudem Diskrepanzen in den Diagnosen wahr, da einige Studierende unterschiedlich streng diagnostizierten, d. h. einige waren »lieb und soft« (FT, Pos. 71) und andere Studierende eher »hart« (FT, Pos. 71) in ihren Urteilen und für einige blieb ungeklärt, welches Urteil akkurat ist.

Anfängliche Unsicherheit und wachsende Sicherheit im Diagnostizieren

Als herausfordernd haben einige Studierende erlebt, dass die Proportionen von Erwachsenen und Kindern so unterschiedlich sind und damit Bewegungen im Wasser kaum vergleichbar zu diagnostizieren sind. Dadurch fiel es ihnen anfangs schwer, eine »Brücke« (FT, Pos. 49) von der eigenen Bewegung (bzw. der Bewegung von Kommiliton:innen) zu der Bewegungsausführung von Schüler:innen zu schlagen. Durch die Thematisierung und Anwendung des Modells der Schwimmerischen Grundbildung »in fast jeder Schwimmstunde« (FW, Pos. 5) im Seminar äußern die Studierenden als Potenzial dieses Vorgehens, dass sie sicherer im summativen Assessment geworden sind und einen »genaueren Blick« (VK, Pos. 69) entwickeln konnten, mit dem sie Schüler:innen »bezüglich Wassergewandheit, Wassergefühl und sicheres Bewegen im Wasser« (FT, Pos. 69) einordnen können.

Zukunftsgerichtete Aussagen zum diagnostischen Handeln als Sportlehrkraft

Als Potenzial für ihr zukünftiges diagnostisches Handeln als Sportlehrkraft äußern die Studierenden, dass durch das summativ Assessment und anschließende Unterrichten der Schüler:innen »ein kompletter Bezug zu dem späteren Beruf« (NS, Pos. 19) ermöglicht wurde. Dies haben die Studierenden als hilfreichen Unterschied zu anderen praxisorientierten Seminaren im Fach Sport wahrgenommen, in denen sich ausschließlich Kommiliton:innen gegenseitig angeleitet haben. Für die Schulpraxis äußern sie, noch »viel Erfahrung« (SJ, Pos. 59) im summativen Assessment von Schüler:innen sammeln zu müssen, um ihre diagnostische Kompetenzen zu verbessern. Als Herausforderung für das zukünftige diagnostische Handeln als Sportlehrkraft haben einige Studierende erlebt, dass sie summativ Assessment zwar als wichtig für den späteren Schulalltag erachten, deren Umsetzung jedoch als kaum realisierbar einschätzen. Als Gründe haben sie fehlende personelle und zeitliche Kapazitäten in der Schulpraxis sowie die »Gegebenheiten im Schwimmbad« (MS, Pos. 93) genannt, wie eine unzureichende Anzahl an Schwimmbahnen.

7 Diskussion

7.1 Methodendiskussion

In der vorliegenden Studie wurden die Sichtweisen von Studierenden bezüglich ihrer professionellen Entwicklung diagnostischer Kompetenzen nach der Teilnahme am Seminar »Bewegen im Wasser« erhoben und ausgewertet. Die Erkenntnisse bringen einen Mehrwert dahingehend, die wahrgenommenen Potenziale und Herausforderungen zu erfahren und, ob sie ihren diagnostischen Kompetenzen einen potentiellen Nutzen und ein Erfordernis für die spätere berufliche Praxis zuschreiben (Molina-Soria et al., 2023). Dies kann ermöglichen, dass Sportlehramtsstudierende summativ und formativ Assessment in ihrer späteren Unterrichtspraxis im Bewegen im Wasser nutzen werden (Black & Wiliam, 2010; Borko et al., 2010). Kritisch zu

diskutieren ist, dass die untersuchte Stichprobe relativ klein ist und dass Verfahren zur Erfassung der Performanz weniger auf Selbstbeurteilungen, sondern bspw. auf Beobachtungsverfahren oder gemischten Verfahren (Fremd- und Selbstbeurteilung) basieren sollten. Eine Selbstbeurteilung der Studierenden kann zu einer starken Verzerrung bei der Bewertung der Qualität der eigenen Performanz führen (Baumgartner, 2022a). Zudem ist kritisch anzumerken, dass die Studierenden nur eine begrenzte Perspektive auf die Potenziale des Seminars haben und weitergehende Potenziale (u. a. Auswahl der Seminarinhalte und des methodischen Vorgehens) nicht beurteilen können (Schweer et al., 2025). Mit Blick auf die methodische Umsetzung ist auch einschränkend zu erwähnen, dass in dieser qualitativen Studie die Akkuratheit der Diagnosen von den Sportlehramtsstudierenden nicht berücksichtigt wurde. In zukünftigen Studien ist zu untersuchen, ob erstens das hier vorgestellte Seminarkonzept zu einer verbesserten Diagnosequalität oder einer höheren Akkuratheit diagnostischer Kompetenzen im Bewegungsfeld »Bewegen im Wasser« führt. Zweitens sollte untersucht werden, inwiefern das vermittelte formative und summative Assessment auch in der späteren Berufspraxis umgesetzt wird (Atienza et al., 2023).

7.2 Inhaltliche Diskussion

Für einen Schwimmunterricht mit zunehmend heterogenen Schüler:innen sollten Sportlehramtsstudierende lernen, wie sie Unterrichtsmethoden individualisieren und Schüler:innen gezielte Unterstützung anbieten können. Diagnostische Kompetenzen sind zentral für diesen Prozess (Langer, 2023), da sie ihnen ermöglichen, lernrelevante Merkmale von Schüler:innen im Bewegen im Wasser zu beurteilen und ihre pädagogischen Entscheidungen dahingehend anzupassen. (Diagnostisch) Kompetent ist eine (angehende) Sportlehrkraft dann, wenn sie formative und summative Assessments zielgerichtet auswählen (Kibble, 2017) und in der Komplexität und Ungewissheit von Sportunterricht auf performativer Ebene umsetzen kann (Baumgartner, 2017; Baumgartner, 2022b). Bisherige empirische Befunde zeigen, dass Sportlehrkräfte bislang im Schwimmunterricht damit Schwierigkeiten haben (Fokken et al., 2023). Die professionelle Entwicklung der Kompetenzfacette W-I-E und die Umsetzung des professionellen Wissens in der eigenen Unterrichtspraxis im Sinne von Performanz stellen folglich eine Leistung dar (Baumgartner, 2022b; Ward et al., 2022), die im Seminar im Bewegungsfeld »Bewegen im Wasser« intendiert wurde.

Mit Blick auf die professionelle Entwicklung der Kompetenzfacette W-I-E durch Peer Assessment wird deutlich, dass die von den Studierenden im Seminar entlang der Fünfer-Schritt-Methode nach Latzko und Gottlebe (2022) umgesetzte Theorie-Praxis-Relationierung sowie die Videoanalyse von ihnen als Potenzial für eine verbesserte Wahrnehmung und Interpretation von Schwimmtechnikausführungen erlebt wurden. Als herausfordernd wird die Wahrnehmung und Interpretation lernrelevanter Merkmale im Bewegen im Wasser und die Entscheidung für passende

Übungen u. a. dann erlebt, wenn die Peers bereits gute Schwimmer:innen sind oder wenn Nachfragen zu dem gegebenen Feedback zu Unsicherheiten bezüglich lernrelevanter Merkmale im Bewegen im Wasser führen. Darüber hinaus sollte das Peer Assessment nicht das Feedback der Dozierenden ersetzen. Es sollte darum gehen, das Peer Assessment als eine pädagogisch fundierte Methode einzusetzen, die Sportlehramtsstudierenden helfen kann, diagnostische Kompetenzen (i. S. des formativen Assessments) zu entwickeln (Patton & Marty-Snyder, 2014).

Für die professionelle Entwicklung der Kompetenzfacette Performanz im summativen Assessment wurde der Zyklus nach Helmke (2017) auf das Seminar »Bewegen im Wasser« übertragen (Schlapkohl & Langer, 2024). Als Potenzial erleben die Studierende eine verbesserte Performanz, da sie gelernt haben, Schüler:innen einer »tatsächliche(n) Schulklassen« (MS, Pos. 101) entlang des Modells der Schwimmerischen Grundbildungen »einzustufen« (MR, Pos. 21). Das Seminarkonzept hat ihnen ermöglicht, diagnostisches Methodenwissen zum summativen Assessment zu erlangen und sie haben gelernt, es in berufsnahen, authentischen Situationen umsetzen zu können (Wiliam, 2011). Ähnliche Ergebnisse (allerdings in der Umsetzung formativer Assessments) zeigen sich auch bei Macken et al. (2020). Die Autor:innen haben herausgefunden, dass Sportlehramtsstudierende ihre professionelle Entwicklung diagnostischer Kompetenzen durch die Umsetzung in Schulpraktika höher wahrnehmen, als wenn sie die Inhalte mit theoretischen Diskussionen in Seminaren lernen. Voraussetzung hierfür ist aber auch eine adäquate Betreuung und Beratung hinsichtlich der Umsetzung formativer Assessments durch die Hochschullehrkraft während der Praktika. Mit Blick auf die professionelle Entwicklung der Kompetenzfacette Performanz im summativen Assessment entlang des Zyklus nach Helmke (2017) zeigen die Ergebnisse allerdings auch, dass bei der Umsetzung im Seminar zwei der fünf Zyklusstufen zur Verbesserung der Diagnosefähigkeit nach Helmke (2017) von den Studierenden nicht wahrgenommen werden. Dies betrifft zum einen die persönliche Einschätzung der Eingangsleistung der Schüler:innen durch die Studierenden und zum anderen den Vergleich dieser ersten persönlichen Einschätzung mit dem anschließend durchgeführten Test. In Zukunft sollte daher mehr Wert auf die Durchführung einer ersten Kurzbeobachtung und den anschließenden Vergleich gelegt werden. Zudem sollte eine kritische Auseinandersetzung zwischen den Kommiliton:innen über ihr summatives Assessment der Schüler:innen initiiert werden. Dies ist relevant, da viele Studierende annehmen, dass sich ihre diagnostische Kompetenz (Performanz) von Bewegungen im Wasser allein durch praktische Erfahrung verbessern würde. Allerdings zeigen bisherige empirische Befunde, dass Erfahrung allein für eine Entwicklung und Verbesserung diagnostischer Kompetenzen nicht ausreicht (Slingerland et al., 2024). »Das eigentliche Potenzial« des Zyklus liegt im »Anstoßen von Diskursen über die Ergebnisse (...) insbesondere, wenn sie von der Erwartung abweichen« (Helmke, 2017, S. 141). Dies kann auch ein Umdenken bewirken. In Bezug auf das zukünftige diagnostische Handeln als Sportlehrkraft zeigt sich, dass die Sportlehramtsstudierende bezweifeln, summative Assessments im

Bewegungsfeld »Bewegen im Wasser« mit einer gesamten Klasse umsetzen zu können (u. a. aufgrund fehlender personeller und zeitlicher Kapazitäten und nicht aufgrund mangelnder professioneller Kompetenz).

Die Ergebnisse der vorliegenden Studie sind mit fachübergreifenden Studien vergleichbar, wonach für Erwerbsprozesse die Verknüpfung von vorhandenem Wissen mit Situationen sowie die Reflexion gewonnener Unterrichtserfahrungen wichtig sind (König et al., 2020). Die untersuchten Studierenden scheinen sowohl das Aneignen und Anwenden von diagnostischem Methodenwissen als auch die berufsnahen Lerngelegenheiten als Potenzial für die professionelle Entwicklung ihrer diagnostischen Kompetenzen zu erleben (Chernikova et al., 2020). Das Seminar setzt damit an dem hohen Potenzial der »practice-based teacher education« (Brataas & Jensen, 2023, S. 1) an, indem in der Sportlehrkräftebildung »coursework more closely to fieldwork« (Brataas & Jensen, 2023, S. 1) herangeführt wird. Dieser Forderung nach möglichst schulpraktischen Lerngelegenheiten (Chernikova et al., 2020; Ward et al., 2022) wird in Deutschland durch das Praxissemester als Teil der ersten, universitären Phase der Sportlehrkräftebildung bereits begegnet. Darüber hinausgehende Kooperationen mit Schulen (wie es in dieser Studie der Fall ist), sind je nach Standort aus strukturellen Gründen nicht überall möglich.

In Bezug auf die (Sport-)Lehrkräfteprofessionalisierung sollten Sportlehramtsstudierende in weiteren Lehr- und Lernsituationen die Möglichkeit erhalten, ihre diagnostischen Kompetenzen in authentischen und praxisnahen Situationen weiterzuentwickeln und zu verbessern (Chernikova et al., 2020; Ward et al., 2022). Dies kann sowohl in der Hochschulausbildung in weiteren Bewegungsfeldern als auch im Rahmen der zweiten Ausbildungsphase (Referendariat) erfolgen.

Literaturverzeichnis

- Atienza, R., Valencia-Peris, A., & López-Pastor, V. M. (2023). Formative assessment and pre-service teacher education: Previous, current and prospective experiences. *Cultura, Ciencia y Deporte*, 18(55), 133–156. <https://doi.org/10.12800/ccd.v18i55.1914>
- Baumert, J., & Kunter, M. (2006). Stichwort: Professionelle Kompetenz von Lehrkräften. *Zeitschrift für Erziehungswissenschaft*, 9(4), 469–520. <https://doi.org/10.1007/s11618-006-0165-2>
- 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 (Eds.), *Cognitive activation in the mathematics classroom and professional competence of teachers: Results from the COACTIV project* (pp. 25–48). Springer. https://doi.org/10.1007/978-1-4614-5149-5_2
- Baumgartner, M. (2017). »Denn sie wissen nicht, was sie können ...!« – die Qualität der Performanzen von angehenden Sportlehrkräften als Verzerrer der Selbstbeurteilung. *German Journal of Exercise and Sport Research*, 47(3), 246–254. <https://doi.org/10.1007/s12662-017-0453-4>
- Baumgartner, M. (2022a). Professional competence(s) of physical education teachers: Terms, traditions, modelling and perspectives. *German Journal of Exercise and Sport Research*, 52(4), 550–557. <https://doi.org/10.1007/s12662-022-00840-z>
- Baumgartner, M. (2022b). Professionelle Kompetenz(en) von Sportlehrkräften – Begriffe, Traditionen, Modellierungen und Perspektiven. In R. Sygusch, J. Hapke, S. Liebl & C.

- Töpfer (Hrsg.), *Kompetenzorientierung im Sport. Grundlagen, Modellentwurf und Anwendungsbeispiele* (S. 27–41). Hofmann.
- Black, P., & Wiliam, D. (2010). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*, 92(1), 81–90. <https://doi.org/10.1177/003172171009200119>
- 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>
- Brataas, G., & Jensen, I. S. (2023). From coursework to fieldwork: How do teacher candidates enact and adapt core practices for instructional scaffolding? *Teaching and Teacher Education*, 132, 104206. <https://doi.org/10.1016/j.tate.2023.104206>
- Borghouts, L. B., Slingerland, M., & Haerens, L. (2017). Assessment quality and practices in secondary PE in the Netherlands. *Physical Education and Sport Pedagogy*, 22(5), 473–489. <http://dx.doi.org/10.1080/17408989.2016.1241226>
- Borko, H., Jacobs, J., & Koellner, K. (2010). Contemporary approaches to teacher professional development. In P. Peterson, E. Baker & B. McGaw (Eds.), *International encyclopedia of education* (7th ed., pp. 548–556). Elsevier. <https://doi.org/10.1016/B978-0-08-044894-7.00654-0>
- Chernikova, O., Heitzmann, N., Fink, M. C., Timothy, V., Seidel, T., & Fischer, F. (2020). Facilitating Diagnostic Competences in Higher Education - a Meta-Analysis in Medical and Teacher Education. *Educational Psychology Review*, 32(1), 157–196. <https://doi.org/10.1007/s10648-019-09492-2>
- Danthony, S., Mascaret, N., & Cury, F. (2019). Test anxiety in physical education: The predictive role of gender, age, and implicit theories of athletic ability. *European Physical Education Review*, 18(1), 128–143. <https://doi.org/10.1177/1356336X19839408>
- Eather, N., Riley, N., Miller, D., & Jones, B. (2017). Evaluating the effectiveness of using peer-dialogue assessment (PDA) for improving pre-service teachers' perceived confidence and competence to teach physical education. *Australian Journal of Teacher Education*, 42(1), 69–83. <https://doi.org/10.14221/ajte.2017v42n1.5>
- Feth, C. (2014). *Benotungspraxis von Sportlehrkräften. Rekonstruktion von Benotungsstrategien im Sportunterricht*. Dissertation, Ruhr-Universität Bochum.
- Fokken, I., Staub, I., & Vogt, T. (2022). Lernstandsdiagnostik im Schwimmunterricht. Lerngruppen analysieren und Unterricht passgenau gestalten. *Sportunterricht*, 71(4), 181–186.
- Fokken, I., Staub, I., & Vogt, T. (2023). Physical education teachers' swimming skill analysis in 6- to 12-year-old children: Findings from an online survey. *Journal of Teaching in Physical Education*, 43(2), 333–342. <https://doi.org/10.1123/jtpe.2022-0297>
- Grossman, P. (2021). *Teaching Core Practices in Teacher Education*. Harvard Education Press.
- Herrero-González, D., López-Pastor, V., Manrique-Arribas, J. C., & Moura, A. (2023). Formative and shared assessment: Literature review on the main contributions in physical education and physical education teacher education. *European Physical Education Review*, 30(3). <https://doi.org/10.1177/1356336X231220995>
- Helmke, A. (2017). *Unterrichtsqualität und Lehrerprofessionalität. Diagnose, Evaluation und Verbesserung des Unterrichts* (7. Aufl.). Klett/Kallmeyer.
- Hoyle, E. (1991). Professionalisierung von Lehrern. Ein Paradox. In E. Terhart (Hrsg.), *Unterrichten als Beruf. Neuere amerikanische und englische Arbeiten zur Berufskultur und Berufsbiographie von Lehrern und Lehrerinnen* (S. 135–144). Böhlau Verlag.
- Kibble, J. (2017). Best practices in summative assessment. *Advances in Physiology Education*, 41(1), 110–119. <https://doi.org/10.1152/advan.00116.2016>
- Kramer, R.-T. (2020) Zum Problem der Professionalisierung im Lehramtsstudium und zum Potenzial der Hochschullernwerkstatt. In K. Kramer, D. Rumpf, M. Schöps & S. Winter (Hrsg.), *Hochschullernwerkstätten – Elemente von Hochschulentwicklung? Ein Rückblick auf 15 Jahre Hochschullernwerkstatt in Halle und andernorts* (S. 275–288). Klinkhardt. <https://doi.org/10.25656/01.21220>
- König, J., Darge, K., & Kramer, C. (2020). Kompetenzentwicklung im Praxissemester. Zur Bedeutung schulpraktischer Lerngelegenheiten auf den Erwerb von pädagogischem Wissen bei Lehramtsstudierenden. In I. Ulrich & A. Gröschner (Hrsg.), *Praxissemester im Lehramtsstudium in Deutschland. Wirkungen auf Studierende* (S. 67–95). Springer VS.

- Kuckartz, U. (2018). *Qualitative Inhaltsanalyse. Methoden, Praxis, Computerunterstützung* (4. Aufl.). Beltz Juventa.
- Lamb, P., Lane, K., & Aldous, D. (2013). Enhancing the spaces of reflection: A buddy peer-review process within physical education initial teacher education. *European Physical Education Review*, 19(1), 21–38. <https://doi.org/10.1177/1356336X12457293>
- Langer, A. (2023). *Pädagogische Diagnosen im Kontext der Unterrichtsgestaltung im Sport*. Hofmann. (Reihe Junge Sportwissenschaft, Bd. 20).
- Latzko, B., & Gottlebe, K. (2022). Der Nutzen der pädagogisch-psychologischen Diagnostik für eine diversitätssensitive und inklusive, pädagogische Praxis. In K. Kramer & B. Hoyer (Hrsg.), *"Individuell fördern" – Wissenschaftlicher Hintergrund sowie Ansatzpunkte aus und für die Praxis* (S. 43–67). FAU University Press.
- Macken, S., MacPhail, A., & Calderon, A. (2020). Exploring primary pre-service teachers' use of 'assessment for learning' while teaching primary physical education during school placement. *Physical Education and Sport Pedagogy*, 25(5), 539–554. <https://doi.org/10.1080/17408989.2020.1752647>
- Molina Soria, M., López-Pastor, V. M., Hortigüela-Alcalá, D., Pascual-Arias, C., & Fernández-Garcimartín, C. (2023). Formative and shared assessment and feedback: an example of good practice in physical education in pre-service teacher education. *Cultura, Ciencia y Deporte*, 18(55), 157–169. <https://doi.org/10.12800/ccd.v18i55.1986>
- Patton, B. J. & Marty-Snyder, M. (2014). Practical Applications for Using Peer Assessment in Physical Education Teacher Education Field Experiences, *Strategies: A Journal for Physical and Sport Educators*, 27(6), 25–31. <https://doi.org/10.1080/08924562.2014.960121>
- Pastore, S., & Andrade, H. L. (2019). Teacher assessment literacy: A three-dimensional model. *Teaching and Teacher Education*, 84, 128–138. <https://doi.org/10.1016/j.tate.2019.05.003>
- Schildkamp, K., van der Kleij, F. M., Heitink, M. C., Kippers, W. B., & Veldkamp, B. P. (2020). Formative assessment: A systematic review of critical teacher prerequisites for classroom practice. *International Journal of Educational Research*, 103, 101602. <https://doi.org/10.1016/j.ijer.2020.101602>
- Schittkowski, B., Woll, A., & Wagner, I. (2024). Multidimensionale Herausforderungen und Strategien der Leistungsbewertung im Sportunterricht. Eine qualitative Analyse. *German Journal of Exercise and Sport Research*, 55(4), 555–565. <https://doi.org/10.1007/s12662-024-01004-x>
- Schlapkohl, N., & Langer, A. (2024). Einbindung pädagogischer Diagnosen in die Lehramtsausbildung - Umsetzungsmöglichkeiten im Seminar Bewegten im Wasser. *Zeitschrift für Studium und Lehre in der Sportwissenschaft*, 7(2), 13–19.
- Schrader, F.-W. (2009). Anmerkungen zum Themenschwerpunkt Diagnostische Kompetenz von Lehrkräften. *Zeitschrift für Pädagogische Psychologie*, 23(3-4), 237–245. <https://doi.org/10.1024/1010-0652.23.34.237>
- Schweer, M., Thies, B., & Lachner, R. (2025) Soziale Wahrnehmungsprozesse und unterrichtliches Handeln. Eine dynamisch-transaktionale Perspektive. In M. Schweer (Hrsg.), *Lehrer-Schüler-Interaktion. Inhaltsfelder, Forschungsperspektiven und methodische Zugänge* (S. 95–120). Springer VS.
- Slingerland, M., Weeldenburg, G., & Borghouts, L. (2024). Formative assessment in physical education: Teachers' experiences when designing and implementing formative assessment activities. *European Physical Education Review*, 30(4), 620–637. <https://doi.org/10.1177/1356336X241237398>
- Snow, S. L., Brown, T. D., Barnett, L. M., Salmon, J., & Lander, N. J. (2025). Understanding the Use of Motor Competence Assessments in Australian Elementary Physical Education: A Qualitative Study. *Journal of Teaching in Physical Education*. <https://doi.org/10.1123/jtpe.2024-0370>
- Starck, J. R., O'Neil, K. M., & Richards, K. A. (2023). Preservice teachers' perceptions of and intentions to utilize assessment. *Curriculum Studies in Health and Physical Education*, 15(3), 324–339. <https://doi.org/10.1080/25742981.2023.2240300>
- Staub, I., & Fokken, I. (2020). Vermittlungsinhalte einer umfassenden schwimmerischen Grundbildung. In T. Vogt (Hrsg.), *Vermittlungskompetenz in Sport, Spiel und Bewegung: sportartspezifische Perspektiven* (S. 129–148). Meyer & Meyer.

- Terhart, E. (2007). Erfassung und Beurteilung der beruflichen Kompetenz von Lehrkräften. In M. Lüders & J. Wissinger (Hrsg.), *Forschung zur Lehrerbildung. Kompetenzentwicklung und Programmevaluation* (S. 37–62). Waxmann.
- Tolgfors, B., Quennerstedt, M., Backman, E., & Nyberg, G. (2021). Enacting assessment for learning in the induction phase of physical education teaching. *European Physical Education Review*, 28(2), 534–551. <https://doi.org/10.1177/1356336X211056208>
- Varveri, D., Flouris, A. D., Smirnios, N., Pollatou, E., Karatzaferi, C., & Sakkas, G. K. (2016). Developing and testing an instrument to assess aquaticity in humans. *Journal of Bodywork and Movement Therapies*, 20(3), 497–503. <https://doi.org/10.1016/j.jbmt.2015.12.013>
- Vogt, T., & Staub, I., (2020). Assessment of basic aquatic skills in children: Interrater reliability of coaches, teachers, students and parents. *Journal of Physical Education and Sport*, 20(2), 577–583. <https://doi.org/10.7752/jpes.2020.02085>
- Wahl, D., Weinert, F. E., & Huber, G. L. (2006). *Psychologie für die Schulpraxis. Ein handlungsorientiertes Lehrbuch für Lehrerinnen und Lehrer*. Sozio-Publishing.
- Ward, P., Dervent, F., Devrilmez, E., Iserbyt, P., Kim, I., Ko, B., Santiago, J. A., Tsuda, E., & Xie, X. (2022). Practice-based teacher education in physical education. *Journal of Teaching in Physical Education*, 42(3), 442–451. <https://doi.org/10.1123/jtpe.2022-0047>
- Weinert, F. E. (2001). Concept of Competence: A Conceptual Clarification. In D. S. Rychen & L. H. Salganik (Eds.), *Defining and selecting key competencies* (1. Aufl., pp. 45–65). Hogrefe & Huber.
- Wiliam, D. (2011) What is assessment for learning? *Studies in Educational Evaluation*, 37(1), 3–14. <https://doi.org/10.1016/J.STUEDUC.2011.03.001>

Ankündigung der Jahrestagung der DGfE 2026 »Ehrliche Theorie – für die Sportdidaktik und die Sportpädagogik«

26.01.2027 – 29.01.2027 in Laax Graubünden, Schweiz

Das Datum mag irritieren, ist aber kein Versehen: Die Jahrestagung 2026 findet im Januar 2027 statt.

Ehrliche Theorie – für die Sportdidaktik und die Sportpädagogik

Die Sportpädagogik und Sportdidaktik sind als Professionswissenschaften immer dem Spannungsfeld von Theorie und Praxis ausgesetzt. Die Jahrestagung der DGfE-Kommission Sportpädagogik widmet sich dieser wechselseitigen Beziehung, die – folgt man hier Oelkers (1994) – als Verhältnis bereits zum Problem geworden ist.

Wir nähern uns dem Thema mehrheitlich aus der Sicht der Theorien. Ziel ist es, die Funktion und Bedeutung von Theorie in der Sportdidaktik und -pädagogik an konkreten Beispielen zu diskutieren. Theorien, die sich nicht der Praxis anbieten, aber auch nicht nur für sich selbst stehen – deshalb *ehrlich*.

Das bedeutet nicht, dass diese Theorien sich nicht der Praxis stellen müssen. Eine Professionswissenschaft wäre nicht professionell, wenn sie sich nicht der Praxis stellen würde. Deshalb sollen an dieser Tagung die vorgestellten Theorien mit Vignetten der Praxis konfrontiert werden. Nicht um zu testen, ob sie sich anwenden lassen, sondern vielmehr um diese Theorien in ihrer Kohärenz und Aussagekraft zu stärken. Das kann aber ohne unmittelbaren Anwendungsanspruch und Alltagsdruck in akademischer Ruhe geschehen.

Diese Ruhe und Distanz wollen wir an der Jahrestagung der DGfE-Kommission Sportpädagogik der Theorie gönnen. Die Ruhe und Distanz suchen wir auf dem Crap Sogn Gion im Skigebiet von LAAX (Graubünden, Schweiz). Gleichsam in Anlehnung an die Davoser Hochschulkurse von 1928-1931 (z.B. zwischen Heidegger und Cassirer) versuchen wir der Theorie Platz und Raum zu geben, ohne gleich mit dem oben erwähnten eklektischen Anwendungsdruck konfrontiert zu sein.

Als Hauptreferent:innen sind vorgesehen:

Hartmut Rosa (Friedrich-Schiller Universität Jena)

Carolin Rotter (Universität Duisburg-Essen)

Ansgar Thiel (Rektor Sporthochschule Köln)

Tagungshomepage: <https://www.sportdidaktik.ch/forschung/dgfe-sportp%C3%A4dagogik-2026>

Wir freuen uns auf Ihre/Eure Teilnahme

stellvertretend für das Tagungsteam,

Prof. Dr. Roland Messmer

Ausschreibung für bis zu zwei Nachfolge-Positionen im Herausgeber:innenkollegium (2027-2030)

Liebe Kolleg:innen,

aus guten Gründen ist es vorgesehen, dass das Herausgeber:innenkollegium der ZSF nicht in permanent gleicher Besetzung agiert, sondern in einem regelmäßigen Rhythmus sukzessive erneuert wird. Planmäßig werden daher einige Mitherausgeber:innen aus dem Kollegium ausscheiden. Das verbleibende Herausgeber:innenkollegium ist deshalb auf der Suche nach bis zu zwei neuen Mitgliedern.

Gerne nehmen wir deshalb Bewerbungen oder Vorschläge für die Nachfolge-Positionen für den Zeitraum von 2027 - 2030 entgegen. **Bewerbungsschluss bzw. Vorschlagsfrist ist der 31. September 2026.**

Es wäre hilfreich, wenn die Bewerber:innen über (umfangreiche) Erfahrungen in der Begutachtung oder Herausgeberschaft von Zeitschriften und/oder Büchern verfügen.

Die Ernennung der neuen Herausgeberinnen oder Herausgeber wird vom Vorstand des Vereins zur Förderung sportpädagogischer Forschung e.V in Abstimmung mit den Sprecherräten der dvs-Sektion Sportpädagogik und der DGfE-Kommission Sportpädagogik sowie den verbliebenen Herausgeber:innen vorgenommen.

Bitte richten Sie Ihre Bewerbung bzw. Ihren Vorschlag (der mit der genannten Person abgestimmt sein sollte) an die geschäftsführende Herausgeberin Prof. Dr. Miriam Seyda (zsf.issw.fk16@tu-dortmund.de).

Mit besten Grüßen,

Miriam Seyda

Autor:innenverzeichnis

Diagnostische Kompetenzen aus sportpädagogischer Perspektive in Theorie und Empirie

Dr. Anneke Langer
Europa-Universität Flensburg
Institut für Sportwissenschaft
Campusalle 2
24943 Flensburg
Anneke.langer@uni-flensburg.de

Dr. Philipp Hendricks
Institut für Sport und Sportwissenschaft der TU Dortmund
Otto-Hahn-Straße 3
44227 Dortmund
E-Mail: philipp.hendricks@tu-dortmund.de
Tel.: +49 (231) 755-4169

Diagnostic Competence and Judgement Accuracy of Primary School PE Teachers in Assessing Basic Motor Competencies

Storni Simone
Scuola Universitaria Professionale della Svizzera Italiana
Dipartimento Formazione e Apprendimento – Alta Scuola Pedagogica
Piazza San Francesco 19
6600 Locarno
simone.storni@supsi.ch

Kress Johanna, Bretz Kathrin, Herrmann Christian, Ferrari Ilaria
Pädagogische Hochschule Zürich
Forschungsgruppe Didaktik Bewegung und Sport
Lagerstrasse 2
8090 Zürich
johanna.kress@phzh.ch
kathrin.bretz@phzh.ch
christian.herrmann@phzh.ch
ilaria.ferrari@phzh.ch

Lenzen Benoît

Université de Genève

Faculté de Psychologie et des Sciences de l'Éducation & Institut Universitaire de
Formation des Enseignants

24 rue du Général-Dufour

1211 Genève 4

benoit.lenzen@unige.ch

**Aufgabenanalytische Kompetenzen angehender Sportlehrkräfte als eine
Voraussetzung zur kognitiven Aktivierung von Schüler*innen im Sportunterricht**

Dr. Philipp Hendricks

Institut für Sport und Sportwissenschaft der TU Dortmund

Otto-Hahn-Straße 3

44227 Dortmund

E-Mail: philipp.hendricks@tu-dortmund.de

Tel.: +49 (231) 755-4169

**Digital Transformation of a Diagnostic Procedure for Swimming Skill Analysis:
Teachers' Evaluation of a Low-Fidelity App Prototype (Click Dummy) of the 'Owl's
Eye'**

Inga Fokken

Deutsche Sporthochschule Köln

Institut für Vermittlungskompetenz in den Sportarten

Am Sportpark Müngersdorf 6

50933 Köln

i.fokken@dshs-koeln.de

Dr. Ilka Staub

Deutsche Sporthochschule Köln

Institut für Vermittlungskompetenz in den Sportarten

Am Sportpark Müngersdorf 6

50933 Köln

i.staub@dshs-koeln.de

Dr. Michel Brinkschulte

Deutsche Sporthochschule Köln

Institut für Vermittlungskompetenz in den Sportarten

Am Sportpark Müngersdorf 6

50933 Köln

m.brinkschulte@dshs-koeln.de

Edgar Sauerbier
Deutsche Sporthochschule Köln
Institut für Vermittlungskompetenz in den Sportarten
Am Sportpark Müngersdorf 6
50933 Köln
e.sauerbier@dshs-koeln.de

Dr. Nele Schlapkohl
Europa-Universität Flensburg
Institut für Sportwissenschaft
Campusallee 2
24943 Flensburg
schlapkohl@uni-flensburg.de

Univ.-Prof. Dr. Tobias Vogt
Deutsche Sporthochschule Köln
Institut für Vermittlungskompetenz in den Sportarten
Am Sportpark Müngersdorf 6
50933 Köln
t.vogt@dshs-koeln.de

Diagnostische Kompetenzen von Sportlehramtsstudierenden durch authentische Situationen professionell entwickeln

Dr. Anneke Langer
Europa-Universität Flensburg
Institut für Sportwissenschaft
Campusallee 2
24943 Flensburg
Anneke.langer@uni-flensburg.de
Tel.: (0461) 805 2847

Dr. Nele Schlapkohl
Europa-Universität Flensburg
Akademische Rätin am Institut für Sportwissenschaft
Campusallee 2
24943 Flensburg
schlapkohl@uni-flensburg.de

Impressum ZSF – Zeitschrift für Sportpädagogische Forschung

ISSN 2196-5218

Organ der Deutschen Vereinigung für Sportwissenschaft

Herausgegeben vom Verein zu Förderung sportpädagogischer Forschung e. V. im Auftrag der Kommission ›Sportpädagogik‹ der Deutschen Gesellschaft für Erziehungswissenschaften (DGfE) sowie der Sektion ›Sportpädagogik‹ der Deutschen Vereinigung für Sportwissenschaft (dvs).

Herausgeberkollegium: Ulrike Burrmann (Berlin), Claus Krieger (Hamburg), Hilke Pallesen (Hamburg), Miriam Seyda (geschäftsführend, Dortmund), Jan Sohnsmeier (Heidelberg), Vera Volkmann (Hannover)

Schriftleitung: Miriam Seyda (V.i.S.d.P.)

Einsendungen bitte an:

Prof. Dr. Miriam Seyda
Institut für Sport und Sportwissenschaft
TU Dortmund
Otto-Hahn-Straße 3
44227 Dortmund
E-Mail: zsf@nomos-journals.de
www.zsf.nomos.de || www.zfso.de

Manuskripte und andere Einsendungen: Alle Einsendungen sind an die o. g. Adresse zu richten. Es besteht keine Haftung für Manuskripte, die unverlangt eingereicht werden. Sie können nur zurückgegeben werden, wenn Rückporto beigelegt ist. Die Annahme zur Veröffentlichung muss in Textform erfolgen. Mit der Annahme zur Veröffentlichung überträgt die Autorin/der Autor der Nomos Verlagsgesellschaft mbH & Co.KG an ihrem/seinem Beitrag für die Dauer des gesetzlichen Urheberrechts das exklusive, räumlich und zeitlich unbeschränkte Recht zur Vervielfältigung und Verbreitung in körperlicher Form, das Recht zur öffentlichen Wiedergabe und Zugänglichmachung, das Recht zur Aufnahme in Datenbanken, das Recht zur Speicherung auf elektronischen Datenträgern und das Recht zu deren Verbreitung und Vervielfältigung sowie das Recht zur sonstigen Verwertung in elektronischer Form. Hierzu zählen auch heute noch nicht bekannte Nutzungsformen. Das in § 38 Abs. 4 UrhG niedergelegte zwingende Zweitverwertungsrecht der Autorin/des Autors nach Ablauf von 12 Monaten nach der Veröffentlichung bleibt hiervon unberührt. Eine eventuelle, dem einzelnen Beitrag oder der jeweiligen Ausgabe beigelegte Creative Commons-Lizenz hat im Zweifel Vorrang. Zum Urheberrecht vgl. auch die allgemeinen Hinweise unter www.nomos.de/urheberrecht. Unverlangt eingesandte Manuskripte – für die keine Haftung übernommen wird – gelten als Veröffentlichungsvorschlag zu den Bedingungen des Verlages. Es werden nur unveröffentlichte Originalarbeiten angenommen. Die Verfasser erklären sich mit einer nicht sinnentstellenden redaktionellen Bearbeitung einverstanden.

Urheber- und Verlagsrechte: Alle in dieser Zeitschrift veröffentlichten Beiträge sind urheberrechtlich geschützt. Der Rechtsschutz gilt auch im Hinblick auf Datenbanken und ähnlichen Einrichtungen. Kein Teil dieser Zeitschrift darf außerhalb der engen Grenzen des Urheberrechtsgesetzes oder über die Grenzen einer eventuellen, für diesen Teil anwendbaren Creative Commons-Lizenz hinaus ohne schriftliche Genehmigung des Verlages in irgendeiner Form vervielfältigt, verbreitet oder öffentlich wiedergegeben oder zugänglich gemacht, in Datenbanken aufgenommen, auf elektronischen Datenträgern gespeichert oder in sonstiger Weise elektronisch vervielfältigt, verbreitet oder verwertet werden. Namentlich gekennzeichnete Artikel müssen nicht die Meinung der Herausgeber/Redaktion wiedergeben. Der Verlag beachtet die Regeln des Börsenvereins des Deutschen Buchhandels e.V. zur Verwendung von Buchrezensionen.

Anzeigen:

Verlag C.H.Beck GmbH & Co. KG
Media Sales
Dr. Jiri Pavelka
Wilhelmstraße 9
80801 München

Tel.: (089) 381 89-687
mediasales@beck.de

Verlag und Gesamtverantwortung für Druck und Herstellung:

Nomos Verlagsgesellschaft mbH & Co. KG,
Waldseestr. 3-5,
76530 Baden-Baden,
Telefon: 07221/2104-0,
Telefax 07221/2104-27,
www.nomos.de

Geschäftsführer: Thomas Gottlöber
HRA 200026, Mannheim

Sparkasse Baden-Baden Gaggenau,
IBAN DE05662500300005002266
(BIC SOLADES1BAD).

Erscheinungsweise: halbjährlich

Preise: Individualkunden: Jahresabo €59 inkl. digitaler Einzelplatzlizenz. Institutionen: Jahresabo € 129,- inkl. digitaler Mehrplatzlizenz. Der Digitalzugang wird über die digitale Plattform des Verlags – Inlibra – bereitgestellt. Einzelheft: € 41,- Die Abopreise verstehen sich einschließlich der gesetzlichen Umsatzsteuer und zuzüglich Vertriebskostenanteil (Inland € 18,80/Ausland € 28,20) bzw. Direktbeorderungsgebühr € 3,50 Die Rechnungsstellung erfolgt nach Erscheinen des ersten Heftes des Jahrgangs.

Bestellungen über jede Buchhandlung und beim Verlag.

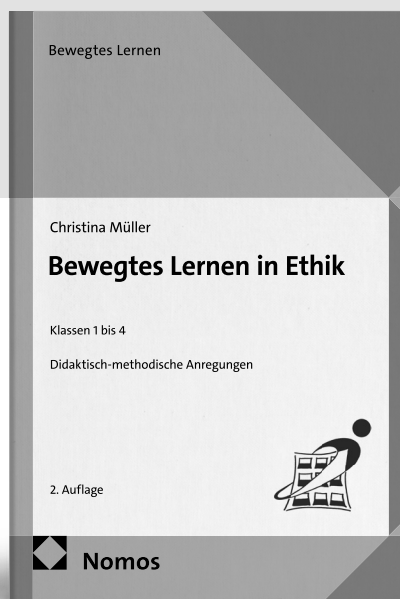
Kundenservice:

Telefon: +49-7221-2104-222
Telefax: +49-7221-2104-285
E-Mail: service@nomos.de

Kündigung: Abbestellungen mit einer Frist von sechs Wochen zum Kalenderjahresende.

Adressenänderungen: Teilen Sie uns rechtzeitig Ihre Adressenänderungen mit. Dabei geben Sie bitte neben dem Titel der Zeitschrift die neue und die alte Adresse an. Hinweis gemäß Art. 21 Abs. 1 DSGVO: Bei Anschriftenänderung kann die Deutsche Post AG dem Verlag die neue Anschrift auch dann mitteilen, wenn kein Nachsendeauftrag gestellt ist. Hiergegen kann jederzeit mit Wirkung für die Zukunft Widerspruch bei der Post AG eingelegt werden.

Lernen mit Kopf und Körper



Bewegtes Lernen in Ethik

Klassen 1 bis 4

Didaktisch-methodische Anregungen

Von Prof. i.R. Dr. Christina Müller

2., aktualisierte Auflage 2026,

120 S., brosch., 22,90 €

ISBN 978-3-7560-3288-4

E-Book 978-3-7489-6675-3

(Bewegtes Lernen, Bd. 14)

Schüler können Informationen zusätzlich zum optischen und akustischen auch über den Bewegungssinn aufnehmen. Bewegungen verbessern bereits mit geringer Intensität die Sauerstoffversorgung des Gehirns und dadurch die Informationsverarbeitung.

Die didaktisch-methodischen Anregungen zeigen Konkretisierungen für bewegtes Lernen im Ethikunterricht der Klassen 1 bis 4 mit folgenden Zielstellungen:

- Zusätzliche Informationszugänge zu erschließen, z. B. von Menschen und Kulturen über Bewegungshandlungen, Ausdrücken von Verhaltensweisen oder Gefühlen mittels Körpersprache bzw. das szenische Darstellen ethischer Probleme
- Die Informationsverarbeitung zu optimieren durch das Einprägen von Normen und Rechten beim Zuspieren eines Balles, das Darlegen von Erfahrungen sowie Positionen beim Gehen mit einem Partner durch den Raum oder den Austausch von Meinungen, verbunden mit Platzwechseln

Digital verfügbar auf  **inlibra.com**

Bestellen Sie im Buchhandel oder versandkostenfrei unter [nomos-shop.de](https://www.nomos-shop.de)

Alle Preise inkl. Mehrwertsteuer



Nomos