

FACHDIDAKTISCHES WISSEN VON SPORTLEHRER*INNEN TESTEN: ÜBERLEGUNGEN ZUR INHALTSVALIDITÄT

von Jonas Wibowo und Tim Heemsoth

ZUSAMMENFASSUNG | In diesem Beitrag werden Eckpunkte einer systematischen Begründung und Prüfung der Inhaltsvalidität theoretisch hergeleitet und anhand einer Analyse von vier ausgewählten Projekten zur Testung von fachdidaktischem Wissen für diesen Bereich spezifiziert. Demnach sollten (theoretische) Itemmengen durch (a) fachinhaltliche Themen, (b) fachdidaktische Wissensfacetten und (c) kognitive Prozesse begründet werden. Zudem sollte die Repräsentation der Itemmenge durch realisierte Items – auch nach Veränderungen im empirischen Prozess – reflektiert werden. Letzteres gilt insbesondere nach Expert*innenbefragungen und Itemselektionsprozessen im Zuge von Konstruktanalysen. Die Itemmenge für die eigene vorliegende Testkonzeption wird durch (a) acht Bewegungsfelder, (b) drei Wissensfacetten und (c) ein handlungsnahes, kognitives Anforderungsprofil bestimmt. Die 40 Items wurden im Rahmen einer Expert*innenbefragung (N = 28) zufriedenstellend hinsichtlich der Relevanz der Plausibilität eingeschätzt. Des Weiteren ordnen die Expert*innen die Items mehrheitlich im Sinne der theoretisch begründeten Systematik zu.

Schlüsselwörter: Inhaltsvalidität; fachdidaktisches Wissen; Sportunterricht; Expertenbefragung; systematische Itemkonstruktion

TESTING PEDAGOGICAL CONTENT KNOWLEDGE OF PHYSICAL EDUCATION TEACHERS: THOUGHTS ON CONTENT VALIDITY

ABSTRACT | In this article, we deduce key points that allow justifying and testing content validity of tests for pedagogical content knowledge. We specify these key points on the basis of four selected projects. Accordingly, the (theoretical) item universe should be specified through (a) a subject-matter dimension, (b) a dimension that contains facets of pedagogical content knowledge, and (c) cognitive processes. Beyond, the representativeness of the item universe with regard to concrete items should be considered during the processes test construction and after changes due to empirical validation steps such as expert reviews and item selection processes during construct analyzes. The item universe of the present test is specified by (a) eight movement-fields, (b) three knowledge facets, and (c) action related cognitive demands. Twenty-eight experts judged the 40 items with satisfactory relevance and satisfactory plausibility. Furthermore, the experts predominantly assigned the items in accordance with the theoretical construct.

Key Words: Content validity; Pedagogical Content Knowledge; Physical Education; Expert Review; systematic Itemconstruction

FACHDIDAKTISCHES WISSEN VON SPORTLEHRER*INNEN TESTEN: ÜBERLEGUNGEN ZUR INHALTSVALIDITÄT

1 | EINLEITUNG

Im Rahmen kompetenzorientierter Ansätze der Lehrkraftbezogenen Professionsforschung wird dem fachdidaktischen Wissen (FDW) von Lehrer*innen ein hoher Stellenwert zum erfolgreichen Unterrichten beigemessen. Es umfasst das Wissen über das „Verständlichmachen von Fachinhalten“ (Krauss, 2011, S. 182) und wird teilweise als Integration von Fachwissen und allgemeinem pädagogischen Wissen angenommen. Die steigende Popularität von FDW als Forschungsgegenstand ergibt sich vor allem aus Befunden, die FDW als starken Prädiktor für Lernzuwächse bei Schüler*innen identifizieren (Kunter, Klusmann, Baumert, Richter, Voss, & Hachfeld, 2013). In der angenommenen Kausalkette, beginnend bei der Lehrer*innenbildung hin zu Schulleistungen, gilt das FDW daher – neben anderen Bestandteilen des Professionswissens und weiteren Kompetenzbereichen, wie Überzeugungen, Werthaltungen, motivationalen Orientierungen und selbstregulativen Fähigkeiten – als besonders bedeutsam (Krauss, 2011). Vor diesem Hintergrund werden mittlerweile in vielen Fächern Tests zur Messung von FDW entwickelt; im Fach Sport sind entsprechende Arbeiten jedoch noch immer im Entwicklungsstadium (Heim & Sohnsmeier, 2015; Vogler, Messmer, Wibowo, Heemsoth & Meier, 2018). Dabei gilt es, bei der Entwicklung entsprechender Tests im Vorfeld der Untersuchungen stets die Inhaltsvalidität als zentrales Gütekriterium zu prüfen. Auf diese Weise soll gewährleistet werden, dass sich die Testentwicklung an einer klaren Konzeptspezifikation orientiert, der wiederum bedeutsame Aspekte des zu erfassenden Konstrukts in seiner Dimensionalität zugrunde liegen (Döring & Bortz, 2016). Vielfach finden sich jedoch Bruchstellen in der theoretisch-argumentativen Begründung der Konzeption eines Tests auf der Basis der ausgewählten Theorien bzw. im Rahmen der empirischen Operationalisierung durch konkrete Items. Dies wird als „Theorie-Empirie-Problem“ der Kompetenzmessung bezeichnet (Zlatkin-Troitschanskaia & Seidel, 2011).

Ziel des vorliegenden Beitrags ist es, zunächst allgemeine Vorgehensweisen einer systematischen Herleitung und Prüfung der Inhaltsvalidität aufzuzeigen (Kap. 2) und darzustellen inwieweit diese in bisherigen Tests zum fachdidaktischen Wissen berücksichtigt wurden (Kap. 3). Dafür werden sowohl fachfremde als auch sportunterrichtsbezogene Beiträge betrachtet. Die Befunde laden dazu ein, eine für FDW-Tests spezifische Heuristik zur systematischen Begründung und Prüfung der Inhaltsvalidität abzuleiten. Diese wird in den Kap. 4 bis 6 durch eine eigene Testkonzeption zum FDW von Sportlehrer*innen verdeutlicht. Insgesamt werden drei Aspekte hinsichtlich der Inhaltsvalidität gesondert berücksichtigt: die Konzeptspezifikation (Kap. 4), die Operationalisierung der Items (Kap. 5) und Einschätzungen von Expert*innen und die Darlegung der Kodierung offener Antwortformate auf Basis einer Expert*innenbefragung (Kap. 6). In Kap. 7 werden Konsequenzen für weitere Forschung diskutiert.

2 | VORGEHENSWEISEN ZUR HERLEITUNG UND PRÜFUNG DER INHALTSVALIDITÄT

Die rationale und nachvollziehbare Testkonstruktion ist eine zentrale Voraussetzung für die Qualität eines Tests (Döring & Bortz, 2016; Klauer, 1984; Messick, 1995). Klauer (1984) betont, dass dieser Prozess regelgeleitet und explizit erfolgen müsse, damit nachvollzogen werden kann, ob ein Test das misst, was er zu messen vorgibt.

Als zentrale Merkmale der *Inhaltsvalidität* benennt Messick (1995) erstens die *inhaltliche Relevanz* und zweitens die *Repräsentativität*. Bezüglich der *inhaltlichen Relevanz* geht es um die theoretische Begründung des fokussierten Inhaltsbereichs, was auch als „Konzeptspezifikation“ (Döring & Bortz, 2016, S. 223) bezeichnet wird. Nach Klauer (1984) sind für die Spezifikation des fokussierten Inhaltsbereichs vor allem eine Situations- und eine Verhaltensdimension von Interesse. Die *Situationsdimension* wird dabei durch den inhaltlichen Sachverhalt des Tests und die *Verhaltensdimension* durch die vom Probanden geforderten Verhaltensweisen bestimmt. Als Beispiel für eine Verhaltensdimension nennt Klauer (1984) – im Sinne der Bloomschen Taxonomie – kognitive Prozesse wie Kennen, Verstehen und Anwenden. Er räumt jedoch ein, dass je nach Inhaltsbereich eines Tests auch andere bzw. weitere Dimensionen eine Rolle spielen können. Zentral ist, dass das (mathematische) Produkt der Anzahlen an Ausprägungen einer jeden Dimension eine *Itemmenge* bestimmt. Werden etwa drei Situationsdimensionen und fünf Verhaltensdimensionen bestimmt, so ergibt sich für den Test eine Itemmenge, die aus 15 verschiedenen Komponenten besteht, die wiederum auf potenzielle (theoretische) Bereiche eines Tests hinweisen. Neben der theoretischen Konzeptspezifikation wird die inhaltliche Relevanz zudem durch Urteile von Expert*innen gestützt, die etwa die Relevanz und Plausibilität eines entwickelten Tests einschätzen (Döring & Bortz, 2016, S. 445; Hartig, Frey & Jude, 2012, S. 149).

Repräsentativität liegt vor, wenn die identifizierte Itemmenge durch die entwickelten Items hinreichend berücksichtigt wird, der Test also „eine repräsentative Stichprobe der Itemmenge darstellt“ (Klauer, 1984, S. 2). Dafür gilt es, vor dem Schritt der Operationalisierung zu spezifizieren, welche Bereiche der Itemmenge wie stark repräsentiert werden sollen. Operationalisierung bedeutet, dass aus der potenziellen Itemmenge reale Items konstruiert werden (Döring & Bortz, 2016, S. 223), deren Gesamtheit als Itempool bezeichnet wird. Damit auch dieser Schritt als *systematisch* bezeichnet werden kann, sollten Itemformen und Transformationsregeln transparent festgelegt werden: Dabei wird die Itemform im Wesentlichen durch die Art des Antwortimpulses (mit mehr oder weniger Kontextinformationen) und durch das geforderte Antwortformat (offen vs. geschlossen) bestimmt (s. auch Döring & Bortz, 2016). Klauer (1984) mahnt bezüglich der Itemform an, dass diese in Kohärenz mit der Verhaltensdimension festgelegt werden soll; so empfiehlt etwa Bloom (1976), für den kognitiven Prozess der *Anwendung* Problemlöseitems einzusetzen, die durch eher umfangreiche Kontextinformationen und offene Antwortformate gekennzeichnet sind. Transformationsregeln haben zum Ziel, „die Übertragung des Inhalts in ein Item der vereinbarten Form eindeutig zu machen“ (Klauer, 1984, S. 8). Der Aspekt der Repräsentativität ist auch im weiteren Verlauf des Entwicklungsprozesses zu berücksichtigen. Jede Veränderung der Beziehung zwischen (theoretischer) Itemmenge und (realisiertem) Itempool sollte reflektiert werden – etwa eine Itemreduktion nach einer Expert*innenbefragung oder nach einer empirischen Konstruktprüfung.

Neben der Inhaltsvalidität gelten die *Konstrukt-* und die *Kriteriumsvalidität* als zentrale Validitätsformen (Bortz & Döring, 2016). Die *Konstruktvalidität* wird gestützt, wenn angenommene Zusammenhänge mit verwandten Konstrukten bestätigt werden können oder die dimensionale Struktur des Tests durch Faktorenanalysen gestützt werden kann. Die *Kriteriumsvalidität* wird gestützt, wenn das Instrument mit einem Außenkriterium korreliert, das durch das Instrument selbst vorhergesagt werden soll. Während die Konstrukt- und die Kriteriumsvalidität erst im Verlauf der Testentwicklung auf der Basis eines gut spezifizierten Konstrukts getestet werden kann, spielt die Inhaltsvalidität schon zu Beginn der Testentwicklung eine zentrale Rolle.

Da die genannten Aspekte zur Begründung und Prüfung der Inhaltsvalidität zunächst allgemeiner Natur sind, sollen sie im Folgenden auf vier Projekte zur Testung von FDW angewandt werden: *TEDS-LT*, *FALKO/COACTIV* und *zwei auf Sportlehrer*innen bezogene Ansätze aus Basel bzw. Köln*. Dabei soll herausgearbeitet werden, inwiefern erstens die inhaltliche Relevanz und zweitens die Repräsentativität der Itemmenge zum Tragen kommt.

3 | ANSÄTZE ZUM FACHDIDAKTISCHEN WISSEN IM VERGLEICH

In der Studie *TEDS-LT* („Teacher Education Development Study – Learning to Teach“) wurde für das Fach Mathematik ein FDW-Test eingesetzt, der eine Weiterentwicklung der Tests aus vorherigen Studien der Autor*innen darstellt (Buchholtz, Kaiser & Blömeke, 2014). Die Itemmenge wird durch zwei Dimensionen bestimmt: zwei Wissensfacetten¹ und drei kognitive Prozesse (Buchholtz et al., 2014, S. 111-113). Die Autor*innen unterscheiden eine *stoffdidaktische Wissensfacette*, die eher fachwissenschaftlich-mathematisch fundiert ist, von einer *unterrichtsbezogenen mathematikdidaktischen Wissensfacette*, die eher auf fachdidaktische, allgemein didaktische und erziehungswissenschaftliche Ansätze Bezug nimmt. Beide Dimensionen werden mit Bezug auf Anderson und Krathwohl (2001) mit den drei kognitiven Prozessen (a) *Erinnern und Abrufen*, (b) *Verstehen und Anwenden* und (c) *Bewerten und Generieren von Handlungsoptionen* gekreuzt. Dadurch entsteht eine theoretische Itemmenge aus sechs Komponenten, die wiederum in einem weiteren Schritt in 30 Komponenten ausdifferenziert wird. Inwiefern die entwickelten 59 Items alle 30 Komponenten repräsentieren, ist nur eingeschränkt nachvollziehbar, wobei die Autor*innen darauf hinweisen, dass die Itemmenge aufgrund von Selektionsprozessen nicht gleichmäßig besetzt wurde. Mit einem Verhältnis von 23 zu 30 Items ist die unterrichtsdidaktische Wissensfacette etwas stärker vertreten als die stoffdidaktische, die drei kognitiven Prozesse werden fast zu gleichen Teilen abgedeckt (Buchholtz & Kaiser, 2013, S. 116-117). Trotz theoretischer Begründung einer Itemmenge mit sechs bzw. 30 Komponenten wird in der Analyse der empirischen Daten eine zweifaktorielle Struktur im Sinne der Wissensfacetten überprüft. Konsequenzen von Selektionsprozessen für die inhaltliche Validität werden nicht diskutiert. Es werden sowohl geschlossene, Multiple-Choice-Items als auch offene Items eingesetzt, wobei keine dieser Itemformen einer bestimmten Komponente bzw. Dimension (z. B. einem konkreten kognitiven Prozess) der Itemmenge zugewiesen ist. Auch wenn sich der Test zum Fachwissen in den Bereichen Arithmetik und Algebra konzentriert,

1 Mit *Wissensfacetten* sind hier Kategorien gemeint, die meist mit Bezug auf Shulman (1986; 1987) als einzelne Bestandteile von FDW dargestellt werden.

ist unklar, ob dies auch für den FDW-Test gilt – eine fachinhaltliche Dimension der Itemmenge wird hier nicht explizit benannt.

In der Studie *FALKO* („Fachspezifische Lehrerkompetenzen“, Krauss, Lindl, Schilcher, Fricke et al., 2017), die konzeptionell auf der Studie *COACTIV* („Cognitive activation in the classroom“, Kunter et al., 2011) aufbaut, werden verschiedene Fächer in einem gemeinsamen Ansatz untersucht. Die Itemmenge wird primär durch drei Wissensfacetten bestimmt: (1) *Erklären und Repräsentieren*, (2) *Schülerkognitionen und fachliche Fehlvorstellungen* und (3) *Potenzial von Lernmaterialien*, wobei die ersten beiden Wissensfacetten für alle Fächer verwendet werden und die dritte Facette stärker fachbezogen variiert wird. So wird etwa für Deutsch, Englisch und Latein das Potenzial von Texten und (inter-)kulturellen Gegenständen und in Physik das Potenzial von Messen und Experimentieren betrachtet (Krauss, Lindl, Schilcher & Tepner, 2017, S. 28). Für das Fach Mathematik wurde auf die Daten und Ergebnisse der *COACTIV-Studie* zurückgegriffen: Die Wissensfacetten werden hier über 12 (Facette 1), 7 (Facette 2) und 4 (Facette 3) Items abgebildet (Krauss et al., 2011, S. 139). Eine Reflexion der unterschiedlichen Gewichtungen der Items im Hinblick auf die Repräsentation der Itemmenge findet nicht statt. Auf Basis der empirischen Daten wird das FDW zudem schlussendlich auf einem Faktor abgebildet, der alle 23 Items umfasst (Krauss et al., 2011, S. 149). Während für das Fach Mathematik keine fachinhaltliche Dimension spezifiziert wurde (Krauss et al., 2011), werden bspw. für das Fach Deutsch die einzelnen Items den Inhaltsbereichen Literaturdidaktik (5 Items) und Sprachdidaktik (8 Items) zugeordnet. Bei der Konstruktion des finalen Itempools für den Deutsch-FDW-Test haben die Autor*innen bei der Selektion darauf geachtet, dass die theoretisch begründeten Inhalte nach wie vor repräsentiert bleiben und haben auch entsprechende Items nachentwickelt (Pissarek & Schilcher, 2017, S. 80-88). Die in *COACTIV* und *FALKO* vorwiegend offenen Items werden von den Autoren als „hypothetische Unterrichtsszenarios“ bezeichnet (Krauss, Lindl, Schilcher & Tepner, 2017, S. 37), die eher eine unterrichtliche Handlungsfähigkeit und somit eher anwendungsbezogene kognitive Prozesse erfordern. Die Fokussierung handlungsnaher kognitiver Anforderungen dürfte zur Vorhersage unterrichtlicher Wirkungen Vorteile zeigen gegenüber der Testung expliziten Wissens durch Erinnerungsaufgaben (s. o.) und eher metakognitiven Aufgaben durch Bewertungen (Baumgartner, 2018; Lüders, 2012). Diese Handlungsorientierung wird in *FALKO* theoretisch aus den Arbeiten Shulmans (1986) begründet.

In Anlehnung an Scherler (2004) werden im sportdidaktischen *Basler Ansatz* drei unterrichtspraktische Handlungen von Sportlehrer*innen fokussiert, die wiederum zur Beschreibung von Wissensfacetten herangezogen werden: (1) *Inhalte präsentieren*, (2) *Bedingungen organisieren* und (3) *mit Schüler*innen interagieren* (Vogler, Messmer & Allemann, 2017, S. 336). Diese drei Wissensfacetten wurden durch eine Delphi-Befragung überprüft, in 16 Komponenten ausdifferenziert und nach einer empirischen Validierung auf neun reduziert (Facette 1 mit zwei Komponenten, Facette 2 mit zwei Komponenten, Facette 3 mit fünf Komponenten). Die neun Komponenten umfassende Itemmenge wird durch einen Itempool von 28 Items repräsentiert, wobei jeder Komponente zwei bis drei Items zugeordnet werden (Vogler et al., 2017, S. 340). Als Itemform werden Textvignetten verwendet, die einen umfangreichen Kontext für eine konkrete Unterrichtssituation formulieren und offen beantwortet werden müssen. Es werden somit – ver-

gleichbar wie in *FALKO/COACTIV* – anwendungsbezogene kognitive Prozesse angesprochen, dies jedoch im Rahmen noch ausführlicherer Kontextualisierung. Eine fachinhaltliche Fokussierung ist nicht festzustellen.

Im ebenfalls auf Sportlehrer*innen bezogenen *Kölner Ansatz* wird die Itemmenge einerseits über eine Dimension mit zwei Wissensfacetten in Anlehnung an Shulman (1986) und andererseits über eine Dimension kognitiver Prozesse bestimmt, wie dies auch in *TEDS-LT* vorgenommen wurde (Meier, 2018). Über eine strukturierende, qualitative Inhaltsanalyse ausgewählter universitärer Ausbildungscurricula der Sportlehrer*innenbildung werden die beiden Wissensfacetten (1) *Wissen über Repräsentationsformen und Instruktionsstrategien* und (2) *Wissen über inhaltspezifische Lernschwierigkeiten und Schülerkognitionen* spezifiziert und wiederum in sechs (Facette 1) bzw. drei (Facette 2) Subfacetten unterteilt. Diese neun Subfacetten bilden mit den drei kognitiven Prozessen (a) *Erinnern*, (b) *Verstehen/Analysieren* und (c) *Kreieren* eine Itemmenge mit 27 Komponenten. Für diese Itemmenge wurden 140 Items erstellt, wobei unklar ist, inwiefern der Itempool die 27 Komponenten gleichermaßen repräsentiert. Für jede Komponente wurden offene und geschlossene Items konstruiert. Die Formulierung des Beispielitems zeigt, dass keine kontextualisierte Erfassung des FDW im Sinne hypothetischer Unterrichtsszenarios (s. o.) vorgenommen wurde (Meier, 2018, S. 80). Nach einem Expert*innenrating wurde der Itempool auf 81 reduziert; inwiefern damit die theoretisch hergeleitete Itemmenge noch gleichermaßen repräsentiert wird, ist nicht nachvollziehbar. Weitere Validierungsschritte wurden (noch) nicht berichtet.

ZWISCHENFAZIT

Die Forderung, dass Tests eine theoretisch hergeleitete Itemmenge repräsentieren sollen (s. Kap. 2), wird in den dargestellten Ansätzen in unterschiedlichem Maße berücksichtigt. Für die Konstruktion der Itemmengen von FDW-Tests werden in den dargestellten Studien drei Dimensionen berücksichtigt: (i) *Wissensfacetten*, (ii) *kognitive Prozesse* und (iii) *eine fachinhaltliche Dimension*.

Die Wissensfacetten werden in den analysierten Studien unterschiedlich begründet. Mit Bezug auf die Arbeiten Shulmans (1986; 1987) wurden unterschiedliche Auslegungen betrachtet. In *TEDS-LT* scheint sich die angenommene zweidimensionale Struktur auch in den Daten widerzuspiegeln; in *FALKO* und *COACTIV* wird, trotz einer angenommenen dreidimensionalen Struktur, ein G-Faktor-Modell berichtet. Die in Anlehnung an die *TEDS-Studien* übernommene Struktur des *Kölner Ansatzes* für Sportlehrpersonen wurde noch nicht überprüft. Einzig der *Basler Ansatz* bezieht sich bei der Konstruktion der Wissensfacetten auf das fachspezifische Modell von Scherler (2004), jedoch ebenfalls ohne eine empirische Prüfung der Dimensionalität anzubieten. Weiterhin werden in den Ansätzen entweder explizit drei unterschiedliche kognitive Prozesse mit den Wissensfacetten gekreuzt (s. *TEDS-LT* und *Kölner Ansatz*) oder es wird ein bestimmter Prozess fokussiert – in den *FALKO-Studien* und bei der *Basler Arbeitsgruppe* werden durch die Kontextualisierung eher anwendungsbezogene Anforderungen angesprochen.

Die unterschiedlichen Konzeptualisierungen sind insofern von besonderer Bedeutung, da sich aus den verschiedenen Ansätzen Unterschiede für die Verortung von FDW im Gefüge pro-

fessioneller Kompetenz von Lehrer*innen ergeben. Trotz verschiedener Definitionen gilt Kompetenz noch immer als „messy construct“ (Blömeke et al., 2015, S. 4), eine Feststellung, die auch auf das FDW als einen Bereich professioneller Kompetenz übertragen werden kann (Depaepe et al., 2013; Ward & Ayzazo, 2016). Um dieser Problematik zu begegnen, sind in der jüngeren Vergangenheit Modelle in den Fokus gerückt, die versuchen, unterschiedliche Verständnisse von Kompetenz zu integrieren (Blömeke et al., 2015; Blömeke & Kaiser, 2017; Santagata & Yeh, 2016; für das Fach Sport Baumgartner, 2018). Dabei werden insbesondere analytische Ansätze, die Kompetenz als Summe kognitiver und affektiv-motivationaler Dispositionen fassen, und holistische Ansätze, die Kompetenz von der tatsächlich erbrachten Performanz abzuleiten versuchen, miteinander verbunden. Als eine Art Konsens kann dabei angesehen werden, dass Kompetenz auf die Bewältigung realer Anforderungssituationen abzielt und einer solchen Bewältigung kognitive, affektiv-motivationale und metakognitive Eigenschaften, aber auch situationsspezifische Fähigkeiten zugrunde liegen. Blömeke et al. (2015) verweisen weiterhin darauf, dass es sinnvoll sei, Ebenen professioneller Kompetenz entlang der Nähe zu echten Anforderungssituationen zu differenzieren. Eine solche Verortung professioneller Kompetenz im Allgemeinen erscheint auch hinsichtlich des FDW im Speziellen plausibel: So führen Kontextualisierungen in Items dazu, dass die zu erbringende Testleistung nicht allein als Disposition gewertet werden kann, sondern sich der Performanz annähert – gleichwohl sie diese nicht gänzlich abbildet. Solch unterschiedliche Operationalisierungen lassen sich auch in den bereits diskutierten Tests ausmachen: In *TEDS-LT* und auch im *Kölner Ansatz* werden stark kontextfreie Items verwendet, z. B. bei der *Reproduktion von Kenntnissen* über fachspezifische Zugangsweisen. Dies hat zur Folge, dass das FDW eher weit entfernt von echten Anforderungssituationen konzeptualisiert wird. Stärker kontextualisiert sind Items im Bereich der kognitiven Prozesse *Verstehen und Anwenden* und *Bewerten und Generieren von Handlungsoptionen*. Hier werden Informationen zu einer konkreten unterrichtlichen Anforderungssituation dargestellt und durch die Item-Prompts werden Aktivitäten, wie z. B. Diagnostizieren oder das Konstruieren von Aufgaben, angeregt (Buchholtz & Kaiser, 2013). In einem noch stärkeren Ausmaß bietet der *Basler Ansatz* Kontextinformationen im Rahmen umfänglicher Fallbeschreibungen an (Vogler et al., 2017). Wir nehmen an, dass ausführliche Kontextualisierung dazu führt, dass insgesamt handlungsnähere und somit ökologisch validere Operationalisierungen des FDW vorliegen. Die Beschreibung der verwendeten Kontexte bzw. die andersartig hergestellte Handlungsnähe (z. B. durch die Simulation von Handlungsdruck) sollte daher in der Darstellung der Testkonzeption und der Operationalisierung eines FDW-Tests ausführlich und explizit erfolgen. Durch eine solche Darstellung würde die Vergleichbarkeit von FDW-Tests entlang der von Blömeke et al. (2015) geforderten Nähe/Distanz zu echten Anforderungssituationen erleichtert werden.

Die fachinhaltliche Dimension der Itemmengen spielte bisher eine eher untergeordnete Rolle. Lediglich in dem *Deutsch-FDW-Test* von *FALKO* werden spezifische Inhaltsbereiche berichtet (Pissarek & Schilcher, 2017). Bemerkenswert ist dieser Punkt daher, weil die von Shulman (1986) eingeforderte und viel zitierte Fachlichkeit in der Kompetenzmessung anscheinend vor den Fachinhalten aufhört. Implizit schwingt mit der ausbleibenden Spezifikation der fachlichen Inhalte die Annahme mit, dass sich Facetten des fachdidaktischen Wissens unabhängig vom unterrichteten

Inhalt messen lassen. Wenn etwa (nur) das Wissen über Fehlvorstellungen in der Arithmetik getestet wird (Buchholtz & Kaiser, 2013), wird zumindest implizit angenommen, dass ein vergleichbares Wissen über Fehlvorstellungen in anderen Bereichen (Algebra, Geometrie usw.) vorhanden ist. Sofern der Test in seiner Reichweite nicht eingegrenzt wird, ist diese Annahme begründungsbedürftig.

Die Operationalisierung durch einen konkreten Itempool kann in allen berücksichtigten Studien als eher intransparent bezeichnet werden, da nicht ersichtlich ist, inwiefern der verwendete Itempool die theoretisch begründete Itemmenge repräsentativ abbildet. Die Nachkonstruktion von Items zur gleichmäßigen Abdeckung der konstruierten Itemmenge nach Itemreduktionen im *Deutsch-FDW-Test* von FALCO ist ein positives Beispiel (Pissarek & Schilcher, 2017). Jedoch ist unklar, warum die einzelnen Komponenten der Itemmengen durch unterschiedliche Mengen an Items repräsentiert werden. Mit Bezug auf Klauer (1984) müssten daher alle betrachteten Studien als zumindest eingeschränkt inhaltsvalide bezeichnet werden. Ob Kompromisse aufgrund statistischer Verfahrensweisen und Elementarisierungsnotwendigkeiten diese eingeschränkte Inhaltsvalidität rechtfertigen, müsste diskutiert werden.

Die oben formulierten Vorgehensweisen zur Herleitung und Prüfung der Inhaltsvalidität eines FDW-Tests wird im Folgenden entlang der Schritte *Konzeptspezifikation*, *Operationalisierung* und *Expert*innenbefragung* anhand eines Beispiels einer eigenen Testentwicklung für das FDW von Sportlehrer*innen verdeutlicht.

4 | KONZEPTSPEZIFIKATION IM EIGENEN ANSATZ

Die fachinhaltliche Dimension der Itemmenge im eigenen Ansatz wird durch die Wissensfacetten (1) *Lernende unterstützen*, (2) *Schwierigkeiten erklären* und (3) *Vorgehensweisen formulieren* strukturiert, die wiederum mit acht Bewegungsfeldern gekreuzt werden. Durch die Verwendung von Itemformen mit hypothetischen Unterrichtsszenarios wird ein handlungsnahes kognitives Anforderungsprofil fokussiert. Damit ist gemeint, dass FDW hier nicht im Sinne einer Reproduktion deklarativen Wissens erfasst wird, sondern unter Simulation von Ausschnitten unterrichtsnaher Kontexte.

Die ersten beiden Wissensfacetten, *Lernende unterstützen* und *Schwierigkeiten erklären*, fokussieren handlungsnaher Anforderungen in der Lehrer*innen-Schüler*innen-Interaktion im laufenden Unterricht und können als Kern der originalen Konzeption von FDW bei Shulman angesehen werden (Depaepe et al., 2013). Die vielfache Verwendung der beiden Wissensfacetten in Forschungsarbeiten kann als Hinweis darauf gedeutet werden, dass sich diese beiden Facetten als zentrale Anforderungen bewährt haben. Bezüglich der Verortung im Rahmen professioneller Kompetenz ist anzunehmen, dass in diesen Wissensfacetten kognitive Wissensbestände, aber auch die situationsspezifischen kognitiven Fähigkeiten des Interpretierens und Entscheidens von Bedeutung sind. Darüber hinaus erscheint uns die Verwendung dieser beiden auch in anderen Fächern verwendeten Facetten insofern als sinnvoll, als dass eine Anschlussfähigkeit ermöglicht wird.

Die dritte Facette, *Vorgehensweisen formulieren*, fokussiert umfänglichere planerische Anforderungen an Lehrpersonen. Planerische Anforderungen wurden auch in *TEDS-LT* (s. o.) integriert,

um in dem Konzept des FDW, neben interaktionsbezogenen Situationen, auch planerische Situationen zu berücksichtigen, die gleichsam in der Diskussion um professionelle Kompetenzen bedacht werden und zwar fachübergreifend (Borko, Roberts & Shavelson, 2008; Bremerich-Voss, Buchholtz & König, 2016), wie auch spezifisch für das Fach Sport (Heemsoth, 2016). Die von Prohl (2012b) geforderte Öffnung bezüglich der Vermittlung von Inhalten im Sportunterricht schlägt sich in der dritten Wissensfacette derart nieder, dass weniger eine direktive technikzentrierte Vermittlung im Vordergrund steht, sondern eine Verschränkung strukturierender und Freiheiten ermöglichender Aspekte. Dabei soll Öffnung jedoch nicht missverstanden werden als maximale Selbstbestimmung der Schüler*innen – dies würde vermutlich zumindest bei einem Teil der Schüler*innen zu einer Überforderung führen (Kirschner, Sweller & Clark, 2006). Vielmehr ist eine Öffnung gemeint, die weg von der Vermittlung normierter Lösungen hin zu einem aktiven Gestaltungsraum für Schüler*innen gedacht ist und einen *Spielraum für Entscheidungen* ermöglicht (Bähr & Wibowo, 2018). Insgesamt leiten wir für diese Facette daher die Notwendigkeit von Entscheidungsfreiheiten (der Lernenden) und andererseits die Notwendigkeit einer fachlichen Strukturierung des Lerngegenstandes (durch die Lehrkraft) ab.

Die Festlegung auf Bewegungsfelder zur Bestimmung der fachinhaltlichen Dimension der Itemmenge erfolgt vor allem aus fachdidaktischen Überlegungen zu einem erziehenden Sportunterricht (Beckers, 2013; Prohl, 2012a). Der sogenannte *Doppelauftrag* einer Erziehung zum Sport und einer Erziehung durch Sport zielt auf eine „Bewegungsbildung im Horizont allgemeiner Bildung“ (Prohl, 2012a, S. 77). Dies bedeutet, dass über die sachliche Auseinandersetzung mit bewegungsbezogenen Inhalten (Erziehung zum Sport) fachspezifische Bildungsprozesse angestoßen werden sollen, die im gleichen Zuge die Entwicklung der Persönlichkeit ermöglichen (Erziehung durch Sport). Der Verweis auf die Autonomie und Nicht-Intendierbarkeit von Bewegungsbildung als ästhetischem Kern des Sportunterrichts hat zur Folge, dass auf den Ebenen von Inhalten und Vermittlung eine fachlich und didaktisch strukturierte Öffnung stattfinden muss (Prohl, 2012b). Auf der Inhaltsebene werden in diesem Kontext Bewegungsfelder als fachliche Inhaltsstruktur diskutiert, die die kulturellen Erscheinungsformen von Bewegung, Spiel und Sport nach strukturellen Gesichtspunkten bündeln (Laging, 2008; Scheid & Prohl, 2012) und heute eine starke Position im Rahmen der Sportlehrpläne einnehmen (Prohl & Krick, 2006). Daraus ist abzuleiten, dass die Struktur der Bewegungsfelder zumindest auf normativer Ebene fachdidaktisches Wissen von Sportlehrkräften (mit-)strukturiert. Insgesamt werden daher Bewegungsfelder aufgegriffen, die in dieser oder ähnlicher Form Einzug in die sportdidaktische Lehre an der Universität als auch in den Lehrplänen der Länder erhalten haben (Scheid & Prohl, 2012, S. 134): (1) Laufen, Springen, Werfen; (2) Bewegen an und mit Geräten; (3) Spielen in und mit Regelstrukturen; (4) Den Körper trainieren, die Fitness verbessern; (5) Bewegung gestalten; (6) Mit/gegen Partner kämpfen; (7) Bewegen im Wasser; (8) Rollen, Gleiten, Fahren. Obwohl in der Praxis des Sportunterrichts eine ungleiche Gewichtung der Bewegungsfelder stattzufinden scheint (Gerlach, Kussin & Brandl-Bredenbeck, 2006), wird hier aufgrund der prinzipiellen Gleichberechtigung der Bewegungsfelder eine gleichmäßige Repräsentation in der Itemmenge vorgenommen. Dem liegt die Annahme zugrunde, dass Sportlehrer*innen nicht nur die (vermeintlich) dominanten Inhalte des Sportunterrichts unterrichten können sollten, sondern das Fach in seiner gesamten inhaltlichen Breite.

Darüber hinaus wird über die Ausbalancierung vermieden, dass ein (biografisch geprägtes) hohes Vorwissen in einem bestimmten Bewegungsfeld zu unverhältnismäßigen Verzerrungen gegenüber anderen Bewegungsfeldern führt.

Aus den dargestellten Aspekten ergibt sich eine Itemmenge mit 24 Komponenten (8 Bewegungsfelder x 3 Wissensfacetten). Aufbauend auf der skizzierten Itemmenge, wird im Folgenden die Operationalisierung der konkreten Items anhand von Transformationsregeln dargestellt.

5 | OPERATIONALISIERUNG DER ITEMMENGE

Die theoretisch hergeleitete Itemmenge wird durch einen Itempool von 40 Items repräsentiert. Die Items verteilen sich in gleichem Maße auf die acht Bewegungsfelder. Für alle Items werden hypothetische Unterrichtsszenarios verwendet, um handlungsnah kognitive Anforderungen zu erstellen. Die Itemformen wurden wie folgt festgelegt: Für die Wissensfacetten *Lernende unterstützen* und *Schwierigkeiten erklären* werden jeweils zwei geschlossene Multiple-Choice-Items pro Bewegungsfeld mit fünf Antwortalternativen formuliert, wobei jede Antwort mit richtig oder falsch zu bewerten ist. Geschlossene Items wurden deshalb gewählt, da sich für beide Facetten anhand der praxisnahen Bezugsliteratur plausibel erscheinende Antworten formulieren ließen. Für die Wissensfacette *Vorgehensweisen formulieren* wird ein offenes Item pro Bewegungsfeld formuliert. Offene Antwortformate wurden gewählt, da eine Vielzahl an Möglichkeiten besteht, für ein Unterrichtsziel einen Spielraum für Entscheidungen einzuräumen und den Lerngegenstand fachlich zu strukturieren.

Für die Operationalisierung der Itemmenge gelten weiterhin folgende Transformationsregeln: Für die Items der Wissensfacette *Lernende unterstützen* wird stets ein unterrichtliches Ziel und/oder eine Lernschwierigkeit dargestellt. Nachfolgend werden fünf Maßnahmen formuliert und es muss entschieden werden, inwieweit diese Maßnahmen geeignet sind, um das unterrichtliche Ziel zu erreichen bzw. die Lernenden zu unterstützen (s. Abb. 1, links oben). Für Items der Wissensfacette *Schwierigkeiten erklären* wird ähnlich wie bei der ersten Wissensfacette eingangs eine unterrichtliche Aufgabe oder eine Lernschwierigkeit dargestellt. Hier muss im Anschluss jedoch entschieden werden, inwieweit mit der Aufgabe einzelne Probleme auftreten können oder inwieweit Lernschwierigkeiten auf bestimmte Erklärungen zurückgeführt werden können (s. Abb. 1, rechts oben). Für die Entwicklung der Items in den ersten beiden Facetten wurde vor allem auf praxisnahe Literatur zurückgegriffen. Sofern verfügbar, wurden empirisch fundierte Beiträge bevorzugt².

In der Wissensfacette *Vorgehensweisen formulieren* werden sowohl ein unterrichtliches Ziel formuliert als auch zwei unterschiedliche Vorgehensweisen skizziert, die sich im Grad ihrer Lenkung durch die Lehrkraft unterscheiden. In der ersten Vorgehensweise wird ein durch die Lehrkraft stark strukturiertes und gelenktes Vorgehen und in der zweiten Vorgehensweise ein offenes, kaum strukturiertes Vorgehen skizziert. Im Anschluss soll eine Vorgehensweise formuliert

² Es ist jedoch anzumerken, dass die fach- und inhaltspezifische Literatur zu Lernschwierigkeiten und Schüler*innenkognitionen bei Weitem nicht so stark empirisch fundiert ist, wie es bspw. für das Fach Mathematik (Aufschnaiter & Rogge, 2010) der Fall ist. Die Güte der entwickelten Items ist daher zukünftig auch im Rahmen weiterer prädiktiver Analysen zu untersuchen.


<p>In Ihrer 7. Klasse soll der Tic-Tac erlernt werden, eine Basistechnik im Parkoursport, um z. B. über einen Kasten zu springen (vgl. Abb.). Einem Schüler gelingt es nicht, sich ausreichend von der Wand abzudrücken. Welche der folgenden Vorübungen ist geeignet, um den Lernprozess zu unterstützen?</p>  <p style="text-align: right;">Ja Nein</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <td style="padding: 2px;">Handstütz auf dem Kasten, ein Schritt gegen die Wand und Aufknien auf dem Kasten.</td> <td style="text-align: center; padding: 2px;"><input checked="" type="checkbox"/></td> <td style="text-align: center; padding: 2px;"><input type="checkbox"/></td> </tr> <tr> <td style="padding: 2px;">Nutzung eines Minitrampolins vor dem Hindernis als Sprunghilfe.</td> <td style="text-align: center; padding: 2px;"><input type="checkbox"/></td> <td style="text-align: center; padding: 2px;"><input checked="" type="checkbox"/></td> </tr> <tr> <td style="padding: 2px;">Springen aus einer Absprungzone nahe am Hindernis.</td> <td style="text-align: center; padding: 2px;"><input type="checkbox"/></td> <td style="text-align: center; padding: 2px;"><input checked="" type="checkbox"/></td> </tr> <tr> <td style="padding: 2px;">Handstütz auf dem Kasten, ein oder mehrere Schritte an der Wand und Überspringen des Kastens ohne Fußkontakt auf dem Kasten.</td> <td style="text-align: center; padding: 2px;"><input checked="" type="checkbox"/></td> <td style="text-align: center; padding: 2px;"><input type="checkbox"/></td> </tr> <tr> <td style="padding: 2px;">Beidbeiniges Abspringen an die Wand.</td> <td style="text-align: center; padding: 2px;"><input type="checkbox"/></td> <td style="text-align: center; padding: 2px;"><input checked="" type="checkbox"/></td> </tr> </table>	Handstütz auf dem Kasten, ein Schritt gegen die Wand und Aufknien auf dem Kasten.	<input checked="" type="checkbox"/>	<input type="checkbox"/>	Nutzung eines Minitrampolins vor dem Hindernis als Sprunghilfe.	<input type="checkbox"/>	<input checked="" type="checkbox"/>	Springen aus einer Absprungzone nahe am Hindernis.	<input type="checkbox"/>	<input checked="" type="checkbox"/>	Handstütz auf dem Kasten, ein oder mehrere Schritte an der Wand und Überspringen des Kastens ohne Fußkontakt auf dem Kasten.	<input checked="" type="checkbox"/>	<input type="checkbox"/>	Beidbeiniges Abspringen an die Wand.	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<p>Beim Inline-Skaten gelingt es einem Schüler Ihrer 7. Klasse nicht, enge Kurven zu fahren. Worauf könnte diese Schwierigkeit zurückgeführt werden?</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th style="width: 80%;"></th> <th style="width: 10%; text-align: center;">Ja</th> <th style="width: 10%; text-align: center;">Nein</th> </tr> </thead> <tbody> <tr> <td style="padding: 2px;">Die Fußstellung ist zu breit.</td> <td style="text-align: center; padding: 2px;"><input checked="" type="checkbox"/></td> <td style="text-align: center; padding: 2px;"><input type="checkbox"/></td> </tr> <tr> <td style="padding: 2px;">Die Fußstellung ist zu eng.</td> <td style="text-align: center; padding: 2px;"><input type="checkbox"/></td> <td style="text-align: center; padding: 2px;"><input checked="" type="checkbox"/></td> </tr> <tr> <td style="padding: 2px;">Es findet kein Belastungswechsel auf den Außenfuß statt.</td> <td style="text-align: center; padding: 2px;"><input checked="" type="checkbox"/></td> <td style="text-align: center; padding: 2px;"><input type="checkbox"/></td> </tr> <tr> <td style="padding: 2px;">Es findet kein Belastungswechsel auf den Innenfuß statt.</td> <td style="text-align: center; padding: 2px;"><input checked="" type="checkbox"/></td> <td style="text-align: center; padding: 2px;"><input type="checkbox"/></td> </tr> <tr> <td style="padding: 2px;">Der Körperschwerpunkt wird ins Kurveninnere verlagert.</td> <td style="text-align: center; padding: 2px;"><input type="checkbox"/></td> <td style="text-align: center; padding: 2px;"><input checked="" type="checkbox"/></td> </tr> </tbody> </table>		Ja	Nein	Die Fußstellung ist zu breit.	<input checked="" type="checkbox"/>	<input type="checkbox"/>	Die Fußstellung ist zu eng.	<input type="checkbox"/>	<input checked="" type="checkbox"/>	Es findet kein Belastungswechsel auf den Außenfuß statt.	<input checked="" type="checkbox"/>	<input type="checkbox"/>	Es findet kein Belastungswechsel auf den Innenfuß statt.	<input checked="" type="checkbox"/>	<input type="checkbox"/>	Der Körperschwerpunkt wird ins Kurveninnere verlagert.	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Handstütz auf dem Kasten, ein Schritt gegen die Wand und Aufknien auf dem Kasten.	<input checked="" type="checkbox"/>	<input type="checkbox"/>																																
Nutzung eines Minitrampolins vor dem Hindernis als Sprunghilfe.	<input type="checkbox"/>	<input checked="" type="checkbox"/>																																
Springen aus einer Absprungzone nahe am Hindernis.	<input type="checkbox"/>	<input checked="" type="checkbox"/>																																
Handstütz auf dem Kasten, ein oder mehrere Schritte an der Wand und Überspringen des Kastens ohne Fußkontakt auf dem Kasten.	<input checked="" type="checkbox"/>	<input type="checkbox"/>																																
Beidbeiniges Abspringen an die Wand.	<input type="checkbox"/>	<input checked="" type="checkbox"/>																																
	Ja	Nein																																
Die Fußstellung ist zu breit.	<input checked="" type="checkbox"/>	<input type="checkbox"/>																																
Die Fußstellung ist zu eng.	<input type="checkbox"/>	<input checked="" type="checkbox"/>																																
Es findet kein Belastungswechsel auf den Außenfuß statt.	<input checked="" type="checkbox"/>	<input type="checkbox"/>																																
Es findet kein Belastungswechsel auf den Innenfuß statt.	<input checked="" type="checkbox"/>	<input type="checkbox"/>																																
Der Körperschwerpunkt wird ins Kurveninnere verlagert.	<input type="checkbox"/>	<input checked="" type="checkbox"/>																																
<p>Ihre 9. Klasse soll das schnelle Überlaufen von Hindernissen erlernen und zentrale Bewegungsmerkmale benennen können. Möglich sind folgende Vorgehensweisen:</p> <p>a) Übt die Bewegung so, wie ich sie euch vormache. Achtet dabei auf die genannten Merkmale.</p> <p>b) Erprobt, wie ihr am schnellsten über die Hindernisse kommt und erklärt, worauf man achten muss.</p> <p>Formulieren Sie eine Vorgehensweise mit einer (mehreren) Aufgabenstellung(en), die den Schülerinnen und Schülern mehr Freiraum als in a) lässt, aber stärker inhaltlich strukturiert als in b). Beschreiben Sie konkret, welche Strategien, Merkmale, Übungen etc. Sie bei ihrer Vorgehensweise berücksichtigen.</p>																																		

Abb. 1: Beispiele für die verschiedenen Itemformen: Lernende unterstützen (oben links, die Bildfolge zeigt die korrekte Bewegungsausführung), Schwierigkeiten erklären (oben rechts), Vorgehensweisen formulieren (unten)

werden, die sich im Grad der Lenkung zwischen den beiden vorgegebenen Vorgehensweisen bewegt (s. Abb. 1, unten). Im Rahmen einer Pilotierung hat sich gezeigt, dass gerade diese mittlere Vorgehensweise dezidierte inhaltlich-methodische Überlegungen abverlangt, während die stark und die wenig lenkenden Vorgehensweisen auch relativ unabhängig vom Inhalt formuliert werden können und somit weniger aussagekräftig für das FDW sind. Die Codierung der Antworten erfolgt auf Basis einer Systematik, die nach einer Analyse von Expert*innenantworten erstellt wurde: Die Expert*innenantworten auf die offenen Items wurden mittels einer strukturierenden qualitativen Inhaltsanalyse verdichtet und zusammengefasst (Kuckartz, 2016). Die Auswertung wurde durch die Software MAXQDA unterstützt. Es wurden folgende, theoretisch begründeten Analysekatgorien gemäß dem fachdidaktischen Anspruch einer fachlich-strukturierten Öffnung bei gleichzeitigem Spielraum für Entscheidungen verwendet: (1) Strukturierungshilfen; (2) Ermöglichen von Entscheidungen. Weiterhin wurden diese ersten beiden Analysekatgorien jeweils ausdifferenziert hinsichtlich (a) Inhalte und (b) Methoden (s. Tab. 1).

Tab. 1: Heuristik zur Codierung strukturierter Lenkung in offenen Antworten der FDW-Subdimension Vorgehensweisen formulieren

	Strukturierungshilfen	Ermöglichen von Entscheidungen
Inhaltlich	1a	2a
Methodisch	1b	2b

Mit *Strukturierungshilfen* sind Vorgaben der Lehrperson gemeint, die sich auf inhaltliche und methodische Aspekte beziehen. Die von den Expert*innen vorgebrachten Antworten lassen sich exemplarisch an der in Abb. 1 dargestellten Aufgabe aufzeigen: Für eine inhaltliche Strukturierung werden von den Expert*innen hier (1a) konkrete Problemstellungen vorgeschlagen („Wie schaffe ich einen gleichbleibenden Rhythmus?“), Bewegungsmerkmale angegeben („flaches Überlaufen, Schwungbeinkick, Abspreizen des Nachziehbeins, Oberkörpervorlage, Gegenarmeinsatz“) und „akustische/optische Hilfen“ empfohlen. Für eine methodische Strukturierung (1b) wird der Einsatz von Arbeitsblättern oder Gruppenarbeit vorgeschlagen oder es werden ganze Übungsreihen zur Orientierung benannt („Rhythmisches Überlaufen von Hindernissen üben, Beidseitigkeit zunächst betonen, um dann festzulegen, mit welchem Bein zuerst überquert wird, optische und akustische Hilfen zur Rhythmusfindung und zum Aufsetzen der Füße geben [...]“).

Mit *Ermöglichen von Entscheidungen* sind Instruktionen gemeint, die Schüler*innen dazu auffordern, inhaltliche (2a) oder methodische (2b) Entscheidungen zu treffen. Ebenfalls für den oben genannten Kontext der Aufgabenstellung in Abb. 1, wurden von den Expert*innen hinsichtlich inhaltlicher Entscheidungen (2a) bspw. Erprobungsaufträge („Probiert verschiedene Lösungen aus und benennt wichtige Merkmale, die für ein schnelles Überqueren sinnvoll sind“) oder das Entwickeln individueller Lösungen („Welcher Abstand ist für euch sinnvoll?“) benannt. Methodische Entscheidungen (2b) liegen im Sinne der Expert*innen etwa dann vor, wenn Ziele, Wege, Übungen, Trainingsschwerpunkte, Zeiten oder Sozialformen zu einer Aufgabenstellung durch die Schüler*innen selbst bestimmt werden können.

Da aus der fachdidaktischen Perspektive eines erziehenden Sportunterrichts sowohl solche Elemente als sinnvoll erachtet werden, die den Unterricht strukturieren, als auch solche, die einen Spielraum für Entscheidungen ermöglichen, werden solche Antworten als am treffendsten eingestuft, die alle vier in Tab. 1 genannten Elemente beinhalten.

Für das Kodiermanual wurden die Kategorien einerseits itemunspezifisch beschrieben und andererseits durch fachinhaltliche Aspekte jeden Items ergänzt. Tab. 2 zeigt das Kodiermanual exemplarisch für das Item *Hindernisse*. Vor Anwendung des Kodiermanuals werden die Rater*innen in mehreren Sitzungen durch die Projektleitung geschult. Erste Befunde hinsichtlich der Interraterreliabilität fallen zufriedenstellend aus: Insgesamt haben zwei studentische Rater*innen unabhängig voneinander alle acht offenen Antworten von insgesamt 65 Studierenden bewertet; das durchschnittliche Cohens Kappa liegt bei $M = .80$ ($SD = .04$, $range = .74$ bis $.85$).

Tab. 2: Kodiermanual für die Kodierung der offenen Antworten in der Wissensfacette Vorgehensweisen formulieren für das Bewegungsfeld Laufen, Springen, Werfen mit dem Thema Hindernisse überqueren

Kriterium	Allgemeine Beschreibung	Itemspezifische Indikatoren
Inhaltliche Strukturierung	Die Lehrer*in strukturiert den Unterricht über inhaltliche Aspekte, die sie vorgibt.	Die Lehrer*in strukturiert den Unterricht durch themenspezifische Merkmale, z. B. Gegenarmbewegung, Körperschwerpunkt, Abstände der Hindernisse, Sprunghöhe, Rhythmus, Schrittlänge, Oberkörperneigung, Schwungbeineinsatz, Nachzieheineinsatz.
Methodische Strukturierung	Die Lehrer*in strukturiert den Unterricht über Vorgaben betreffend der Ziele, der Wege oder der Zeit.	Die Lehrer*in strukturiert den Unterricht z. B. mit Bezug auf folgende Ansätze: Erfahrungsorientierung, Kontrastübungen, Video-Feedback, Beobachtungsbögen.
Inhaltliche Entscheidungen	Die Schüler*innen müssen durch das Vorgehen Entscheidungen im Hinblick auf Inhalte treffen.	Die Schüler*innen müssen durch das Vorgehen Probleme beim schnellen Überlaufen erarbeiten oder Lösungen für das Problem erarbeiten.
Methodische Entscheidungen	Die Schüler*innen müssen durch das Vorgehen Entscheidungen über Ziele oder Wege oder Zeit oder Personenkonstellationen treffen.	Die Schüler*innen müssen durch das Vorgehen aus unterschiedlichen Hindernissen passende wählen oder den Zeitpunkt des nächsten Schwierigkeitsniveaus wählen.

6 | EINSCHÄTZUNG DER RELEVANZ DER TESTITEMS, DER PLAUSIBILITÄT DER ANTWORTEN UND DER ZUORDNUNG ZU DEN WISSENSFACETTEN DURCH EXPERT*INNEN

Für die vorliegende Untersuchung wurde auf Basis einer Expert*innenbefragung überprüft, ob erstens die Items relevante Situationen des Sportunterrichts umfassen und damit als handlungsnah eingestuft werden und zweitens die Antwortmöglichkeiten plausibel sind und damit als fachlich angemessen eingestuft werden können.

6.1 | METHODISCHES VORGEHEN DER EXPERT*INNENBEFRAGUNG

Stichprobe. Für die Studie konnten 28 Expert*innen gewonnen werden. Es handelt sich um 13 Fachseminarleiter*innen der zweiten Ausbildungsphase der Lehrer*innenbildung (4 weiblich) und 15 an einer Universität arbeitende Fachdidaktiker*innen (5 weiblich). Sie waren im Durchschnitt $M = 48,60$ Jahre alt ($SD = 9,32$) und hatten eine durchschnittliche Berufserfahrung von $M = 13,41$ Jahren ($SD = 8,13$). Es wurden folgende Auswahlkriterien für die Expert*innen herangezogen: (i) ein Minimum von fünf Jahren Berufserfahrung – nach einem solchen Zeitraum ist die erste Phase der Rollenfindung überwunden und es werden allmählich Routinen etabliert (Miethling, 2013); (ii) ein Lehramtsstudium im Fach Sport – wir nehmen an, dass damit ein Grundkanon an theoretischem Wissen vorhanden ist; (iii) ein Sek-I-Bezug in der Ausbildungstätigkeit – wir gehen davon aus, dass hierdurch vielfältige unterrichtliche Situationen im Sinne der reflektierten Praxis (mit Referendar*innen) durchdacht wurden.

Insgesamt sollte so gewährleistet werden, dass die Expert*innen über ausreichend Kenntnisse und Fähigkeiten in Bezug auf unterrichtliche Praxis verfügen und damit die Qualität der hypothetischen Unterrichtsszenarios einschätzen können. Den Expert*innen wurde für die Teilnahme eine Aufwandsentschädigung ausgezahlt.

Datenerhebung und -auswertung. Alle Daten wurden mittels einer Online-Befragung über die Software *Unipark* (Questback GmbH, 2018) erhoben. Die Expert*innen beantworteten die Items zunächst selbst und wurden stets im Anschluss um die Bewertung der Items gebeten. Hierzu wurde erstens die Relevanz der Items auf einer fünfstufigen Likert-Skala (0 = *trifft nicht zu* bis 4 = *trifft zu*) eingeschätzt („Der hier beschriebene Kontext ist für den Sportunterricht relevant“). Zweitens wurde auf Basis der gleichen Skala die Plausibilität der Antworten eingeschätzt (geschlossene Items der Facetten *Lernende unterstützen* und *Schwierigkeiten erklären*: „Die Zuordnung richtiger und falscher Antworten scheint mir plausibel“; offene Items der Facette *Vorgehensweisen formulieren*; „Es gelingt mir, eine geeignete Antwort zu formulieren“). Drittens sollte im Rahmen der Expert*innenbefragung die theoretisch angenommene Zuordnung der Items zu den Wissensfacetten eingeschätzt werden. Da die Wissensfacetten als Antwortalternativen jedoch umfänglicherer Erklärungen bedürft hätten, wurden die Expert*innen aufgefordert, ein jedes Item einer spezifischen handlungsnahen Anforderung zuzuordnen, bezüglich derer eine große Nähe zur entsprechenden Wissensfacette angenommen wurde („Welche Fähigkeit einer Lehrkraft wird bei dieser Aufgabe vor allem gefordert?“). Als Antwortalternativen standen *Diagnostizieren* (Wissensfacette: Schwierigkeiten erklären), *Unterstützen* (Wissensfacette: Lernende unterstützen) und *Planen* (Wissensfacette: Vorgehensweisen formulieren) zur Auswahl.

6.2 | ERGEBNISDARSTELLUNG UND INTERPRETATION

Insgesamt kann die Einschätzung der Relevanz der einzelnen Items als gut bis sehr gut eingestuft werden. Über alle Items zeigt sich ein Mittelwert von $M = 3,35$ ($SD = 0,93$; range: 1,96-3,92). Die Werte liegen damit in einem ähnlichen Bereich wie die Expertenbefragungen im Rahmen der *FALKO-Studie* (z. B. im Deutsch-FDW Test $M = 3,10$; $SD = 0,40$; range 2,54 - 3,48 (Pissarek & Schilcher, 2017, S. 91)).

Hinsichtlich der Plausibilität der Antworten zeigt sich über alle Items ein Mittelwert von $M = 3,01$ ($SD = 1,06$; range: 2,27-3,85). Vergleicht man die Items nach Bewegungsfeldern, so zeigt sich für die Bewegungsfelder *Fitness* und *Rollen, Gleiten, Fahren* eine insgesamt etwas niedrigere Relevanz als in den anderen Bewegungsfeldern. Die erhöhte Standardabweichung in den betreffenden Items kann als Uneinigkeit der Expert*innen für die entsprechenden Items aufgefasst werden. Trotz kleinerer Unterschiede kann jedoch festgestellt werden, dass die aus der Fachliteratur abgeleiteten Antworten und die entwickelten Distraktoren bezüglich der Handlungsnähe zum Unterrichtsgeschehen als angemessen wahrgenommen werden.

Tab. 3: Expert*inneneinschätzung zur Relevanz und Plausibilität der einzelnen Items nach Bewegungsfeldern

Wissensfacette	Inhalt	Relevanz		Plausibilität	
		M	SD	M	SD
Laufen, Springen, Werfen		3,55	0,69	3,11	0,96
Lernende unterstützen	Schlagballwurf	3,82	0,48	3,43	0,79
	Weitsprung	3,69	0,55	2,73	1,15
Schwierigkeiten erklären	Kugelstoßen	3,15	1,12	3,12	0,78
	Hochsprung	3,41	0,75	2,88	1,14
Vorgehensweisen formulieren	Hindernisse	3,69	0,55	3,38	0,94
Bewegen an und mit Geräten		3,59	0,83	3,30	0,93
Lernende unterstützen	Parkour	3,50	0,71	3,85	0,37
	Felgaufschwung	3,52	0,96	3,32	0,90
Schwierigkeiten erklären	Radwende	3,56	0,80	3,19	0,92
	Akrobatik	3,73	0,60	3,08	1,16
Vorgehensweisen formulieren	Handstand	3,64	1,08	3,08	1,29
Spielen in und mit Regelstrukturen		3,69	0,79	2,93	1,27
Lernende unterstützen	Zieltreffen	3,46	1,10	2,27	1,19
	Volleyball	3,85	0,61	2,85	1,32
Schwierigkeiten erklären	Dribbling	3,80	0,65	3,04	1,17
	Basketball	3,80	0,41	3,48	1,05
Vorgehensweisen formulieren	Badminton	3,54	1,17	3,00	1,63
Den Körper trainieren, die Fitness verbessern		2,97	1,10	2,80	1,03
Lernende unterstützen	Sling-Trainer	1,96	1,46	2,54	1,03
	Ausdauer	3,60	0,76	2,52	1,05
Schwierigkeiten erklären	Grätschsitz	2,59	1,42	2,67	1,18
	Stabilisation	3,40	0,76	3,04	0,79
Vorgehensweisen formulieren	Bauchmuskulatur	3,31	1,09	3,23	1,11

Wissensfacette	Inhalt	Relevanz		Plausibilität	
Bewegung gestalten		3,23	1,01	2,78	1,21
Lernende unterstützen	Jump-Style	3,31	0,93	2,96	1,04
	Gummitwist	2,77	1,21	2,50	1,07
Schwierigkeiten erklären	Volkstanz	3,19	1,10	3,31	0,97
	Präsentationen	3,65	0,63	2,65	1,36
Vorgehensweisen formulieren	Aerobic	3,24	1,17	2,50	1,64
Mit/gegen Partner kämpfen		3,39	0,79	3,14	0,94
Lernende unterstützen	Fixieren	2,96	1,22	2,85	1,01
	Regeln	3,68	0,56	3,36	1,00
Schwierigkeiten erklären	Judo	3,12	1,14	2,96	1,00
	Linie	3,28	0,74	3,16	0,75
Vorgehensweisen formulieren	Raufregeln	3,92	0,28	3,36	0,95
Bewegen im Wasser		3,63	0,90	3,13	1,02
Lernende unterstützen	Brustschwimmen	3,76	0,83	3,32	0,99
	Startsprung	3,54	0,95	2,73	1,19
Schwierigkeiten erklären	Kraulschwimmen	3,50	1,03	3,00	0,85
	Streckentauchen	3,72	0,84	3,72	0,79
Vorgehensweisen formulieren	Auftrieb	3,64	0,86	2,88	1,30
Rollen, Gleiten, Fahren		2,76	1,35	2,86	1,12
Lernende unterstützen	Waveboard	2,78	1,19	2,44	0,96
	Skifahren	2,85	1,35	3,00	0,96
Schwierigkeiten erklären	Schlittschuhfahren	2,73	1,46	3,42	1,10
	Inlineskate	3,00	1,27	3,12	1,05
Vorgehensweisen formulieren	Skateboard	2,46	1,48	2,30	1,51

Hinsichtlich der Zuordnung der Items zu den Fähigkeiten zeigt sich, dass insbesondere der Fähigkeit *Planen* mehr Items zugeordnet wurden, als es in der theoretischen Konzeption angedacht war. So werden neben den Items der Wissensfacette *Vorgehensweisen formulieren* (100 %, hier geht es um das Planen) auch Items der Wissensfacetten *Lernende unterstützen* (19 %) und Items der Wissensfacette *Schwierigkeiten erklären* (31 %) der Fähigkeit zugeordnet. Dies deutet darauf hin, dass die so zugeordneten Items nicht eindeutig genug der angedachten Fähigkeit zugeordnet werden können, bzw. es könnten auch Unstimmigkeiten bei der Interpretation der Begriffe *Diagnostizieren*, *Unterstützen* und *Planen* vorliegen. Um potenzielle Probleme der Itemzuordnungen und damit der Konstruktvalidität im Rahmen späterer Studierendenbefragungen vorzubeugen, wurden die entsprechenden Items nochmals dezidiert überarbeitet.

Über diese Einschätzungen hinaus waren auch die freien Anmerkungen der Expert*innen hilfreich. Sie haben dazu beigetragen, die hypothetischen Unterrichtsszenarios und Antwortmöglich-

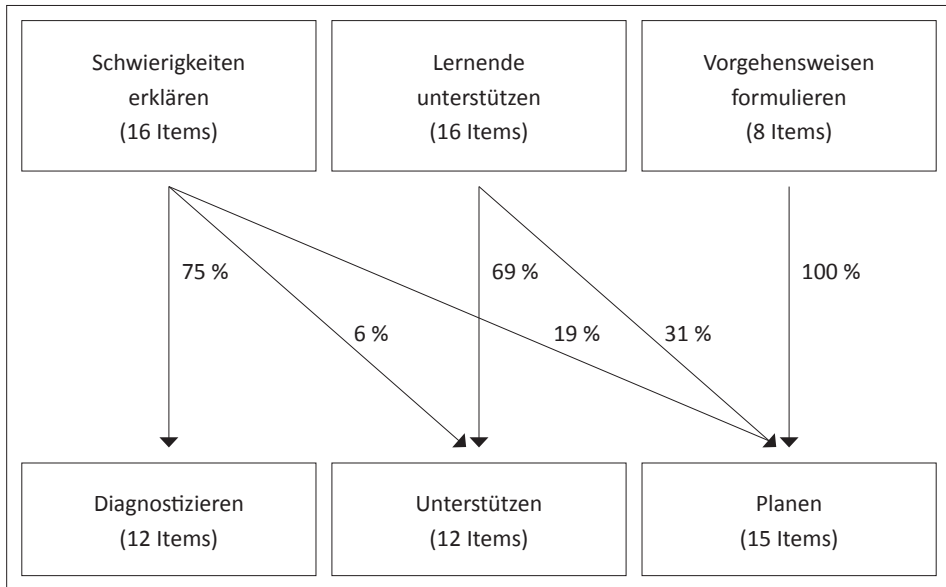


Abb. 2: Zuordnung von Items zu Fähigkeiten durch die befragten Expert*innen

keiten weiterzuentwickeln. Hinsichtlich der hypothetischen Unterrichtsszenarios wurden insbesondere Lernvoraussetzungen konkretisiert, Klassenstufen verändert und Inhaltskontexte präzisiert, hinsichtlich der Antworten wurden sachliche Richtigstellungen sowie Begriffsschärfungen vorgenommen.

Zusammengenommen sprechen die Werte bezüglich der Relevanz des Itemkontextes und der Plausibilität der Antworten dafür, die Items beizubehalten, wenngleich bestimmte Aspekte mit Unterstützung durch bewegungsfeldspezifische Expert*innen überarbeitet wurden. Da die Rückmeldungen zur Weiterentwicklung nicht die Zusammensetzung des Itempools grundlegend verändern, wird die aufgestellte Repräsentation der Itemmenge beibehalten.

7 | ZUSAMMENFASSUNG UND FAZIT

Die systematische Herleitung und Prüfung der Inhaltsvalidität wird zwar immer wieder eingefordert, jedoch zeigt die dargelegte Analyse, dass sie in bisherigen Veröffentlichungen nur eingeschränkt nachvollziehbar ist. Zentrale Merkmale der Inhaltsvalidität sind die inhaltliche Relevanz und die Beachtung der Repräsentation der theoretisch hergeleiteten Itemmenge und des anschließenden Itempools der entwickelten Testaufgaben. Abweichungen sollten auf der Ebene der Herleitung der Itemmenge mitbegründet werden (s. Kap. 2).

Für die Konstruktion von Tests zum FDW wird nach Analyse exemplarischer Studien folgendes Vorgehen vorgeschlagen:

1. Bei der Konzeptspezifikation sollte die Itemmenge systematisch durch die Spezifizierung der Dimensionen *Wissensfacetten*, *kognitive Prozesse* und *fachliche Inhalte* hergeleitet werden. Die Form der verwendeten Items – Aufgabenstellungen und Antwortformate – sollte be-

gründet werden. Hypothetische Unterrichtsszenarios scheinen zur Kontextualisierung geeignet, um kognitive Prozesse auf ein handlungsnahes kognitives Anforderungsprofil einzuengen. Jedoch sollten solche Reduktionen auf eine einzelne Ausprägung explizit reflektiert werden. Das (mathematische) Produkt der drei Dimensionen gibt schließlich an, wie viele Komponenten eine Itemmenge hat.

2. Beim Schritt der Operationalisierung sollte dargestellt werden, inwiefern der entwickelte Itempool die zuvor aufgestellte Itemmenge repräsentiert. Einschränkungen (z. B. unterschiedliche Gewichtungen einzelner Komponenten oder Dimensionen) sollten diskutiert und begründet werden.
3. Nach weiteren Überarbeitungsschritten sollte reflektiert werden, inwiefern der Itempool die Itemmenge nach wie vor repräsentiert, ggf. sollten Items überarbeitet und/oder nachentwickelt werden.

In dem vorliegenden eigenen Ansatz zur Messung von FDW wurde die Itemmenge durch acht Bewegungsfelder, drei Wissensfacetten und ein handlungsnahes kognitives Anforderungsprofil bestimmt. Letzteres erfolgte über den Einsatz hypothetischer Unterrichtsszenarios. Damit wurde eine Itemmenge mit 24 Komponenten hergeleitet, die durch einen Itempool mit 40 Items repräsentiert wird. Die Expertenbefragung weist insgesamt darauf hin, dass relevante und plausible Items entwickelt wurden. Während die theoretisch angedachte Zuordnung der Items zu den drei Wissensfacetten im Großen und Ganzen gelingt (s. Abb. 2), bedurfte es bei einzelnen Items einer leichten Überarbeitung. Die Konstrukt- und Kriteriumsvalidität des Tests gilt es zukünftig z. B. anhand von Studierendenstichproben zu überprüfen.

Es erscheint zudem für zukünftige Forschung lohnenswert, zu klären, in welchem Verhältnis Itemmenge und Itempool stehen. Um einem für die Erhebungen zu großen Itempool vorzubeugen, könnte eine Lösung eine Fokussierung auf einzelne Inhaltsbereiche sein. Auch eine Orientierung an spezifischen Altersstufen könnte einen Ansatzpunkt darstellen – so konzentriert sich die *TEDS-M-Studie* etwa auf einen Primarstufen- und einen Sek-I-Test (Blömeke, Kaiser & Lehmann, 2010a; 2010b). Eine solche Fokussierung erfordert jedoch auch, zu klären, inwieweit dann generisches bzw. auf andere Inhaltsbereiche transferierbares oder aber spezifisches FDW erhoben wird. So könnte etwa der Befund, dass die sportmotorische Expertise die professionelle Unterrichtswahrnehmung stark beeinflusst (Reuter, 2012, S. 244), auch ein Hinweis darauf sein, dass Sportlehrkräften ein wissensbezogener Transfer von einem Inhalt zum anderen eher schwerfällt. Diese Annahme gilt es, in zukünftigen Studien zu untersuchen.

Abschließend sei darauf verwiesen, dass das FDW von Sportlehrkräften zwar als notwendige, nicht aber hinreichende Voraussetzung für erfolgreichen Unterricht angenommen werden kann. So gilt es, zukünftig weitere Kompetenzbereiche (z. B. Überzeugungen, s. Kap. 1) unter fachdidaktischer Perspektive zu fokussieren und auszudifferenzieren und auch in einen Zusammenhang zu unterrichtlichen Prozessen zu bringen, wie es in anderen Fächern bereits erfolgt ist (Kunter et al., 2011).

LITERATUR

- Anderson, L. W. & Krathwohl, D. R. (2001). *A taxonomy for learning, teaching, and assessing. A revision of Bloom's taxonomy of educational objectives*. New York: Longman.
- Aufschnaiter, C. v. & Rogge, C. (2010). Misconceptions or Missing Conceptions? *Eurasia Journal of Mathematics, Science and Technology Education*, 6, 3-18.
- Bähr, I. & Wibowo, J. (2018). Theoretische Grundlagen - zum Verständnis Kooperativen Lernens. In J. Wibowo & I. Bähr (Hrsg.), *Kooperatives Lernen im Sportunterricht*. (S. 15-39). Baltmannsweiler: Schneider.
- Baumgartner, M. (2018). „... Kompetenz ohne Performanz ist leer! Performanz ohne Kompetenz blind ...!“ Zu einem integrativen Kompetenzstrukturmodell von Sportlehrkräften. *Zeitschrift für sportpädagogische Forschung*, 6 (1), 49-68.
- Beckers, E. (2013). Prinzipien eines erziehenden Sportunterrichts. In H. Aschebrock & G. Stibbe (Hrsg.), *Didaktische Konzepte für den Schulsport* (S. 178-196). Aachen: Meyer & Meyer.
- Blömeke, S., Gustafsson, J.-E. & Shavelson, R. J. (2015). Beyond dichotomies. Competence viewed as a continuum. *Zeitschrift für Psychologie*, 223 (1), 3-13.
- Blömeke, S. & Kaiser, G. (2017). Understanding the development of teachers' professional competencies as personally, situationally and socially determined. In D. J. Clandinin & J. Husu (Eds.), *International handbook of research on teacher education* (pp. 783-802). Thousand Oaks: SAGE Publications.
- Blömeke, S., Kaiser, G. & Lehmann, R. (Hrsg.). (2010a). *TEDS-M 2008. Professionelle Kompetenz und Lerngelegenheiten angehender Mathematiklehrkräfte für die Sekundarstufe I im internationalen Vergleich*. Münster: Waxmann.
- Blömeke, S., Kaiser, G. & Lehmann, R. (Hrsg.). (2010b). *TEDS-M 2008. Professionelle Kompetenz und Lerngelegenheiten angehender Primarstufenlehrkräfte im internationalen Vergleich*. Münster: Waxmann.
- Bloom, B. (1976). *Taxonomie von Lernzielen im kognitiven Bereich*. Weinheim: Beltz.
- Borko, H., Roberts, S. A. & Shavelson, R. J. (2008). Teachers' decision making: from Alan J. Bishop to today. In P. Clarkson (Ed.), *Critical issues in mathematics education* (pp. 37-67). New York: Springer Science.
- Bremerich-Vos, A., Buchholtz, C. & König, J. (2016). PlanvoLL-D – The role of professional knowledge of pre-service german teachers in their lesson planning: Validation and methodological innovation. In H. A. Pant, O. Zlatkin-Troitschanskaia, C. Lautenbach, M. Toepper & D. Molerov (Eds.), *Modeling and measuring competencies in higher education* (pp. 31-34). Berlin: Humboldt University; Johannes Gutenberg University.
- Buchholtz, N. & Kaiser, G. (2013). Professionelles Wissen im Studienverlauf: Lehramt Mathematik. In S. Blömeke, A. Bremerich-Vos, G. Kaiser, G. Nold, H. Haudeck, J.-U. Keßler et al. (Hrsg.), *Professionelle Kompetenzen im Studienverlauf. Weitere Ergebnisse zur Deutsch-, Englisch- und Mathematiklehrausbildung aus TEDS-LT* (S. 107-143). Münster: Waxmann.
- Buchholtz, N., Kaiser, G. & Blömeke, S. (2014). Die Erhebung mathematikdidaktischen Wissens – Konzeptualisierung einer komplexen Domäne. *Journal für Mathematik-Didaktik*, 35 (1), 101-128.
- Depaepe, F., Verschaffel, L. & Kelchtermans, G. (2013). Pedagogical content knowledge. A systematic review of the way in which the concept has pervaded mathematics educational research. *Teaching and Teacher Education*, 34, 12-25.
- Döring, N. & Bortz, J. (2016). *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften*. Berlin, Heidelberg: Springer.
- Gerlach, E., Kussin, U., Brandl-Bredenbeck, H. P. & Brettschneider, W.-D. (2006). Der Sportunterricht aus Schülerperspektive. In Deutscher Sportbund (Hrsg.), *DSB-Sprint-Studie. Sportunterricht in Deutschland. Eine Untersuchung zur Situation des Schulsports in Deutschland* (S. 107-144). Aachen: Meyer & Meyer.
- Hartig, J., Frey, A. & Jude, N. (2012). Validität. In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (S. 143-171). Berlin: Springer.
- 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*, 3 (2), 41-60.
- Heim, R. & Sohnmeyer, J. (2015). Schulsport. In W. Schmidt, N. Neuber, T. Rauschenbach, H. P. Brandl-Bredenbeck, J. Süßenbach & C. Breuer (Hrsg.), *Dritter*

- Deutscher Kinder- und Jugendsportbericht. Kinder- und Jugendsport im Umbruch* (S. 118-139). Schorndorf: Hofmann.
- Kirschner, P. A., Sweller, J. & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist*, 41 (2), 75-86.
- Klauer, K. J. (1984). Kontentvalidität. *Diagnostica*, 30, 1-23.
- Krauss, S. (2011). Das Experten-Paradigma in der Forschung zum Lehrerberuf. In E. Terhart, H. Bennewitz & M. Rothland (Hrsg.), *Handbuch der Forschung zum Lehrerberuf*. (S. 171-191). Münster: Waxmann.
- Krauss, S., Blum, W., Brunner, M., Neubrand, M., Baumert, J., Kunter, M. et al. (2011). Konzeptualisierung und Testkonstruktion zum fachbezogenen Professionswissen 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. 135-160). Münster: Waxmann.
- Krauss, S., Lindl, A., Schilcher, A., Fricke, M., Göhring, A., Hofmann, B. et al. (Hrsg.). (2017). *FALKO: Fachspezifische Lehrerkompetenzen. Konzeption von Professionswissenstests in den Fächern Deutsch, Englisch, Latein, Physik, Musik, Evangelische Religion und Pädagogik: mit neuen Daten aus der COACTIV-Studie*. Münster: Waxmann.
- Krauss, S., Lindl, A., Schilcher, A. & Tepner, O. (2017). Das Forschungsprojekt FALKO – ein einleitender Überblick. In S. Krauss, A. Lindl, A. Schilcher, M. Fricke, A. Göhring, B. Hofmann et al. (Hrsg.), *FALKO: Fachspezifische Lehrerkompetenzen. Konzeption von Professionswissenstests in den Fächern Deutsch, Englisch, Latein, Physik, Musik, Evangelische Religion und Pädagogik: mit neuen Daten aus der COACTIV-Studie* (S. 9-65). Münster: Waxmann.
- Kuckartz, U. (2016). *Qualitative Inhaltsanalyse. Methoden, Praxis, Computerunterstützung*. Wertheim: Beltz Juventa.
- Kunter, M., Kleickmann, T., Klusmann, U. & Richter, D. (2013). The development of teachers' professional competence. In M. Kunter, J. Baumert & W. Blum (Eds.), *Cognitive activation in the mathematics classroom and professional competence of teachers* (pp. 63-78). Dordrecht: Springer.
- Kunter, M., Baumert, J., Blum, W., Klusmann, U., Krauss, S. & Neubrand, M. (Hrsg.). (2011). *Professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV*. Münster: Waxmann.
- 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 (3), 805-820.
- Laging, R. (Hrsg.). (2008). *Inhalte und Themen des Bewegungs- und Sportunterrichts*. Baltmannsweiler: Schneider.
- Lüders, M. (2012). „Pädagogisches Unterrichtswissen“ – eine Testkritik. *Zeitschrift für Erziehungswissenschaft*, 15 (4), 775-791.
- Meier, S. (2018). Fachdidaktisches Wissen angehender Sportlehrkräfte. Ein Konzeptualisierungsvorschlag. *Zeitschrift für sportpädagogische Forschung*, 6 (1), 69-84.
- Miethling, W.-D. (2013). Zur Entwicklung von Sportlehrer/innen. *Sportwissenschaft*, 43 (3), 197-205.
- Messick, S. (1995). Validity of Psychological Assessment. *American Psychologist*, 50 (9), 741-749.
- Neuweg, G. H. (2011). Das Wissen der Wissensvermittler. In E. Terhart, H. Bennewitz & M. Rothland (Hrsg.), *Handbuch der Forschung zum Lehrerberuf* (S. 451-477). Münster: Waxmann.
- Pissarek, M. & Schilcher, A. (2017). FALKO-D: Die Untersuchung des Professionswissens von Deutschlehrenden. Entwicklung eines Messinstruments zur fachspezifischen Lehrerkompetenz und Ergebnisse zu dessen Validierung. In S. Krauss, A. Lindl, A. Schilcher, M. Fricke, A. Göhring, B. Hofmann et al. (Hrsg.), *FALKO: Fachspezifische Lehrerkompetenzen. Konzeption von Professionswissenstests in den Fächern Deutsch, Englisch, Latein, Physik, Musik, Evangelische Religion und Pädagogik: mit neuen Daten aus der COACTIV-Studie* (S. 67-111). Münster: Waxmann.
- Prohl, R. (2012a). Der Doppelauftrag des Erziehenden Sportunterrichts. In V. Scheid & R. Prohl (Hrsg.), *Sportdidaktik. Grundlagen, Lehrplan, Bewegungsfelder* (S. 70-91). Wiebelsheim: Limpert.
- Prohl, R. (2012b). Vermittlungsformen im Erziehenden Sportunterricht. In V. Scheid & R. Prohl (Hrsg.), *Sportdidaktik. Grundlagen, Lehrplan, Bewegungsfelder* (S. 92-112). Wiebelsheim: Limpert.

- Prohl, R. & Krick, F. (2006). Lehrplan und Lehrplanelentwicklung – programmatische Grundlagen des Schulsports. In Deutscher Sportbund (Hrsg.), *DSB-Sprint-Studie. Sportunterricht in Deutschland. Eine Untersuchung zur Situation des Schulsports in Deutschland* (S. 19-52). Aachen: Meyer & Meyer.
- Questback GmbH. (2018). *EFS Survey, Version Spring 2018*. Köln: Questback GmbH.
- Renkl, A. (2011). Aktives Lernen: Von sinnvollen und weniger sinnvollen theoretischen Perspektiven zu einem schillernden Konstrukt. *Unterrichtswissenschaft*, 39 (3), 197-212.
- Reuker, S. (2012). Der professionelle Blick von Sportlehrkräften. *Sportwissenschaft*, 42 (4), 240-246.
- Reuker, S. (2018). „Ich unterrichte so, wie es die Ereignisse erfordern“ – Der professionelle Blick von Sportlehrkräften und seine Bedeutung für adaptiven Unterricht. *Zeitschrift für sportpädagogische Forschung*, 6 (2), 31-52.
- Santagata, R. & Yeh, C. (2016). The role of perception, interpretation, and decision making in the development of beginning teachers' competence. *ZDM Mathematics Education*, 48 (1), 153-165.
- Scheid, V. & Prohl, R. (Hrsg.). (2012). *Sportdidaktik. Grundlagen, Lehrplan, Bewegungsfelder*. Wiebelsheim: Limpert.
- Scherler, K. (2004). *Sportunterricht auswerten. Eine Unterrichtslehre*. Hamburg: Czwalina Verlag.
- Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, 15 (2), 4-14.
- Shulman, L. S. (1987). Knowledge and teaching. *Harvard Educational Review*, 57 (1), 1-21.
- Vogler, J., Messmer, R. & Allemann, D. (2017). Das fachdidaktische Wissen und Können von Sportlehrpersonen (PCK-Sport). *German Journal of Exercise and Sport Research*, 47 (4), 335-347.
- Vogler, J., Messmer, R., Wibowo, J., Heemsoth, T. & Meier, S. (2018). Drei Zugänge zur Modellierung fachdidaktischen Wissens von Sportlehrpersonen. In E. Balz & D. Kuhlmann (Hrsg.), *Sportwissenschaft in pädagogischem Interesse* (S. 47-55). Hamburg: Feldhaus.
- Ward, P. & Ayvazo, S. (2016). Pedagogical content knowledge: Conceptions and findings in physical education. *Journal of Teaching in Physical Education*, 35 (3), 294-207.
- Zlatkin-Troitschanskaia, O. & Seidel, J. (2011). Kompetenz und ihre Erfassung – das neue „Theorie-Empirie-Problem“ der empirischen Bildungsforschung? In O. Zlatkin-Troitschanskaia (Hrsg.), *Stationen empirischer Bildungsforschung. Traditionslinien und Perspektiven* (S. 219-233). Wiesbaden: VS Verlag.