

DAVID WOITKOWSKI, JOSEF RIESE UND PETER REINHOLD

Modellierung fachwissenschaftlicher Kompetenz angehender Physiklehrkräfte Modelling Content Knowledge of Prospective Physics Teachers

ZUSAMMENFASSUNG

Momentan gelingt es der fachwissenschaftlichen Ausbildung von Physiklehrkräften an deutschen Hochschulen nicht, Leistungsunterschiede und Differenzen im Vorwissen der Studierenden auszugleichen, da bisher kaum systematische Konzepte zur gezielten und adaptiven Förderung fachlicher Kompetenz existieren. Dies ist u. a. auf einen Mangel an differenzierten Instrumenten zur validen und präzisen Diagnose fachlicher Kompetenz zurückzuführen. So setzt ein Großteil der Forschung zur Modellierung und Messung fachlicher Kompetenz bislang auf Schülerebene an, während die wenigen aktuellen Arbeiten auf Lehrer- bzw. Hochschulebene primär der Grundlagenforschung zuzurechnen sind und nicht den Anforderungen einer präzisen Diagnostik genügen, wie es für adaptive Förderkonzepte erforderlich wäre. Vor diesem Hintergrund wird im Beitrag ein Modell fachlicher Kompetenz angehender Physiklehrkräfte im Inhaltsbereich Mechanik erarbeitet und mit Hilfe einer Reanalyse vorhandener Items einer ersten Plausibilitätsprüfung unterzogen, welches die Entwicklung eines umfassenden Fachwissenstests vorbereitet und die Beschreibung von Testitems anhand schwierigkeiterzeugender Aufgabenmerkmale ermöglicht. Perspektivisch soll ausgehend vom vorgestellten Modell die Erstellung und empirische Überprüfung eines Kompetenzniveauomodells erfolgen, um zukünftig eine angemessene und zielgerichtete Förderung Studierender auf einzelnen Niveaustufen möglich zu machen.

Schlüsselwörter: Kompetenzmessung, physikalisches Fachwissen, Modellentwicklung, schwierigkeiterzeugende Aufgabenmerkmale

ABSTRACT

At present, education of prospective physics teachers at German universities is not able to compensate differences in competencies and prior knowledge because there are almost no systematic concepts for a specific and adaptive support available. This is, amongst others, due to a lack of precise, sophisticated and valid instruments for measuring content knowledge. In this context, most of the available approaches are limited to physics education at school level. The few existing approaches focusing on teachers can be characterized as fundamental research and offer no adequate links for support concepts. Against this background, a model of prospective physics teachers' content knowledge in the domain of mechanics is developed in this paper. This model will be used to develop a corresponding comprehensive test and to describe the test items by using difficulty generating item characteristics. Based on the introduced model, the results of a

planned inquiry will be used for empirical verification of a model of competence levels which in the future can be the basis of adequate support of students on specific levels.

Keywords: measurement of competence, physical content knowledge, model development, difficulty generating item characteristics

1 Einleitung

Fachwissen spielt als eine von drei Wissensfacetten neben pädagogischem und fachdidaktischem Wissen eine wesentliche Rolle in der (universitären) Lehrerbildung (vgl. Shulman, 1986) und wird als zentrale Komponente der professionellen Kompetenz von Fachlehrkräften gesehen (vgl. Baumert & Kunter, 2006; Bromme, 1997). Kognitiv anregender Unterricht kann nur mit entsprechend fundiertem Verständnis fachlicher Konzepte auf Seiten der Lehrkräfte gestaltet werden. Dabei sollten Lehrpersonen ihren Schülern nicht nur stofflich „voraus“ sein, sondern ein tieferes Verständnis fachlicher Inhalte des Schulcurriculums besitzen (Krauss et al. 2008), um beispielsweise sicher auf Schülerfragen reagieren zu können oder das fachliche Schüler(vor)verständnis sicher diagnostizieren zu können. Fachwissen bildet damit zusammen mit fachdidaktischem Wissen die Grundlage, auf der Anforderungen im Zusammenhang mit der Planung und Durchführung von Unterricht bewältigt werden können. Das Repräsentations- und Erklärungsrepertoire hängt wesentlich von der Breite und Tiefe des

Fachwissens ab (Baumert & Kunter, 2006), darüber hinaus hat sich Fachwissen – vermittelt über fachdidaktisches Wissen – als bedeutsamer Prädiktor eine kognitiv herausfordernde und konstruktive Unterstützung gewährende Unterrichtsführung erwiesen (Baumert et al., 2010). Fehlt angemessenes Fachwissen, schlägt sich das direkt negativ auf die Schülerleistungen nieder (vgl. z. B. Anderson, 1979). Weiterhin wird es als wesentliche Voraussetzungen für den Erwerb fachdidaktischen Wissens gesehen (vgl. z. B. Krauss et al., 2008; Terhart, 2002).

Trotzdem weist die Lehrerbildung an Hochschulen (in Deutschland, aber auch darüber hinaus; vgl. Abell, 2007) in diesem Bereich wesentliche Schwächen auf: Stark heterogene Lernvoraussetzungen in Bezug auf das aus der Schule eingebrachte fachliche Wissen können im Studium nur unzureichend ausgeglichen werden. Im Gegenteil, es zeigt sich sogar eine mit zunehmender Studiendauer größere Streuung der Leistungen – obwohl angenommen werden könnte, dass Studierende mit schlechteren Voraussetzungen in den ersten Semestern aus dem Studium ausscheiden. Auch zeigen sich große Leistungsunterschiede zwi-

schen verschiedenen Studiengängen (z. B. Haupt-/Realschule vs. Gymnasium; vgl. Riese, 2009, 2010). Andere Studien stellen bei Lehrpersonen dieselben konzeptionellen Fehlvorstellungen wie bei Schülern fest, die bis hin zu groben fachlichen Mängeln reichen (vgl. Girwitzt, Kurz & Kautz, 2003; Abell, 2007; de Jong, Korthagen & Wubbels, 1998).

Um diese Probleme mittelfristig angehen zu können, ist es nötig, adaptive – d. h. auf die spezifischen Schwächen der Studierenden ausgerichtete – Interventions- und Fördermaßnahmen zu installieren, welchen der Einsatz eines differenzierten und präzisen Diagnoseinstruments vorausgehen muss.

Die Voraussetzung für eine valide Kompetenzmessung wiederum stellt dabei ein Kompetenzmodell dar, das Auskunft über die Struktur und eventuell auch die Entwicklung oder Stufung innerhalb des betrachteten Kompetenzbereichs gibt. Modelle dienen hierbei sowohl der breiten Abdeckung der jeweiligen Domäne im Zusammenhang mit der Entwicklung eines Kompetenztests als auch der strukturierten Interpretation der Testresultate (vgl. Klieme & Leutner, 2006). Eine Kompetenzmessung muss also immer bei der theoriegeleiteten Erstellung eines Kompetenzmodells anfangen.

Die bisher vorhandenen Modelle fachlicher Kompetenz und die entsprechenden Diagnoseinstrumente zielen jedoch vor allem auf die Schülerebene ab oder erlauben auf Lehrerebene nur globale Modellprüfungen (siehe Abschnitt 2.1). In diesem Beitrag soll daher aufbauend auf dem Paderborner Kompetenzmodell (Riese,

2009) ein Strukturmodell der fachlichen Kompetenz angehender Physiklehrkräfte erarbeitet werden, in dem hypothetisch schwierigkeiterzeugende Aufgabenmerkmale formuliert werden können. Dabei soll das Modell sowohl die Nutzung von Fachwissen als auch den Umgang mit Experimenten und Modellen abdecken.

Die Modellentwicklung stellt hierbei den ersten Arbeitsschritt in einem Projekt dar, in dem aufbauend auf dem hier vorgestellten Modell ein Kompetenztest entwickelt, pilotiert und für eine Querschnittserhebung unter Physik-Lehramts-Studenten eingesetzt wird, mit dem Ziel, Niveaustufen der physikalischen Fachkompetenz zu beschreiben. Durch Einordnung von Studenten in diese Niveaus kann ein Ansatzpunkt für differenzierte Fördermaßnahmen geschaffen werden.

2 Modellierung physikalischer Fachkompetenz

Kompetenzmodelle dienen der genaueren Bestimmung (professioneller) Kompetenzen, die notwendig zur Bewältigung bestimmter Anforderungen innerhalb einer Domäne sind. Sie stellen eine Voraussetzung für valide Kompetenzmessung dar (vgl. Klieme & Leutner, 2006; Hartig & Jude, 2007). Nach Schecker & Parchmann (2006) können hierbei verschiedene Ausprägungen von Kompetenzmodellen unterschieden werden: Auf der einen Seite beschreiben normative Kompetenzmodelle Voraussetzungen, über die ein Lernender verfügen soll, um gewisse Aufgaben und Probleme lösen zu können.

Davon abgrenzen lassen sich deskriptive Modelle, die lediglich die typischerweise vorliegenden Voraussetzungen für die Aufgabenlösung beschreiben. Auf der anderen Seite kann zwischen Kompetenzniveaumodellen, in denen die Differenzierung unterschiedlicher Ausprägungsgrade eines Konstrukts im Vordergrund steht, Kompetenzentwicklungsmodellen, die Sequenzen eines effektiven Kompetenzentwicklungsprozesses beschreiben, und Kompetenzstrukturmodellen unterschieden werden, welche ihrerseits Aussagen zur Binnenstruktur und Dimensionalität von Kompetenzbereichen machen (Klieme, Maag-Merki & Hartig, 2007). Für eine deduktive Testkonstruktion werden solche Strukturaussagen benötigt, um eine gleichmäßige und repräsentative Verteilung der zu entwickelnden Testitems innerhalb eines Anforderungsbereichs sicherzustellen und eine differenzierte Auswertung entlang einzelner Teildimensionen zu ermöglichen (vgl. Hartig & Jude, 2007). Zur Entwicklung solcher Kompetenzstrukturmodelle stehen verschiedene Zugänge und Methoden zur Verfügung. Zum einen können Kompetenzmodelle auf der Basis empirischer, eher induktiver Methoden wie Anforderungsanalysen oder Expertenbefragungen generiert werden. Zum anderen ist eine Modellierung auf der Grundlage normativer Annahmen, fachsystematischer Konzeptionen oder empirisch fundierter Konzeptualisierungen aus Untersuchungen angrenzender Domänen durch deduktives Ableiten aus bestehenden Kompetenzkategorien möglich (vgl. Klieme et al., 2003; Terhart, 2002; Schaper, 2009). Bei der Mo-

dellierung fachlicher Kompetenzaspekte wird dabei in der Regel eher auf deduktive Strategien zurückgegriffen (vgl. Schecker & Parchmann, 2006; Blömeke, Kaiser & Lehmann, 2008), da sich deren inhaltliche Strukturelemente bis zu einem gewissen Grad aus der sachlogischen Struktur der Fachdisziplin ableiten lassen (wenngleich die konkrete sachlogische Reihenfolge in Vermittlungssituationen im Detail natürlich deutlich variieren kann). Üblicherweise verweisen Autoren hier jeweils auf typische Inhaltsbereiche bzw. Themen der Physik (z. B. Borowski et al., 2010; Riese, 2009) oder Leitideen (z. B. Kauertz, 2010). Darüber hinaus ist bei der Modellierung auch die Frage der Dimensionierung zu stellen, in Bezug auf physikalisches Fachwissen finden sich häufig mehrdimensionale Modelle (etwa Inhaltsbereiche, kognitive Aktivitäten und Niveaustufen, vgl. Riese, 2009). Letztlich gibt es aber keine allgemein eindeutigen empirischen Belege für die Dimensionierung oder die Anzahl der Bereiche pro Teildimension, zumal sich Subdimensionen nicht immer explorativ ausmachen lassen (für COACTIV vgl. Krauss et al., 2008). Unabhängig davon ist bei jeder Untersuchung stets zu prüfen, ob sich die theoretisch begründeten Klassifikationen des Modells empirisch bewähren.

Neben dem Paderborner Kompetenzmodell (vgl. Abschnitt 3.1), auf dem in dieser Arbeit aufgebaut werden soll, existieren verschiedene Ansätze zur Modellierung fachlicher Kompetenz, die im Folgenden vorgestellt werden. Neben den umfassenden Strukturmodellen stehen auch Teilmodelle zur Aufgabenkomplexität (maßgeblich zur

Beschreibung schwierigkeiterzeugender Merkmale) und zur Erkenntnisgewinnung (zum Umgang mit Experimenten und Modellen, die im Paderborner Modell noch nicht abgedeckt sind) im Fokus.

2.1 Strukturmodelle fachlicher Kompetenz

An Modellen auf Schülerebene sind zunächst die Nationalen Bildungsstandards Physik (KMK, 2005) als normatives Modell sowie mehrere daran aufbauende oder zumindest anschlussfähige Modelle zu nennen (vgl. Schecker & Parchmann, 2006; ESNaS-Projekt zur Evaluation der Bildungsstandards in den naturwissenschaftlichen Fächern Sekundarstufe I: Kauertz, Fischer, Mayer, Sumfleth & Walpuski, 2010). Modelle auf Schülerebene können allerdings nur begrenzt auf die Ebene der Lehrerbildung übertragen werden, da für die Unterrichtsgestaltung vor allem das Fachwissen vom höheren Standpunkt relevant ist (vgl. Krauss et al., 2008) und das didaktische Repertoire der Lehrkräfte vor allem von der Breite und Tiefe ihres konzeptionellen (und damit nicht dem faktenbezogenen) Fachverständnisses abhängt (vgl. Baumert & Kunter, 2006). Lehrkräfte müssen demzufolge über deutlich umfangreicheres Wissen verfügen, als sie ihren Schülerinnen und Schülern vermitteln wollen bzw. dass Schülerinnen und Schülern üblicherweise vermittelt wird. Mit einem Modell auf Schülerebene werden somit wesentliche Teile der Handlungskompetenz von Lehrkräften systematisch ausgeblendet.

Auf der Ebene der Lehrkräfte ist aus dem Bereich der Mathematik die MT21-Studie (vgl. Blömeke et al., 2008) zu nennen. Hier wird ein Modell verwendet, das das Fachwissen nach Stufen (*Sekundarstufe I, Sekundarstufe II, Schulmathematik vom höheren Standpunkt, universitäre Mathematik*), nach Verknüpfungsleistungen mit dem Unterrichtshandeln (*keine, eine einfache, eine anspruchsvolle, mehrere*) und der Art der kognitiven Auseinandersetzung (*keine, einfache und komplexe kognitive Anstrengung*) differenziert. Auf der Grundlage dieses Strukturmodells gelingt die Bildung von Kompetenzniveaus allerdings nur eingeschränkt. Es werden auf Grundlage der empirisch ermittelten Aufgabenschwierigkeiten einzelne Abschnitte aller drei Dimensionen in a priori nicht vorhersehbarer Weise zur Beschreibung von Niveaus herangezogen. Dieses Vorgehen wäre prinzipiell mit jedem Kompetenzstrukturmodell möglich, erschwert aber die systematische Charakterisierung der Niveaus erheblich und ist auch aus anderen theoretischen Gesichtspunkten nachteilig (z. B. Übertragbarkeit auf andere Itembestände, vgl. Hartig, 2007).

Die entsprechende Nachfolgestudie TEDS-M (vgl. Blömeke, Kaiser, Lehmann, 2010) unterscheidet drei theoretische Schwierigkeitsgrade: *Elementares Niveau, Mittleres Niveau* und *Fortgeschrittenes Niveau*. Sie beziehen sich jeweils auf die Klassenstufen, in denen die jeweiligen Probleme eine Rolle spielen. Diese Schwierigkeitsgrade mussten empirisch zum Teil korrigiert werden. Die schließlich festgelegten Kompetenzniveaus im

Fachwissen basieren allein auf Schwellwerten in den empirischen Daten.

Ebenfalls aus dem Bereich der Mathematik wurde im Rahmen der COACTIV-Studie (vgl. z. B. Krauss et al., 2008) das professionelle Wissen von Mathematik-Lehrerinnen und -Lehrern modelliert. Das Modell differenziert eher bei den fachdidaktischen Anteilen und unterteilt das Fachwissen lediglich eindimensional in drei Ebenen (*Mathematisches Alltagswissen, Schulstoff, Tieferes Verständnis der Fachinhalte des Curriculums der Sekundarstufe, reines Universitätswissen*), von denen die 3. Stufe für den Unterricht am relevantesten eingestuft und vorrangig getestet wird.

Im Bereich der Naturwissenschaften entstand neben den Paderborner Arbeiten (vgl. Abschnitt 3.1) in neuerer Zeit das ProwiN-Projekt (Borowski et al., 2010), in welchem ein Strukturmodell auf Lehrerebene für die drei Naturwissenschaften Biologie, Chemie und Physik entwickelt und genutzt wird. Im Fokus steht hier die Struktur des Professionswissens bzw. der Zusammenhang zwischen Fachwissen, fachdidaktischem und pädagogischem Wissen untereinander sowie deren jeweiliger Zusammenhang mit dem unterrichtlichen Handeln der Lehrkräfte. Das Strukturmodell für das Fachwissen fokussiert dementsprechend auf die Wissensbereiche, für die ein starker Zusammenhang mit dem Unterrichtshandeln der Lehrkräfte und dem Lernerfolg der Schüler angenommen wird. Das sind in der Sprechweise des Paderborner Modells vor allem die Fach-Stufen (vgl. Abschnitt 3.2) des Schulwissens und des vertieften Wis-

sens. Bei Modellen, in denen (wie im hier vorgestellten Modell) der Fokus stärker auf der Evaluation und Diagnose innerhalb der universitären Ausbildungsphase liegt, erscheint die zusätzliche Aufnahme des rein universitären Wissens angebracht, da dieser Bereich insbesondere bei den höheren Lehrämtern in den ersten Semestern im Schwerpunkt der Lehre steht. Ob das universitäre Wissen tatsächlich Relevanz im Hinblick auf das professionelle Handeln von Physiklehrkräften besitzt, kann zum jetzigen Zeitpunkt nicht abschließend beurteilt werden (vgl. auch Riese, 2009). Insbesondere daher erscheint es somit geboten, diesen Bereich in einem hochschulbezogenen Modell der fachlichen Kompetenz von (zukünftigen) Lehrkräften zu berücksichtigen, um auf der Grundlage entsprechender Testinstrumente bzw. Kompetenzmessungen hier einen Erkenntnisfortschritt zu erzielen.

Im Rahmen des ProwiN-Projektes wird die professionelle Kompetenz von Lehrkräften als komplexes, allumfassendes Konstrukt (fachliches, fachdidaktisches, pädagogisches Wissen, Beliefs, motivationale Orientierungen) modelliert und gemessen, so dass bei einer handhabbaren Testdauer keine differenzierten Aussagen auf Teilskalenebene zu erwarten sind. Auch die bisherigen Paderborner Arbeiten erlauben keine präzisen Kompetenzdiagnosen (vgl. Abschnitt 3.1). Damit sind beide auf Hochschulebene im Bereich Physik verfügbaren Kompetenzmodelle und Messinstrumente noch nicht differenziert genug und liefern im Wesentlichen lediglich Aussagen auf Gesamtscore-Ebene für das Fachwissen. Es können aber keine

differenzierten Aussagen über die spezifischen Stärken und Schwächen der jeweiligen Probanden gemacht werden.

2.2 Modelle der Aufgabenkomplexität

Als zentraler Aspekt eines Modells, das Testitems a priori nach ihrer Aufgabenschwierigkeit anordnen will, hat sich in den letzten Jahren die Aufgabenkomplexität herausgestellt. Diese ordnet die Aufgaben aufgrund von strukturellen Merkmalen der zur Lösung notwendigen kognitiven Vorgänge verschiedenen Komplexitätsstufen zu.

Als wegweisend kann hier das Modell von Kauertz (2008) für die Schülerebene im Fach Physik genannt werden. Es verwendet die Stufen *Fakt*, *Fakten*, *Zusammenhang*, *unverbundene Zusammenhänge*, *verbundene Zusammenhänge*, *übergeordnete Konzepte*. Weitere zur Aufgabenkonstruktion verwendete Eigenschaften sind *Leitidee*, *Kognitive Aktivität* und *Kontext*, die jedoch jeweils nur geringen Einfluss auf die Aufgabenschwierigkeit haben. Die empirische Überprüfung ergab hier im Wesentlichen eine Anordnung der mittleren Aufgabenschwierigkeiten nach diesen Komplexitäts-Stufen, wobei allerdings vor allem die Stufen *Fakten* und *unverbundene Zusammenhänge* empirisch höhere Schwierigkeiten aufweisen, als die Reihung der Stufen vermuten ließe. Der mittlere Schwierigkeitsunterschied zwischen den Stufen *verbundene Zusammenhänge* und *übergeordnete Konzepte* fällt ausgesprochen gering aus.

Komplexitätsmodelle mit einer ähnlichen Struktur wie bei Kauertz (2008) finden vielfältige Anwendung, z. B. im ESNaS-Projekt (Kauertz et al., 2010). Besonders im ESNaS-Projekt ist die Kombination der Komplexitätsstufen *ein Fakt*, *zwei Fakten*, *ein Zusammenhang*, *zwei Zusammenhänge* und *übergeordnetes Konzept* mit einer Skala kognitiver Prozesse mit den Ausprägungen *Reproduzieren*, *Selegieren*, *Organisieren* und *Integrieren* bemerkenswert, die genutzt wird, um die in den Bildungsstandards definierten, nominalen Anforderungsbereiche auf empirisch besser handhabbare, ordinale Dimensionen abzubilden (vgl. Walpuski et al., 2010).

Aus der Chemiedidaktik ist, ebenfalls auf Schülerebene, das Modell von Bernholt (2010) zu nennen, welches sich auf das domänenunspezifische Modell der hierarchischen Komplexität (Commons, Trudeau, Stein, Richards, Krause, 1998) bezieht. Letzteres definiert allgemein: "Actions at a higher order of hierarchical complexity: (a) Are defined in terms of the actions at the next lower order of hierarchical complexity; (b) Organize and transform the lower order actions; (c) Produce organizations of lower order actions that are new and not arbitrary and cannot be accomplished by those lower order actions alone." (Commons et al., 1998, S. 240) Abzugrenzen ist die hierarchische von der nichthierarchischen Komplexität, welche nur die Aneinanderreihung mehrerer Tasks gleicher hierarchischer Komplexität beschreibt. Nichthierarchische Komplexität erhöht zwar ebenfalls die Aufgabenschwierigkeit, da sie im Sinne einer Cognitive Load Theory (vgl. Paas, Renkl, Sweller, 2004) eine

größere Auslastung des Arbeitsgedächtnisses bedingt, macht aber keine Aussage über eine Fähigkeit zum Umgang mit untereinander stärker verknüpften kognitiven Inhalten.

Bernholt (2010) konkretisiert nun dieses allgemeine Modell zu den fünf Stufen *unreflektiertes Erfahrungswissen, Fakten, Prozessbeschreibungen, lineare Kausalität* und *multivariate Interdependenz*.

Die Modelle von Bernholt und Kauertz scheinen analoge Aussagen zur Aufgabenkomplexität zu treffen, unterscheiden sich aber in einigen, für das vorgestellte Projekt wesentlichen Punkten. Zunächst findet sich im Modell von Kauertz jeweils der nichthierarchische Komplexitätsunterschied zwischen *Fakt* und *Fakten* sowie zwischen *Zusammenhang* und *unverbundenen Zusammenhängen*. Gleiches findet sich zwischen den Komplexitäts-Stufen bei ESNaS. Dies erzeugt zwar eine geringere Lösungswahrscheinlichkeit (*Fakten* und *unverbundene Zusammenhänge* sind empirisch jeweils wesentlich schwieriger als im Rahmen des Modells erwartet), was unter dem Aspekt der Cognitive Load Theory (vgl. Paas et al., 2004) unmittelbar einleuchtet. Es ist aber nicht unmittelbar klar, ob dies tatsächlich die für die Perspektive eines Assessments interessante Information über die Probandenfähigkeit ist. Kalyuga et al. (2003) gehen davon aus, dass für eine erfolgreiche Förderung viel mehr von Interesse ist, mit welchem Verknüpfungsgrad (im Sinne hierarchischer Komplexität) ein Lernender umgehen kann. Für Rückschlüsse auf diesen Aspekt sollte also der nichthierarchische Komplexitätsaspekt der zu entwickelnden

Testitems klein gehalten und möglichst nicht variiert werden.

Aufgaben, die im Kauertz-Modell auf den beiden höchsten Stufen (*verbundene Zusammenhänge, übergeordnete Konzepte*) verortet sind, würden im Bernholt-Modell auf derselben Stufe (*multivariate Interdependenz*) zusammengefasst (vgl. Abb. 1). Aufgrund der sehr ähnlichen mittleren Aufgabenschwierigkeiten bei Kauertz erscheint dies aber nur als minimaler Informationsverlust.

Ein weiterer Unterschied zwischen den beiden Modellen besteht in ihrer Fähigkeit, auch schwache Probanden abzubilden. Eine Ausblendung der leistungsschwächsten 25 % der Probanden, wie dies bei PISA der Fall war (kritisiert z. B. in Labudde et al., 2009, S. 350), ist hier sicher nicht zielführend, da in einem Assessment, das als Grundlage für eine individuelle Förderung dienen soll, gerade über die schwächeren Probanden differenzierte Aussagen getroffen werden müssen. Während das Modell von Bernholt unterhalb der *Fakten* das *unreflektierte Erfahrungswissen* enthält, welches sich gegenüber den *Fakten* durch einen Mangel an (wissenschaftlicher) Reflexion auszeichnet, ist das Modell von Kauertz vor allem durch den Einbezug weiterer Dimensionen (wie z. B. bei ESNaS) erweiterbar. Auch wenn dieses Modell in der Lage ist, schwächere Probanden differenziert abzubilden (Walpuski et al., 2010), bleibt ein Ökonomievorteil zu Gunsten der eindimensionalen Modellierung nach Bernholt.

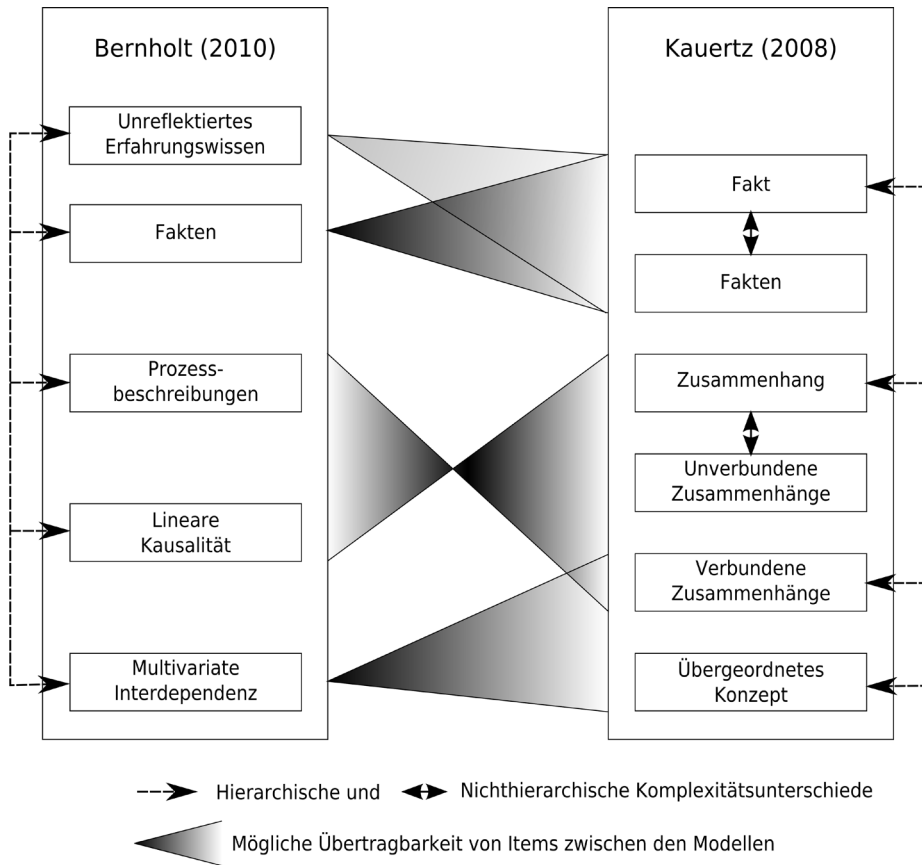


Abb. 1: Die Komplexitätsmodelle von Bernholt (2010) und Kauertz (2008) gegenübergestellt.

2.3 Modelle zum Umgang mit Experimenten und Modellen

Das Paderborner Kompetenzmodell deckt den Bereich der Erkenntnisgewinnung bzw. der Erkenntnismethodik nur marginal ab (vgl. Abschnitt 3.1). Im Folgenden wird daher ein möglicher Ansatz vorgestellt, in welchem Erkenntnisgewinnung als Problemlöseprozess aufgefasst wird. Er liegt u. a. auch dem entsprechenden Teilmodell im ESNaS-Projekt (vgl. Kauertz et al., 2010) zu Grunde.

Rahmenmodell

Nach Mayer (2007) lässt sich der Bereich der Erkenntnisgewinnung in die Kompetenzbereiche *Charakteristika der Naturwissenschaften* (Nature of Science), *Wissenschaftliche Erkenntnismethoden* (Scientific Inquiry) und *Praktische Arbeitstechniken* (Practical Work) unterteilen. Allerdings ist es nicht für alle Bereiche gleichermaßen möglich, eine praktikable und valide Kompetenzmessung vorzunehmen. Während sich der erstgenannte

Bereich noch verhältnismäßig valide mit Paper-Pencil-Test erfassen lässt (z. B. durch Likertskalen), ist dies insbesondere für den letztgenannten Bereich weitaus problematischer. Im Falle der Experimentierkompetenz zweifeln Schreiber, Theyßen & Schecker (2009) die Testbarkeit in einem Paper-Pencil-Test prinzipiell an, da prozessbezogene Aspekte des Experimentierens, die wesentlich für diese Form der Erkenntnisgewinnung sind, systematisch nicht erfasst werden (können). Das von ihnen beschriebene Projekt *eXkomp* nutzt daher andere Testverfahren unter Einsatz von computerbasierten Experimentierumgebungen. Ein schriftlicher Kompetenztest, wie er in dem hier vorgestellten Projekt entwickelt werden soll, ist zwangsläufig Einschränkungen unterworfen und kann nicht alle Bereiche der Erkenntnisgewinnung gleichermaßen gut abdecken.

Das SDDS-Modell

Das Modell des Science Discovery as Dual Search von Klahr (2000) fasst wissenschaftliche Erkenntnisgewinnung als Problemlösen (vgl. auch Mayer, 2007) auf und betrachtet so lediglich die kognitiven Anteile einer umfassenderen Experimentierkompetenz, wie sie z. B. *eXkomp* erfassen würde. Der Erkenntnisweg wird modelliert als Suche in zwei abstrakten Räumen: Dem Hypothesen-Raum, der alle möglichen Hypothesen zur gewählten Fragestellung enthält, diese ordnet und miteinander in Beziehung setzt sowie dem Experiment-Raum, der in analoger Weise die möglichen Experimente enthält.

Zu wissenschaftlicher Erkenntnis gelangt man nun auf einem Weg, der die drei Hauptkomponenten des SDDS-Modells durchläuft (vgl. Abb. 2). Die Suche im Hypothesen-Raum generiert zunächst die zur Fragestellung passenden Hypothesen bzw. wählt einige, plausibel erscheinende Hypothesen aus. Im zweiten Schritt werden diese dann getestet, indem ein passendes Experiment aus dem Experiment-Raum ausgewählt wird. Passend heißt hier, dass es geeignet ist, in einem dritten Schritt, der Analyse von Evidenzen, Rückschlüsse auf die Hypothesen zu ziehen, also Hypothesen auszuschließen, zu bestätigen oder weitergehende Operationen auf dem Hypothesen-Raum durchzuführen (für eine umfassende Übersicht siehe van Joolingen & de Jong, 1997). Für ein gegebenes Problem lassen sich diese Hauptkomponenten gegebenenfalls noch detaillierter auffächern, wie dies z. B. von Klahr & Dunbar (1988) oder Hammann (2007) gezeigt wird. Auch andere, im wissenschaftlichen Alltag vorkommende Abläufe wurden von Klahr & Dunbar (1988) im Rahmen dieses Modells beschrieben.

Operationalisierung

Neben der nicht-naturwissenschaftsspezifischen Operationalisierung (z. B. Klahr & Dunbar, 1988) ist vor allem die Operationalisierung von Hammann, Phan & Bayrhuber (2007) aus dem Bereich der Biologie zu erwähnen. Beide Ansätze ordnen jeweils Aufgabenstellungen deutlich einer der drei Hauptkomponenten zu. Für das Testen von Hypothesen wird jeweils

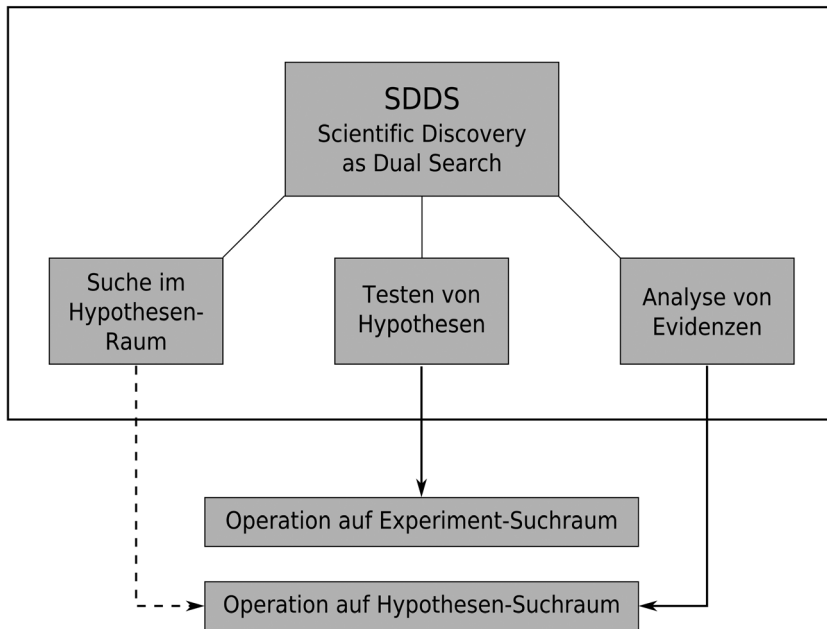


Abb. 2: Die drei Hauptkomponenten des SDDS-Modells (oben) und die beiden abstrakten Suchräume, auf denen bei der Aufgabenlösung jeweils operiert wird (unten).

eine Experimentierumgebung vorgegeben oder wenigstens skizziert, in deren Rahmen ein passendes Experiment ausgewählt werden muss. Für die Analyse von Evidenzen wird andersherum ein Ausgang eines Experiments vorgegeben und Folgerungen daraus für die zur Diskussion stehenden Hypothesen gefordert. Auffällig ist die Operationalisierung der ersten Hauptkomponente. Während Klahr & Dunbar (1988) von ihren Probanden explizit die Generierung von Hypothesen zu einer Fragestellung fordern, geben Hammann et al. (2007; vgl. auch Hammann, Phan, Ehmer & Bayrhuber, 2006) ein Experiment vor und stellen die Frage, welche Hypothese diesem zugrunde liegen könnte.

Auf diese Weise wird also die erste Hauptkomponente strukturell sehr ähnlich der dritten Hauptkomponente operationalisiert: In beiden Fällen muss von einem gegebenen Experiment auf Auswirkungen auf den Hypothesen-Raum geschlossen werden. Der Unterschied besteht lediglich darin, dass einmal der Ausgang des Experiments bekannt und einmal unbekannt ist.

Empirische Ergebnisse

Hammann et al. (2007) finden im Rahmen einer Faktorenanalyse im Wesentlichen zwei Faktoren, die sich mit Operationen auf dem Hypothesen- und auf dem Experiment-Suchraum beschreiben lassen.

Nimmt man eine Drei-Faktoren-Lösung an, erhält man zwei Faktoren, die sich mit Operationen auf dem Hypothesen-Suchraum beschreiben lassen, wobei eine Differenzierung nach Inhaltsbereichen geschieht. Der dritte Faktor lässt sich wiederum als Operationen auf dem Experiment-Suchraum beschreiben. Die beiden Hauptkomponenten, die auf dem Hypothesen-Suchraum operieren, lassen sich also im Rahmen dieser Faktorenanalyse empirisch nicht trennen.

Die Fähigkeiten der Probanden korrelieren hier nur sehr schwach mit ihren Leistungen im Fachwissen in den jeweiligen Inhaltsbereichen. Außerdem zeigt sich, dass das Testen von Hypothesen, also die Operation auf dem Experiment-Suchraum, etwas schwieriger zu sein scheint als die beiden anderen Itemgruppen.

3 Modellentwicklung

Da sich ein (naturgemäß eher grobes) Modell der physikalischen Fachkompetenz im Hinblick auf dessen inhaltsbezogene, domänenspezifische Strukturierung im Wesentlichen an zentralen, sachlogischen Elementen bzw. typischen Themen der betreffenden Fachdomäne orientiert, wird in der Regel ein deduktiv-normatives Vorgehen bei inhaltlichen Aspekten der Modellentwicklung gewählt (s. o., vgl. auch Schecker & Parchmann, 2006; Blömeke, Kaiser & Lehmann, 2008).

Aufbauend auf dem Paderborner Kompetenzmodell (Riese, 2009) soll im Folgenden ein Kompetenzstrukturmodell

entwickelt werden, das bekannte Probleme des Paderborner Modells behebt, um ein Teilmodell zur naturwissenschaftlichen Erkenntnisgewinnung erweitert ist und die hypothetische Formulierung schwierigkeitserzeugender Aufgabenmerkmale ermöglicht.

3.1 Ausgangspunkt: Das Paderborner Kompetenzmodell

Rahmenmodell

Das in der Arbeit von Riese (2009) entwickelte Strukturmodell der professionellen Kompetenz von Physiklehrkräften (vgl. Abb. 3) definiert Kompetenz gemäß Weinert (2001) als die bei Individuen verfügbaren oder von ihnen erlernbaren kognitiven Fähigkeiten und Fertigkeiten, bestimmte Probleme zu lösen, sowie die damit verbundenen, motivationalen, volitionalen und sozialen Bereitschaften und Fähigkeiten, die Problemlösungen in variablen Situationen erfolgreich und verantwortungsvoll nutzen zu können. In Anlehnung an Shulman (1986), Bromme (1997) und Baumert & Kunter (2006) wird ein Rahmenmodell verwendet, das auf der einen Seite kognitive Fähigkeiten und Professionswissen (ausdifferenziert in fachliches, fachdidaktisches und pädagogisches Wissen) und auf der anderen Seite motivationale, volitionale und soziale Bereitschaften und Fähigkeiten (ausdifferenziert in Belief Systems und motivationale Orientierungen) umfasst. Das Fachwissen, das im vorliegenden Artikel vertieft betrachtet werden soll, stellt



Abb. 3: Rahmenmodell der professionellen Handlungskompetenz von Physiklehrkräften von Riese (2009) in Anlehnung an Baumert & Kunter (2006), Weinert (2001), Bromme (1997), Shulman (1986).

also als Teil des Professionswissens nur einen Teilaspekt eines umfassenderen Kompetenzkonstrukts dar. Die Operationalisierung geschieht bei Riese (2009) in Form von Paper-Pencil Leistungstests für die drei Bereiche des Professionswissens (ca. 30 Minuten Testzeit je Bereich) und in Form von 5-Stufigen Likert-Skalen für die Beliefs und motivationalen Orientierungen (10 Minuten).

Strukturmodell Fachwissen

Für das Fachwissen beschränkt sich Riese auf den Inhaltsbereich Mechanik. Dieser Inhaltsbereich wird bereits in den fachlichen Einführungsveranstaltungen thematisiert und ist damit bei Studenten prinzipiell ab dem ersten Semester testbar. Zudem werden üblicherweise wesentliche Konzepte der Physik (Kraft, Impuls,

Energie etc.) im Zusammenhang mit der Mechanik eingeführt, so dass hier die Grundlage für das Verständnis weiterer Bereiche geschaffen wird. Diese Tatsache spiegelt sich darin wider, die Kompetenz im Inhaltsbereich Mechanik ein guter Prädiktor für die Kompetenz in anderen Inhaltsbereichen ist (vgl. z. B. Friege & Lind, 2004), auch bei Riese (2009) konnte ein Kompetenzzuwachs mit zunehmender Semesterzahl beobachtet werden, obwohl nur typische Inhalte des ersten Semesters getestet wurden. Darin zeigt sich, dass Physiklernen hoch kumulativ ist.

Als Grundlage für die zielgerichtete Konstruktion der 28 Testitems wurde von Riese (2009) ein 3-dimensionales Strukturmodell verwendet. Neben einer Inhalts-Dimension mit den Ausprägungen *Kinematik*, *Kraft*, *Energie*, *Impuls* und einer Dimension mit den kognitiven Aktivitäten *Reproduzieren*, *Verstehen* sowie

Beurteilen und Analysieren umfasst es die Niveaustufen *Schulwissen*, *vertieftes Wissen* (im Sinne eines Schulwissens vom höheren Standpunkt in Anlehnung an die Definitionen bei Krauss et al. (2008) und Blömeke, Kaiser & Lehmann (2008)) und *universitäres Wissen*. Diese stellen jedoch keine Kompetenzniveaus im Sinne des vorgestellten Projekts dar, wie im Folgenden dargelegt werden soll.

Grenzen des bisherigen Modells

Obwohl der Test von Riese (2009) insgesamt eine gute Testgüte aufweist (Cronbachs Alpha 0,81 bei 28 Items im Fachwissen), ist es nicht gelungen, im Rahmen dieses Modells Kompetenzniveaus und Aufgabenschwierigkeiten zu verorten. Die Faktorstruktur der Niveaustufen *Schulwissen*, *vertieftes Wissen* und *universitäres Wissen* konnte zwar empirisch bestätigt werden, sie weisen aber keine in dieser Reihenfolge steigende Schwierigkeit auf. Die Faktorstruktur der Dimension der kognitiven Aktivität konnte darüber hinaus nicht bestätigt werden, was wahrscheinlich auch mit schlecht abgrenzenden Definitionen in diesem Bereich zusammenhängt.

Weiterhin bildet das Modell Formen der naturwissenschaftlichen Erkenntnisgewinnung (Experimentieren, Umgang mit Modellen) nur marginal ab, obwohl dies vielfach als zentraler Bestandteil physikalischer Kompetenz gesehen wird (vgl. KMK, 2005; Schecker & Parchmann, 2006; Metzger & Labudde, 2007; Kauertz et al., 2010).

Diese bislang unzureichenden Aspekte im vorhandenen Modell sollen in den folgenden Abschnitten bei der Darstellung einzelner Teildimensionen im Zusammenhang mit der Modellentwicklung besonders berücksichtigt werden. Damit sollen die bisherigen Erfahrungen genutzt werden und es soll gewährleistet werden, dass ein auf dem Modell aufbauender Test Ergebnisse liefert, die zur Definition und Charakterisierung von Kompetenzniveaus geeignet sind.

3.2 Fach-Stufen

Riese (2009) benennt drei Niveaustufen (*Schulwissen*, *vertieftes Wissen*, *universitäres Wissen*), welche aber, wie oben dargestellt, keine Niveaustufen mit steigender Schwierigkeit im Sinne eines Kompetenzniveaumodells darstellen. Daher wird hier im Folgenden von Fach-Stufen die Rede sein. Die bei Riese getroffenen Definitionen werden wie folgt präzisiert: Unter *Schulwissen* verstehen wir, in Anlehnung an die Definition bei Krauss et al. (2008), eine Anforderung, die ein durchschnittlicher Schüler am Ende der Sekundarstufe I erfüllen können sollte, operationalisiert durch Aufgaben, die in der Schule verwendet werden können. Als Referenzmaßstab dienen hier die Bildungsstandards Physik für den mittleren Schulabschluss (KMK, 2005).

Unter *universitärem Wissen* verstehen wir vollständig von der Schule losgelöstes Wissen (vgl. Krauss et al., 2008, und Blömeke, Kaiser & Lehmann, 2008). Die entsprechenden Items können auch von sehr

guten Schülern in der Regel nicht gelöst werden.

Das *vertiefte Wissen* schlägt die Brücke zwischen den vorherigen beiden Stufen und vermindert so die „doppelte Diskontinuität“ (scheinbare Irrelevanz des schulischen Wissens im Fachstudium und ebenso scheinbare Irrelevanz des an der Universität erworbenen Fachwissens im späteren Lehrerberuf; vgl. Klein, 1933) zwischen Schule und Hochschule. Aus Klein (1933), Kirsch (1987) und den sich darauf beziehenden Studien mit einem ähnlichen Konstrukt (Blömeke et al., 2008; Krauss et al., 2008; Riese, 2009) ergibt sich folgende (notwendigerweise unvollständige) Liste von Charakteristika dieser Wissensstufe:

- i. Explizite Kombination von Schul- und universitärem Wissen.
- ii. Anwendung von universitären Denkweisen auf typische Probleme der Schulphysik.
- iii. Systematisierung und Herstellung von Zusammenhängen des Schulwissens vor dem Hintergrund universitärer Physik; eine umfassende Sicht auf die Physik als Ganzes (im Sinne Kleins für die Mathematik).
- iv. Einübung (bzw. Erfassung) von elaborierten Denk-, Sprech- und Verhaltensweisen im Umgang mit Physik anhand von Gegenständen der Schulphysik.
- v. Reflexion von Bedeutung, Genese und Verwendung von Begriffen der Schulphysik.
- vi. Als Konkretisierung der vorhergehenden beiden Punkte: Items, die gezielt Fehlvorstellungen abtesten.

Diese Stufe kann von sehr guten Schülern (z. B. in einem Leistungskurs) in Teilen erreicht werden, ist aber charakteristisch für das vernetzte Verständnis, das Baumert & Kunter (2006) als Basis für das Erklärungsrepertoire einer Lehrkraft ansehen. Ob die obige Auflistung präzise genug ist, um Testitems zweifelsfrei charakterisieren zu können, wird im Zusammenhang mit der Itementwicklung im beschriebenen Projekt empirisch überprüft (Bestimmung der Interraterreliabilität).

3.3 Inhaltsbereiche

Bei der Auswahl der Inhaltsbereiche erscheinen die Einschränkungen durch die Curricula in den Haupt-/Realschul-Lehramtsstudiengängen am stärksten. Hier werden, verglichen mit anderen Studiengängen, die wenigsten Veranstaltungen zu schulfernen physikalischen Inhalten besucht.

Das Modell von Riese (2009) verwendet die vier Inhaltsbereiche *Kinematik*, *Kraft*, *Energie* und *Impuls*. Um eine breite Einsetzbarkeit des zu entwickelnden Modells und des darauf aufbauenden Testinstruments für verschiedene Studiengänge (Vollfach, Gymnasial- und Haupt-/Realschul-Lehramt) zu gewährleisten, ist es notwendig, dass die Inhaltsbereiche die Schnittmenge der Schul- und Hochschul-Curricula abdecken. Aus den weiter oben genannten Gründen (Abschnitt 3.1) und um Anschlussfähigkeit an das Modell von Riese zu gewährleisten, bleibt der Fokus dabei auf der Mechanik.

Der Inhaltsbereich *Schwingungen und Wellen* wird zwar auch in allen genann-

ten Curricula angesprochen und befasst sich mit einem System, das modellhaft auf andere Bereiche übertragen werden könnte und an dem auch komplexe Fragestellungen bearbeitet werden könnten. Allerdings ist hier problematisch, dass dieser Inhaltsbereich in der Regel erst in der Sekundarstufe II tiefergehend behandelt wird und dementsprechend in vielen Lehramtsstudiengängen für Haupt-/Realschule mutmaßlich nur am Rande vorkommt. Der Einbezug dieses Bereichs würde daher vermutlich zu einer Bevorzugung der höheren Studiengänge führen, was uns nicht angestrebt werden soll. Eine Erweiterung des Modells um weitere Inhaltsbereiche der Mechanik scheint mit Blick auf die geplante Generierung von Kompetenzniveaus allerdings auch nicht zwingend erforderlich zu sein, da sich in den Daten von Riese (2009) zeigt, dass die Lösungshäufigkeiten der einzelnen Inhaltsbereiche kaum voneinander abweichen; wo dies doch der Fall ist, kann von einem Curriculums-Effekt ausgegangen werden. Insofern ist auch im Hinblick auf die Entwicklung des Kompetenzstrukturmodells und des Testinstruments keine erschöpfende Abdeckung aller denkbaren Inhaltsbereiche nötig.

3.4 Komplexität

Um im Modell schwierigkeiterzeugende Aufgabenmerkmale abbilden zu können, wird die Aufgabenkomplexität als eine Dimension einbezogen. Aus den oben genannten Gründen scheint in Bezug auf das Projektziel das Modell von Bernholt (2010)

geeignet. Da jedoch die professionelle Fachkompetenz getestet wird, kann hier zunächst die unterste Stufe, das unreflektierte Erfahrungswissen, entfallen. Sollte sich bei der Pilotierung des Instruments allerdings die Notwendigkeit ergeben, die Schwierigkeiten nach unten zu erweitern, kann auf diese Stufe zurückgegriffen werden.

Eine Nutzung zweier Dimensionen im Zusammenhang mit Komplexität wie bei ES-NaS (Kombination von Komplexität und kognitiven Anforderungen, vgl. Walpuski et al., 2010) würde zwar prinzipiell eine differenziertere Beschreibung schwierigkeiterzeugender Merkmale ermöglichen, erscheint in Bezug auf das hier beschriebene Modell vor allem aus ökonomischen Gesichtspunkten unvorteilhaft. Gegenüber dem Modell von Bernholt bietet ein solches Vorgehen keine Vorteile im Hinblick auf die geplante Nutzung des Modells bzw. Instruments im Zusammenhang mit Förderungsmaßnahmen (vgl. auch Abschnitt 2.2).

Anstatt daher Komplexität als vierte Dimension in das vorhandene Paderborner Modell einzubauen, erscheint es somit zielführend, die bislang verwendete Dimension kognitiver Aktivität zu Gunsten der Komplexität fallen zu lassen:

Zum einen beschreiben Komplexität nach Bernholt und kognitive Aktivität konzeptionell ähnliche Konstrukte – nämlich eine Graduierung der kognitiven Anstrengung, die zur Aufgabenlösung erforderlich ist. Hier wäre eine Doppelung unnötig und stände einer prägnanten Charakterisierung von Niveaustufen im Wege.

Zum anderen enthält die kognitive-Aktivitäts-Dimension bei Riese nicht immer

trennscharf abgrenzende Definitionen, was dazu führt, dass Items mitunter nicht eindeutig nach ihrer kognitiven Aktivität klassifiziert werden können. Demgegenüber sind die Stufen bei Bernholt klarer abgrenzend definiert. Eine Kategorisierung der Fachwissen-Items aus dem Test von Riese nach beiden Skalen (s. Tabelle 1) zeigt zudem, dass Items zum Reproduzieren nur auf den unteren beiden und Items zum Beurteilen und Analysieren nur auf dem obersten Komplexitätsniveau nach Bernholt verortet sind. Die Komplexität verhält sich hier also bis zu einem gewissen Grad parallel zur kognitiven Aktivität. Items zum Verstehen finden sich auf allen Komplexitätsstufen, was mit dem Befund einer schlecht abgrenzenden Definition dieser Ausprägung korrespondiert. Die Aufgabenkomplexität nach Bernholt verhält sich außerdem orthogonal zu den Inhaltsdimensionen – es finden sich also Items zu allen Inhaltsbereichen mit allen Komplexitäten, was ebenfalls für dieses Vorgehen spricht.

Insgesamt scheint es also zielführend, Testitems zur Nutzung von Fachwissen mit den Dimensionen hierarchische Komplexität, Inhaltsbereich und Fach-Stufe zu beschreiben.

3.5 Umgang mit Experimenten und Modellen

Aufgrund der oben dargelegten Eigenschaften und der Passung mit dem Ziel dieser Untersuchung sollen die Bereiche Experimentieren und Modellieren unter Nutzung des SDDS-Modells in das zu entwickelnde Kompetenzstrukturmodell eingearbeitet werden. Für eine Adaption des SDDS-Modells ist zunächst zu klären, in wie weit damit tatsächlich der Bereich naturwissenschaftlicher Erkenntnisgewinnung voll abgedeckt werden kann. Das Experimentieren findet sich in der Operation auf dem Experiment-Suchraum. Das Modellieren, also die theoriebezogene Transformation eines fachlichen Ge-

Tab 1: Einstufung der Items im Fachwissen-Test von Riese im Hinblick auf deren kognitive Aktivität (Reproduzieren, Verstehen, Beurteilen und Analysieren) nach Riese (2009; dort findet sich auch eine Beschreibung der Stufen) und im Hinblick auf deren hierarchische Komplexität (Fakten, Prozessbeschreibungen, lineare Kausalität, multivariate Interdependenz) nach Bernholt (2010). Eingetragen ist jeweils die Anzahl der Items

	Fakten	Prozessbeschreibungen	Lineare Kausalität	Multivariate Interdependenz	Gesamt
Reproduzieren	7	2			9
Verstehen	4	4	2	4	14
Beurteilen und Analysieren				5	5
Gesamt	11	6	2	9	28

genstandes in ein materielles Objekt oder theoretisches Konstrukt (vgl. Justi & Gilbert, 2002), kann bei höheren Komplexitäten bei den Operationen auf dem Hypothesen-Suchraum verortet werden, da (mentale) Modelle letztlich komplexe Hypothesen-Systeme darstellen.

Auch die Prüfung anderer Aufgabestämme zur Erkenntnisgewinnung zeigt, dass das SDDS-Modell den relevanten Bereich gut abdeckt: Die vom IQB (2010) aus dem ESNaS-Projekt veröffentlichten Items sind problemlos den Hauptkomponenten zuzuordnen (und weisen für die erste Hauptkomponente sogar die selbe Besonderheit wie die Items von Hammann auf). Die von Hammann et al. (2006) gezeigte Aufgabe aus der TIMS-Studie gliedert sich ebenso ins Kompetenz-Modell ein. Gleiches gilt für die im Rahmen des HarmoS-Projektes (2007) veröffentlichten Items.

Weiterhin stellt sich die Frage, ob die zwei- oder drei-Faktoren-Lösung von Hammann et al. (2007) zur Adaption geeigneter erscheint.

Die Items zur ersten Hauptkomponente nehmen wie dargestellt eine Sonderstellung ein. Bereits Popper (1994, Erstausgabe 1934) schloss den Schritt der Hypothesengenerierung aus seiner Betrachtung der Logik der Forschung aus und ordnete ihn dem Bereich der Kreativität zu. Vor diesem Hintergrund und angesichts der ähnlichen Operationalisierung bei Hammann et al. (2007) scheint es gerechtfertigt, die erste mit der dritten Hauptkomponente zu den Operationen auf dem Hypothesen-Suchraum zusammenzufassen und den Operationen auf

dem Experiment-Suchraum gegenüberzustellen.

Bei der Einarbeitung des SDDS-Modells stellt sich darüber hinaus die Frage, ob es sinnvoll ist, wiederum eine neue Dimension in das Kompetenzmodell aufzunehmen, oder diesen Anforderungsbereich vollständig vom bisherigen Modellteil, der die Nutzung von Fachwissen behandelt, abzukoppeln. Die zweite Vorgehensweise scheint hier sinnvoll: Wie Hammann et al. (2007) zeigen, korreliert die Erkenntnisgewinnung nur schwach mit der Nutzung von Fachwissen in den jeweiligen Inhaltsbereichen. Dies rechtfertigt die Annahme, dass eine getrennte Niveaudefinition für die Nutzung von Fachwissen und die Erkenntnisgewinnung und dementsprechend auch eine konsequente Trennung der beiden Teilmodelle die realen Kompetenzausprägungen besser wiedergibt.

Die Fach-Stufen erscheinen zur Beschreibung von Items in diesem Bereich aus folgenden Gründen nicht zielführend: Auch wenn der Umgang mit Experimenten und Modellen in der Hochschule häufig stringenter und formalisierter behandelt wird, als dies in der Schule der Fall ist, können und sollen (vgl. KMK, 2005) die im SDDS-Modell abgebildeten kognitiven Fähigkeiten an beiden Stellen gleichermaßen erworben werden. Während die Fach-Stufen im Teilmodell zur Nutzung von Fachwissen den Bezug der jeweiligen Wissensfacette auf die jeweilige vermittelnde Institution und in Ansätzen auch die Nähe zum Unterrichtshandeln modellieren, steht in diesem Modellteil eher die (reflektierte) wissenschaftliche Arbeitsweise im Vordergrund.

Es ergibt sich also für das entwickelte Kompetenzmodell ein Teil-Modell zum Umgang mit Experimenten und Modellen, das als Dimensionen die bekannten Inhaltsbereiche und die hierarchische Komplexität enthält. Als dritte Dimension soll hier der zu bearbeitende Suchraum aus dem SDDS-Modell (Experiment-Raum, Hypothesen-Raum) verwendet werden, die etwas weniger abstrakt vielleicht als das Entwickeln von Experimenten bzw. die Bezugnahme auf Modelle umschrieben werden können. Die empirischen Befunde von Hammann et al. (2007) und die oben aufgeführte Argumentation rechtfertigen hier, die Analyse von Evidenzen und die Suche im Hypothesen-Raum zusammenzufassen.

Zur Operationalisierung der Kompetenzbereiche zum Umgang mit Experimenten und Modellen können hier die von Hammann et al. (2006, 2007), Kauertz et al. (2010) und vom IQB (2010) gezeigten Items als Beispiele verstanden werden. Dieses Vorgehen sichert auch die Anschlussfähigkeit an das ESNaS-Modell und damit an die Vorgaben der KMK für den Schulunterricht.

Der Bereich *Nature of Science* wird im Paderborner Modell bei den Beliefs verortet und mit Likert-Skalen abgefragt. Diese Skalen können problemlos für das vorliegende Projekt übernommen werden. Hiermit verbindet sich auch die Möglichkeit, eventuelle auffällige Befunde in

diesem Bereich mit zur Analyse und Charakterisierung der Kompetenzniveaus heranziehen zu können.

4 Modell-Übersicht

Das bisher erarbeitete Strukturmodell der fachlichen Kompetenz wird im Folgenden noch einmal zusammenfassend dargestellt. Es ist eingebettet in das Rahmenmodell professioneller Handlungskompetenz von Lehrkräften (Abb. 3; vgl. Riese, 2009). Es teilt sich in zwei Teilmodelle mit je drei Dimensionen (vgl. Abb. 4). Die beiden Di-

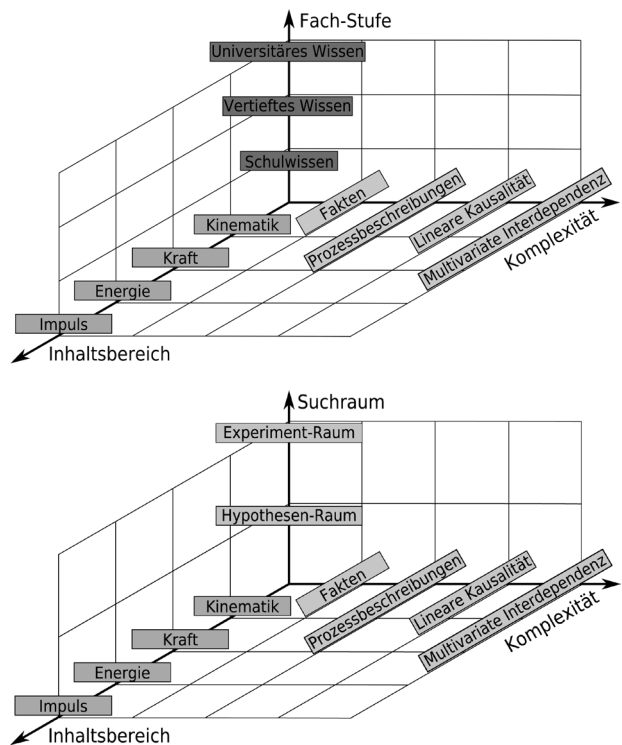


Abb. 4: Strukturmodell der fachlichen Kompetenz zur Nutzung von Fachwissen (oben) und zum Umgang mit Experimenten und Modellen (unten).

mensionen *Komplexität* und *Inhaltsbereich* sind für beide Teilmodelle identisch. Dazu kommt im Teilmodell zur Nutzung von Fachwissen die *Fach-Stufe* und im Teilmodell zum Umgang mit Experimenten und Modellen der *Suchraum* im Rahmen des SDDS-Modells nach Klahr (2000).

4.1 Schwierigkeitserzeugende Merkmale

Als Grundlage für die Definition von Niveaustufen muss hypothetisch formuliert werden, welche Merkmale einer Anforderung schwierigkeiterzeugend sind.

Als dominante Dimension wird hierbei die hierarchische Komplexität angenommen. Dabei sollten *Fakten*-Aufgaben die einfachsten und Aufgaben zu *multivariaten Interdependenzen* die schwierigsten sein. Als Sekundäre Faktoren kann für die Nutzung von Fachwissen nach den Erkenntnissen von Riese (2009) angenommen werden, dass Anforderungen auf der Stufe des *vertieften Wissens* schwieriger sind als auf den anderen beiden Stufen. Ähnlich kann mit den Erkenntnissen von Hammann, Phan & Bayrhuber (2007) angenommen werden, dass Anforderungen zur Nutzung von Experimenten, die auf dem *Experiment-Raum* operieren etwas schwieriger ausfallen als solche, die auf dem *Hypothesen-Raum* operieren.

In wie weit diese sekundären schwierigkeiterzeugenden Merkmale gegenüber der als dominant angenommenen Aufgabenkomplexität eine Einstufung in höhere Niveaus rechtfertigen, kann an dieser Stelle noch nicht beantwortet werden.

Konkret wären hier zwei Möglichkeiten denkbar:

Das sekundäre Merkmal könnte sich so stark auf die Schwierigkeit auswirken, dass Items aus diesem Bereich grundsätzlich in einem höheren Niveau eingestuft werden als andere Items derselben Komplexität. Andererseits wäre es auch denkbar, dass dieser Effekt nur bei kleinen Komplexitäten auftritt und Items höherer Komplexität auf Niveaus unabhängig ihrer Fach-Stufe verortet werden.

4.2 Erste Hinweise zur Validität der formulierten schwierigkeiterzeugenden Merkmale

Als erster Schritt zur Validierung und Plausibilitätsprüfung dieses Modells und der darin formulierten schwierigkeiterzeugenden Merkmale werden im Folgenden die Testitems von Riese (2009) einer Reanalyse unterzogen in das präzisierte bzw. neu entwickelte Modell eingeordnet. Da diese Items jedoch auf der Grundlage eines Modells mit weniger Matrixzellen erstellt wurden, sind nicht alle Anforderungsbereiche vollständig abgedeckt. Somit kann dieses Vorgehen nur erste Hinweise auf die Anwendbarkeit des Modells liefern, sie ersetzen keine vollständige Überprüfung mit Hilfe eines neu zu erstellenden Tests.

Wie erwartet steigt die Aufgabenschwierigkeit im Wesentlichen mit der hierarchischen Komplexität, wie Tabelle 2 zeigt. Dies unterstreicht die prinzipielle Eignung des Modells von Bernholt für unsere Zwecke. In der Tabelle ist der jeweilige mittlere Itemparameter eingetragen,

Tab. 2: Explorative Validierung mit Hilfe der Testitems von Riese (angegeben ist jeweils der Itemparameter, die Standardabweichung und in Klammern die Itemanzahl; dabei zeigt ein größerer Itemparameter eine höhere Schwierigkeit an)

	Fakten	Prozess- beschrei- bungen	Lineare Kausalität	Multivariate Inter- dependenz	Gesamt
Schulwissen	(0 Items)	.14±.12 (2 Items)	1.31 (1 Item)	1.68±.37 (2 Items)	.88±.96 (5 Items)
Vertieftes Wissen	.15±.51 (5 Items)	.53±.21 (4 Items)	.09 (1 Item)	1.79±.66 (4 Items)	.72±.84 (14 Items)
Universitäres Wissen	.05±.68 (6 Items)	(0 Items)	(0 Items)	2.00±.38 (3 Items)	.70±1.13 (9 Items)
Gesamt	.09±.58 (11 Items)	.31±.38 (6 Items)	.70±.86 (2 Items)	1.83±.48 (9 Items)	.74±.93 (28 Items)

welchen man aus einer Rasch-Analyse der empirischen Daten als Schätzwert für die Itemschwierigkeit erhält. Ein größerer Wert steht hier für eine höhere Aufgabenschwierigkeit.

Der Einfluss der Fach-Stufe auf die Itemschwierigkeit lässt sich anhand der hier vorliegenden Daten nur sehr schwer abschätzen. Auf den Stufen *Fakten* und *Prozessbeschreibungen* zeigt sich (wie erwartet) die erhöhte Schwierigkeit des *vertieften Wissens*. Für die *lineare Kausalität* ist aufgrund der geringen Itemzahl keine (belastbare) Aussage möglich und auf der höchsten Komplexitätsstufe zeigt sich eher eine Tendenz zu einer höheren Schwierigkeit des *Universitären Wissens*.

Es fällt auf, dass dieser Befund nicht mit der eventuell vermuteten Reihung *Schul-*, *vertieftes* und *universitäres Wissen* übereinstimmt (vgl. auch Riese, 2009). Für eine belastbare Erklärung sind allerdings weitere empirische Untersuchungen auf der Basis einer präziseren Beschreibung der Stufen nötig.

Eine Analyse der Inhaltsbereiche zeigt weiterhin, dass Items zur *Kraft* schwieriger ausfallen als zu den anderen Inhalten (vgl. auch Riese, 2009). Dies kann allerdings darauf zurückgeführt werden, dass die Kraft-Items spezifisch Fehlvorstellungen abfragen, was im hier entwickelten Modell als Merkmal des *vertieften Wissens* angesehen wird.

Insgesamt kann das Teil-Modell zur Nutzung von Fachwissen die Items von Riese also korrekt nach ihrer Schwierigkeit anordnen. Es zeigt sich aber auch, dass für eine komplette Modellabdeckung deutlich mehr Items benötigt werden. Auf dieser Grundlage könnte dann auch der Einfluss der sekundären schwierigkeiterzeugenden Merkmale genauer geprüft werden.

Da bei Riese keine Items zur Erkenntnisgewinnung verwendet wurden, muss an dieser Stelle auf eine analoge Analyse des Teilmodells zum Umgang mit Experimenten und Modellen verzichtet werden. Die umfassende Überprüfung des Modells muss also mittels eines noch zu er-

stellenden Tests erfolgen, der das gesamte Modell abdeckt.

5 Fazit

Der Mangel an adaptiven Fördermaßnahmen im Bereich der physikalischen Fachkompetenz hängt eng mit einem Mangel an präzisen und differenzierten Kompetenztests auf diesem Gebiet zusammen. Die Entwicklung entsprechender Tests wiederum erfordert differenzierte und valide Kompetenzstrukturmodelle im Hintergrund.

Vor diesem Hintergrund wurde ausgehend vom Strukturmodell des fachbezogenen Wissens (Riese, 2010) ein Modell des Fachwissens angehender Physik Lehrkräfte entwickelt, das in einem Teilmodell die Nutzung von Fachwissen (*Schulwissen*, *vertieftes* und *universitäres Wissen*) und in einem zweiten Teilmodell den Umgang mit Experimenten und Modellen abdeckt. Das Modell beschränkt sich aus Gründen der Testpraktikabilität auf Inhaltsbereiche der Mechanik (*Kinematik*, *Kraft*, *Energie*, *Impuls*) als Prädiktor für allgemeines physikalisches Fachwissen. Es nutzt das empirisch bereits bestätigte Modell hierarchischer Komplexität nach Bernholt (2010) zur Einstufung von Aufgaben (*Fakten*, *Prozessbeschreibungen*, *lineare Kausalität*, *multivariate Interdependenz*), welches gegenüber dem verbreiteten Komplexitätsmodell von Kauertz der Problemstellung dieser Arbeit besser angepasst erscheint. Das Teilmodell zum Umgang mit Experimenten und Modellen adaptiert für den Bereich des Scientific Inquiry das SDDS-

Modell von Hammann (2007) bzw. Klahr (2000) und nutzt so eine Modellierung des naturwissenschaftlichen Erkenntnisprozesses als Problemlösen, welche anschlussfähig an andere Projekte operationalisiert werden kann. Charakterisierendes Aufgabenmerkmal ist hier, ob die Aufgaben auf dem Experiment- oder Hypothesen-Suchraum operiert. Darüber hinaus werden Skalen zur Nature of Science von Riese (2009) übernommen.

Im Hinblick auf die spätere Bildung von Kompetenzniveaus wird hypothetisch die Aufgabenkomplexität als primäres und die Fach-Stufe bzw. der Suchraum als sekundäres schwierigkeiterzeugendes Aufgabenmerkmal im Modell angenommen. Erste Hinweise zur Validität des Modells für die Nutzung von Fachwissen liefert eine Analyse anhand der Items von Riese. Sie bestätigt insgesamt die Tauglichkeit des Modells für die Einordnung von Items nach ihrer Schwierigkeit, lässt aber mangels Items die Frage nach dem Einfluss der sekundären Merkmale offen.

Auf der Grundlage dieses Modells ist im nächsten Schritt ein umfassender Test zu entwickeln, der einerseits zur Validierung des Modells und andererseits zur Definition und Analyse von Kompetenzniveaus der physikalischen Fachkompetenz genutzt werden kann. Aufbauend auf den so gewonnenen Daten können dann die Niveaustufen endgültig im Modell verortet und die dazugehörigen Itemgruppen charakterisiert werden. Auf der empirischen Einstufung von Probanden in Niveaus, die sich durch ähnliche Fähigkeitsprofile auszeichnen, können dann Fördermaßnahmen aufbauen, die adaptiv an den je-

weiligen Niveaus ansetzen und darauf abgestimmt sind.

Literatur

- Abell, S. K. (2007). Research on Science Teacher Knowledge. In Abell, S. K., & Lederman, N. G. (Eds.): Handbook of research on science education. Mahwah, NJ: Lawrence Erlbaum, S. 1105–1149.
- Anderson, C.W. (1979). An observational study of classroom management and information structuring in elementary school science lessons. Dissertation Abstracts International, 40 (11), 5810A (UMI no. 8009823).
- Baumert, J. & Kunter, M. (2006). Stichwort: Professionelle Kompetenz von Lehrkräften. Zeitschrift für Erziehungswissenschaft, 9 (4), S. 469–520.
- Baumert, J., Kunter, M., Blum, W., Brunner, M., Voss, T., Jordan, A., Klusmann, U., Krauss, S., Neubrand, M., & Tsai, Y. M. (2010). Teachers' mathematical knowledge, cognitive activation in the classroom, and student progress. American Educational Research Journal, pp 133–180.
- Bernholt, S. (2010). Kompetenzmodellierung in der Chemie – Theoretische und empirische Reflexion am Beispiel des Modells hierarchischer Komplexität. Dissertation. Berlin: Logos Verlag.
- Blömeke, S., Kaiser, G. & Lehmann, R. (Hrsg.) (2008). Professionelle Kompetenz angehender Lehrerinnen und Lehrer – Wissen, Überzeugungen und Lerngelegenheiten deutscher Mathematikstudierender und –referendare – Erste Ergebnisse zur Wirksamkeit der Lehrerausbildung. Münster: Waxmann Verlag.
- Blömeke, S., Kaiser, G. & Lehmann, R. (Hrsg.) (2010). TEDS-M 2008 – Professionelle Kompetenz und Lerngelegenheiten angehender Primarstufenlehrkräfte im internationalen Vergleich. Münster: Waxmann.
- Borowski, A.; Neuhaus, B. J.; Tepner, O.; Wirth, J.; Fischer, H. E.; Leutner, D. et al. (2010). Professionswissen von Lehrkräften in den Naturwissenschaften (ProwiN) – Kurzdarstellung des BMBF-Projekts. Zeitschrift für Didaktik der Naturwissenschaften, 16.
- Bromme, R. (1997). Kompetenzen, Funktionen und unterrichtliches Handeln des Lehrers. In F.E. Weinert (Hrsg.), Enzyklopädie der Psychologie, Bd.3: Psychologie des Unterrichts und der Schule. Göttingen: Hogrefe, S. 177–212.
- Commons, M.L., Trudeau, E.J., Stein, S.A., Richards, F.A. & Krause, S.R. (1998). Hierarchical Complexity of Tasks Shows the Existence of Developmental Stages. In: Developmental Review, 18, S. 237–278.
- De Jong, O., Korthagen, F. & Wubbels, Th. (1998). Research on science teacher education in Europe: teacher thinking and conceptual change. In: B. J. Frazer and K. G. Tobin (Eds.), International Handbook of Science Education, (pp. 745–758). Dordrecht: Kluwer Academic Publishers.
- Friege, G. & Lind, G. (2004). Leistungsmessung im Leistungskurs. In: Der Mathematische und Naturwissenschaftliche Unterricht – MNU, 57(5), S. 259–265.
- Girwidz, R., Kurz, G. & Kautz, Chr. (2003). Zum Verständnis der newtonschen Mechanik bei Studienanfängern – der Test 'Force Concept Inventory – FCI'. In Nordmeier, V. (Hrsg.) Beiträge zur Frühjahrstagung des Fachverbandes Didaktik der Physik der Deutschen Physikalischen Gesellschaft. Tagungs-CD: Didaktik der Physik – Augsburg 2003. Berlin: Lehmanns Media.
- Hammann, M. (2007). Das Scientific Discovery as Dual Search-Modell. In: Dirk Krüger/ Helmut Vogt: Theorien in der biologiedidaktischen Forschung. Springer Verlag, Berlin, 187–196.
- Hammann, M., Phan, T.H. & Bayrhuber, H. (2007). Experimentieren als Problemlösen: Lässt sich das SDDS-Modell nutzen, um unterschiedliche Dimensionen beim Experimentieren zu messen? In Zeitschrift für Erziehungswissenschaft, 10, Sonderheft 8, S. 33–49
- Hammann, M., Phan, T.H., Ehmer, M. & Bayrhuber, H. (2006). Fehlerfrei Experimentieren. In: Der Mathematische und Naturwissenschaftliche Unterricht – MNU, 59(5), S. 292–299.
- HarmoS-Konsortium Naturwissenschaften+ und EDK (2007). Aufgaben, Situationen. Zugriff am 30.11.2010 von <http://harmos.phbern.ch/index.php?id=113>

- Hartig, J. (2007): Skalierung und Definition von Kompetenzniveaus. In: Beck, B. & Klieme, E. (Hrsg.): Sprachliche Kompetenzen. Konzepte und Messung. Weinheim, S. 83–99.
- IQB (2010) Beispielaufgaben für den Kompetenzbereich Erkenntnisgewinnung. Zugriff am 23.11.2010 von http://www.iqb.hu-berlin.de/arbbereiche/projekte?pg=p_34&spg=r_7
- van Joolingen, W. & de Jong, T. (1997). An extended dual search space model of scientific discovery learning. *Instructional Science* 25, S. 307–346.
- Justi, R. S. & Gilbert, J. K. (2002). Modelling, teachers' views on the nature of modelling, and implications for the education of modellers, *International Journal of Science Education*, 24 (4), 369–387.
- Kalyuga, S., Ayres, P., Chandler, P. & Sweller, J. (2003). The expertise reversal effect. *Educational Psychologist* 38. S. 23–32.
- Kauertz, A. (2008). Schwierigkeitserzeugende Merkmale physikalischer Leistungstestaufgaben. In H. Niedderer, H. Fischler & E. Sumfleth (Hrsg.): *Studien zum Physik- und Chemielernen*, Band 79. Berlin: Logos.
- Kauertz, A., Fischer, H. E., Mayer, J., Sumfleth, E. & Walpuski, M. (2010). Standardbezogene Kompetenzmodellierung in den Naturwissenschaften der Sekundarstufe I. *Zeitschrift für Didaktik der Naturwissenschaften*, 16, 135–153.
- Kirsch, A. (1987). *Mathematik wirklich verstehen*. Köln: Aulis Verlag Deubner.
- Klahr, D. (2000) *Exploring Science: The Cognition and Development of Discovery Processes*. Cambridge, MA: MIT Press.
- Klahr, D. & Dunbar, K. (1988). Dual Space Search During Scientific Reasoning. *Cognitive Science* 12, S. 1–48
- Klein, F. (1933). *Elementarmathematik vom höheren Standpunkte aus*. 4. Auflage, Nachdruck 1968, Berlin: Springer.
- Klieme, E.; Avenarius, H.; Blum, W.; Döbrich, P.; Gruber, H.; Prenzel, M.; Reiss, K.; Riquarts, K.; Rost, J.; Tenorth, H.-E. & Vollmer, H. J. (2003). *Zur Entwicklung nationaler Bildungsstandards. Eine Expertise*. Berlin.
- Klieme, E. & Leutner, D. (2006). Kompetenzmodelle zur Erfassung individueller Lernergebnisse und zur Bilanzierung von Bildungsprozessen. Beschreibung eines neu eingerichteten Schwerpunktprogramms der DFG. In *Zeitschrift für Pädagogik*, 52 (6), S. 876–903.
- Klieme, E., Maag-Merki, K. & Hartig, J. (2007). Kompetenzbegriff und Bedeutung von Kompetenzen im Bildungswesen. In J. Hartig & E. Klieme (Hrsg.), *Möglichkeiten und Voraussetzungen technologiebasierter Kompetenzdiagnostik* (S. 5–15). Berlin: BMBF.
- KMK, Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland (Hrsg.) (2005). *Bildungsstandards im Fach Physik (Chemie / Biologie) für den Mittleren Schulabschluss*. München: Luchterhand.
- Krauss, S., Neubrand, M., Blum, W., Baumert, J., Brunner, M., Kunter, M. et al. (2008). Die Untersuchung des professionellen Wissens deutscher Mathematik-Lehrerinnen und -Lehrer im Rahmen der COACTIV-Studie. *Journal für Mathematikdidaktik*, 29 (3/4), S. 223–258.
- Labudde, P., Duit, R., Fickermann, D., Fischer, H. E., Harms, U., Mikelskis, H. et al. (2009). Schwerpunkttagung „Kompetenzmodelle und Bildungsstandards: Aufgaben für die naturwissenschaftsdidaktische Forschung“ *Zeitschrift für Didaktik der Naturwissenschaften*, 15, 343–370
- Mayer, J. (2007). Erkenntnisgewinnung als wissenschaftliches Problemlösen. In: Krüger, D. & Vogt, H.: *Theorien in der biologiedidaktischen Forschung*. Berlin/Heidelberg: Springer, S. 177–186.
- Metzger, S. & Labudde, P. (2007). HarmoS Naturwissenschaften+. Bildungsstandards für die Schweiz. In *Praxis der Naturwissenschaften – Physik in der Schule*, 56 (6), S. 14–18.
- Paas, F. G. W. C., Renkl, A. & Sweller, J. (2004). Cognitive load theory: Instructional implications of the interaction between information structures and cognitive architecture. *Instructional Science*, 32, S. 1–8.
- Popper, K. (1994). *Logik der Forschung*. Mohr: Tübingen. 10. Auflage

- Riese, J. (2009). Professionelles Wissen und professionelle Handlungskompetenz von (angehenden) Physiklehrkräften. Dissertation. Berlin: Logos Verlag.
- Riese, J. (2010). Empirische Erkenntnisse zur Wirksamkeit der universitären Lehrerbildung - Indizien für notwendige Veränderungen der fachlichen Ausbildung von Physiklehrkräften. In *PhyDid – Physik und Didaktik in Schule und Hochschule*, 9, S.25–33.
- Schaper, N. (2009). Aufgabenfelder und Perspektiven bei der Kompetenzmodellierung und -messung in der Lehrerbildung. Themenheft der Zeitschrift *Lehrerbildung auf dem Prüfstand*, 2 (1), S. 166-199.
- Schecker, H. & Parchmann, I. (2006). Modellierung naturwissenschaftlicher Kompetenz. In *Zeitschrift für Didaktik der Naturwissenschaften*, 12, S. 45–66.
- Schreiber, N., Theyßen, H. & Schecker, H. (2009) Experimentelle Kompetenz messen?! In: *Physik und Didaktik in Schule und Hochschule* 3(8), S. 92–101
- Shulman, L. (1986). Those who Understand: Knowledge Growth in Teaching. In *Educational Researcher*, 15, S. 4–14.
- Terhart, E. (2002). Standards für die Lehrerbildung. Eine Expertise für die Kultusministerkonferenz (ZKL-Texte Nr. 23). Universität Münster: Zentrale Koordination Lehrerbildung.
- Walpuski, M., Kampa, N., Kauertz, A. & Wellnitz, N. (2008). Evaluation der Bildungsstandards in den Naturwissenschaften. In: *Der Mathematische und Naturwissenschaftliche Unterricht – MNU*, 61(6), S. 323–326.
- Walpuski, M., Kauertz, A., Kampa, N., Fischer, H. E., Mayer, J., Sumfleth, E., Wellnitz, N.: *ESNaS – Evaluation der Standards für die Naturwissenschaften in der Sekundarstufe I*. In: *Bildungsstandards und Kompetenzmodelle – Beiträge zu einer aktuellen Diskussion über Schule, Lehrerbildung und Unterricht*. A. Gehrman, U. Hericks & M. Lüders (Hrsg.); Bad Heilbrunn: Verlag Julius Klinkhardt. S. 163–176.
- Weinert, F. E. (2001). Concept of Competence: A Conceptual Clarification. In D.S. Rychen & L.H. Salganik (Hrsg.), *Defining and Selecting Key Competencies*. Göttingen: Hogrefe, S. 45–66.

KONTAKT

David Woitkowski
 Universität Paderborn
 Department Physik
 Warburger Str. 100
 33098 Paderborn
david.woitkowski@upb.de

AUTORENINFORMATION

David Woitkowski
 ist Doktorand in der Arbeitsgruppe Didaktik der Physik an der Universität Paderborn. Er beschäftigt sich mit der Entwicklung und Validierung eines Modells und dazugehörigen Testinstruments zur Definition und Charakterisierung von Niveaustufen im Bereich des physikalischen Fachwissens.

Dr. Josef Riese

ist zurzeit wissenschaftlicher Mitarbeiter in der Arbeitsgruppe Didaktik der Physik an der Universität Paderborn mit den Forschungsschwerpunkten Kompetenzmodellierung, Testentwicklung sowie Struktur und Entwicklung professioneller Handlungskompetenz von Physiklehrkräften.

Prof. Dr. Peter Reinhold

ist Professor für Didaktik der Physik an der Universität Paderborn. Seine Forschungsschwerpunkte sind Lehrerbildungsforschung, Physik Lernen und Lehren mit digitalen Medien, kontextorientierter und Interesse fördernder Physikuterricht und kooperatives Lernen.

