

Messung des Fachwissens von Physikstudenten - Erste Ergebnisse einer Pilotstudie -

David Woitkowski*, Josef Riese*, Peter Reinhold*

*Universität Paderborn, AG Didaktik der Physik, Warburger Str. 100, 33098 Paderborn
david.woitkowski@upb.de

Kurzfassung

Physikalisches Fachwissen spielt eine grundlegende Rolle in der Ausbildung von Physiklehrkräften, Fach-Physikern und Studierenden weiterer Fächer. Gleichzeitig zeigen frühere Studien ernsthafte Probleme bei der universitären Wissensvermittlung in diesem Bereich. Um angestrebte Maßnahmen zur Verbesserung dieser Situation evaluieren zu können, soll im vorgestellten DFG-Projekt ein präzises und differenziertes Testinstrument auf Hochschulebene erarbeitet werden, welches die Nutzung von Fachwissen und die Nutzung von Experimenten und Modellen zur Erkenntnisgewinnung abdeckt. Das dem Instrument zugrunde gelegte Strukturmodell fachlichen Wissens wird im Beitrag überblicksartig vorgestellt. Aus den Daten einer Pilotierungsstudie des erarbeiteten Tests können erste Hinweise auf Bedingungen für den Erwerb fachlichen Wissens im Physikstudium, Unterschiede und Gemeinsamkeiten zwischen verschiedenen Studiengängen und die weitere Verbesserung des Instruments gezogen werden.

1. Einleitung

Im Bereich der universitären Lehrerbildung spielt Fachwissen neben fachdidaktischem und pädagogischem Wissen als eine von drei Wissensfacetten eine wesentliche Rolle (vgl. [1]) und bildet so eine zentrale Komponente der professionellen Kompetenz von Fachlehrkräften (vgl. [2] und [3]). Kognitiv anregender Unterricht baut wesentlich auf ein fundiertes und vernetztes Fachwissen der Lehrkräfte auf. Dabei reicht es nicht aus, wenn die Lehrpersonen den Lernenden fachlich „voraus“ sind, viel mehr benötigen sie ein tieferes Verständnis der fachlichen Inhalte des Schulcurriculums [4], um flexibel mit auftretenden Unterrichtssituationen wie Schülerfragen und Schwierigkeiten im Vorverständnis umgehen zu können. Es ist davon auszugehen, dass das Repräsentations- und Erklärungsrepertoire von Lehrkräften wesentlich von der Breite und Tiefe ihres Fachwissens abhängt [2], darüber hinaus kann Fachwissen, vermittelt über das fachdidaktische Wissen, als bedeutsamer Prädiktor für eine kognitiv herausfordernde und konstruktiv unterstützende Unterrichtsgestaltung gelten [5]. Fehlt Fachwissen, so schlägt sich dies direkt auf die Schülerleistungen nieder (vgl. z.B. [6]). Weiterhin hat es sich als notwendige Voraussetzung zum Erwerb von fachdidaktischem Wissen erwiesen (vgl. z.B. [4] und [7]).

Vor diesem Hintergrund erscheint es problematisch, dass die Vermittlung von Fachwissen an deutschen Hochschulen (aber auch international, vgl. [8]) wesentliche Schwächen aufweist: Die heterogenen Lernvoraussetzungen, die die Studierenden aus der Schule an die Hochschule mitbringen, können dort nicht ausgeglichen werden [9]. Obwohl man erwarten könnte, dass gerade schwächere Studierende früh aus dem Studium ausscheiden, zeigt sich mit zunehmender Studiendauer eine größer werdende Spannweite der Leistungsfähigkeit (vgl. [9], [10]). Mehrere Studien stellen bei Lehrpersonen dieselben fachlich-konzeptuellen Fehlvorstellungen fest, wie sie von Schülern bekannt sind (vgl. [8], [11] und [12]).

Um diese Probleme mittelfristig angehen zu können, ist es notwendig, eine adaptive Förderintervention zu installieren, die an den spezifischen Schwächen der Studenten ausgerichtet ist. Diese Adaption kann aber nur geleistet werden, wenn zuvor eine präzise und differenzierte Bestimmung der Fähigkeiten und Schwächen eines einzelnen Studenten durchgeführt werden kann. Ein Diagnoseinstrument, das dies leisten kann, wird im vorliegenden DFG-Projekt entwickelt. Es basiert auf einem Modell professioneller Kompetenz, das im Folgenden vorgestellt wird. Anschließend wird das Testinstrument kurz vorgestellt und erste Ergebnisse aus der Pilotstudie werden berichtet.

2. Zu Grunde gelegtes Modell

Das der Studie zu Grunde liegende Modell professioneller Kompetenz von Physiklehrkräften ist in [13] ausführlich dokumentiert. Es handelt sich dabei um ein normatives Strukturmodell, das fachliches Wissen zunächst neben physikdidaktischem und erziehungswissenschaftlichem Wissen als Teil professioneller Kompetenz von Physiklehrkräften auffasst. Dieses Strukturmodell dient einerseits als Basis für die Erstellung eines differenzierten Tests und trifft andererseits hypothetische Aussagen über Schwierigkeitserzeugende Merkmale einzelner Anforderungen, die sich in den Testaufgaben wiederfinden.

Obwohl die genutzten Ansätze ihren Ursprung in der Lehrerbildungsforschung haben, kann zunächst angenommen werden, dass das Kompetenzkonstrukt des fachlichen physikalischen Wissens auch bei Fach- und Nebenfach-Studenten der Physik anwendbar ist.

2.1 Rahmenmodell

Das von Riese [9] auf der Grundlage von [1], [2], [3], [14] und [15] erstellte Modell professioneller Kompetenz von Lehrpersonen (Abb. 1) gliedert sich im Sinne Weinerts [14] einerseits in die Bereiche der kognitiven Fähigkeiten und des Professionswissens und der motivationalen, volitionalen und sozialen Bereitschaften und Fähigkeiten andererseits. Während letzterer Bereich hier nur am Rande von Interesse ist, lässt sich ersterer wiederum (analog zu [1]) in die drei Wissensarten *fachliches Wissen*, *fachdidaktisches Wissen* und *erziehungswissenschaftliches Wissen* aufteilen. Das *physikalische Fachwissen* steht im Weiteren im Fokus.

Mehrere aktuelle Ansätze unterteilen nun das fachliche Wissen, häufig in Anlehnung an die Bildungsstandards [16], wiederum in die Bereiche Fachwissen (im engeren Sinne), Erkenntnisgewinnung, Kommunikation und Bewertung. Während *Kommunikation* und *Bewertung* sicherlich tätigkeitspezifisch (d.h. für Lehrpersonen und Fachphysiker unterschiedlich strukturiert und ausgeprägt) gedacht werden muss, können die beiden Teilbereiche der *Nutzung von Fachwissen* und der *Nutzung von Experimenten und Modellen zur Erkenntnisgewinnung*

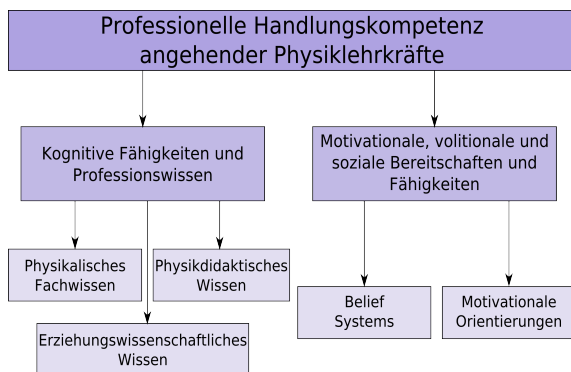


Abb. 1: Rahmenmodell professioneller Handlungskompetenz von Lehrpersonen nach [9] und [13].

weitgehend tätigkeitsübergreifend formuliert und strukturiert werden. Diese beiden Bereiche finden sich in ähnlicher Form auch bei z.B. [17], [18] und [19] wieder.

Für beide Teilbereiche wurde ein eigenes Strukturmodell erstellt, um später empirische Aussagen über deren innere Struktur und Zusammenhänge treffen zu können.

2.2 Nutzung von Fachwissen

Im Rahmen dieses Teilmodells werden Anforderungen beschrieben, bei denen physikalisches Fachwissen genutzt werden muss, um gegebene innerphysikalische oder alltägliche Probleme zu lösen. Jede Anforderung wird dabei anhand von drei Merkmalen beschrieben (vgl. Abb. 2).

2.2.1 Inhaltsbereich

Wie auch schon in den anderen Paderborner Arbeiten beschränkt sich das vorliegende Projekt auf den Inhaltsbereich Mechanik. Er wird im Physik-Studium in der Regel bereits im ersten Semester thematisiert und kann folglich schon früh im Studium getestet werden. Grundlegende Konzepte der Physik wie Kraft, Impuls, Energie u.a. werden in der Mechanik eingeführt, so dass später darauf zurückgegriffen werden kann. Damit hängt zusammen, dass die Mechanik empirisch ein guter Prädiktor für die Kompetenz in anderen Inhaltsbereichen ist (vgl. z.B. [20]). Bei Riese [9] konnte ein Kompetenzzuwachs über die gesamte Studiendauer festgestellt werden, obwohl nur typische Inhalte des ersten Studienseesters getestet wurden. Wie dort konzentriert sich dieses Projekt auf die vier Inhaltsbereiche Kinematik, Kraft, Energie und Impuls.

2.2.2 Komplexität

In mehreren aktuellen Arbeiten (vgl. z.B. [21] und [22]) konnte gezeigt werden, dass die Komplexität als strukturelles Merkmal einer Aufgabe einen wesentlichen Einfluss auf ihre empirische Schwierigkeit hat. Nach einer Abwägung (vgl. [13]) zwischen verschiedenen Komplexitäts-Modellen hat sich das Modell von Bernholt [22] als für dieses Projekt angemessen gezeigt. Es enthält die vier Stufen *Fakten*, *Prozessbeschreibung*, *lineare Kausalität* und *multivariate Interdependenz*. Die unterste Stufe *unreflek-*

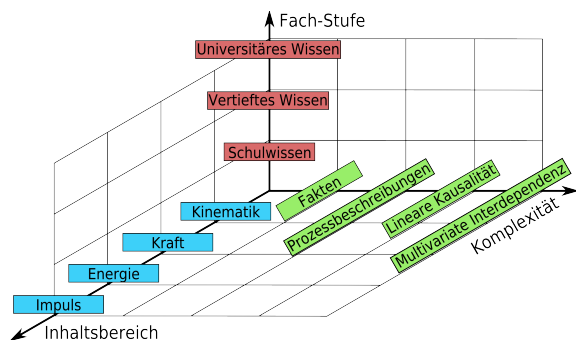


Abb. 2: Strukturmodell der fachlichen Kompetenz, Teilmodell zur Nutzung von Fachwissen.

tiertes Erfahrungswissen wird hier nicht weiter verwendet, da hier die professionelle Kompetenz von Interesse ist, bei der diese Stufe keine Rolle mehr spielen sollte.

2.2.3 Fach-Stufe

Analog zu der in verschiedenen Studien aus dem Bereich der Mathematik (z.B. [4], [15]) und Physik (z.B. [9]) auftretenden Staffelung nach schulischem und universitärem Wissen und teilweise weiteren Unterteilungen in Sekundarstufen-I und -II-Wissen o.ä. teilt auch das beschriebene Modell zunächst grob in *Schulwissen* (Maßstab sind hier die Bildungsstandards für den mittleren Schulabschluss, [16]) einerseits und völlig von der Schule losgelöstes, *universitäres Wissen* andererseits ein. Auf eine Stufe wie vertieftes Schulwissen, Sekundarstufe-II-Wissen o.ä. wurde zu Gunsten einer deutlichen Abgrenzbarkeit verzichtet.

Darüber hinaus sieht das Modell eine dritte Fach-Stufe, genannt *vertieftes Wissen*, vor. Aufbauend auf Arbeiten wie [23] und [24] baut dieses auf dem schulischen und universitären Wissen gleichermaßen auf, vertieft, reflektiert und verknüpft es. Auf dieser Stufe sind bekannte fachliche Fehlvorstellungen überwunden und Begriffe und Konzepte gründlich verstanden. Es stellt somit das vernetzte Verständnis dar, das als Basis für das Erklärungsrepertoire einer Lehrkraft (vgl. [2]) angesehen werden kann.

Aufgrund von Daten aus [9] können Aufgaben zum *vertieften Wissen* hypothetisch als etwas schwieriger als andere Aufgaben angenommen werden.

2.3 Nutzung von Experimenten und Modellen

Das Rahmenmodell wissenschaftsmethodischer Kompetenzen [25], auf das hier Bezug genommen wird, unterscheidet die drei Kompetenzkonstrukte *Wissenschaftsverständnis*, *Wissenschaftliches Denken* und *Manuelle Fertigkeiten*. Dabei kann das Wissenschaftsverständnis im Bereich der *Belief Systems* (vgl. Abb. 1) verortet werden, ist also keine kognitive Fähigkeit im Sinne des Rahmenmodells. Die *manuellen Fertigkeiten* können in einem Paper-Pencil-Test nicht ohne weiteres valide getestet werden, so dass sie hier nicht weiter betrachtet werden. Somit

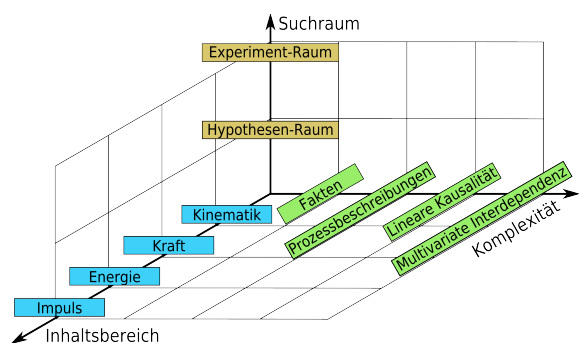


Abb. 3: Strukturmodell der fachlichen Kompetenz, Teilmodell zur Nutzung von Experimenten und Modellen zur Erkenntnisgewinnung.

steht das *wissenschaftliche Denken* im Fokus der folgenden Betrachtungen.

Wie in [13] dargestellt, wird im Rahmen des vorliegenden Projektes das Modell des Scientific Discovery as Dual Search (im Folgenden „SDDS-Modell“, vgl. [25] und [26]) genutzt, um Anforderungen im Bereich des *wissenschaftlichen Denkens* zu beschreiben. Dabei werden Anforderungen unterschieden, bei denen auf dem *Experiment-* bzw. auf dem *Hypothesen-Suchraum* operiert werden muss. Im ersten Fall wird dem Probanden in einer Aufgabe eine Hypothese oder eine wissenschaftliche Fragestellung und entsprechendes physikalisches Hintergrundwissen präsentiert. Auf dieser Grundlage muss ein Experiment geplant, ausgewählt oder zusammengestellt werden, welches geeignet ist, die Hypothese zu prüfen bzw. die Frage zu beantworten. Umgekehrt liegen bei einer Anforderung auf dem *Hypothesen-Raum* das experimentelle Setting und die Beobachtungen, Messwerte oder sonstige Resultate vor, die in Hinblick auf die ebenfalls gegebene Hypothese bzw. Frage ausgewertet werden müssen. Diese beiden Typen von Anforderungen entsprechen dabei der zweiten und dritten Hauptkomponente im SDDS-Modell. Auf eine weitere Auffächerung der Anforderungen oder die Testung der Fähigkeit zur Hypothesenbildung oder Formulierung wissenschaftlicher Fragen wird, wie in [13] erläutert, verzichtet.

Wie in Abb. 3 dargestellt, wird die beschriebene Unterscheidung wiederum in ein dreidimensionales Strukturmodell eingebettet. Neben dem Suchraum wird jede Anforderung durch den Inhaltsbereich und ihre Komplexität beschrieben. Diese beiden Skalen sind mit den entsprechenden Merkmalen im Teilmodell zur Nutzung von Fachwissen identisch.

Hypothetisch wird hier wieder die *Komplexität* als primäres schwierigkeiterzeugendes Merkmal angenommen. Außerdem kann angenommen werden, dass Anforderungen auf dem *Experiment-Raum* etwas schwieriger sind (vgl. [27]).

3. Eingesetztes Testinstrument

Aufbauend auf dem dargestellten Strukturmodell fachlicher Kompetenz wurden etwa 130 Test-Items mit geschlossenem und offenem Antwortformat entwickelt und im Modell verortet. Es ergibt sich so eine fast vollständige Abdeckung aller Matrixzellen (d.h. aller im Rahmen des Modells möglicher Merkmalskombinationen). Lediglich im Teilmodell zur *Nutzung von Experimenten und Modellen* blieben einige Zellen unbesetzt, da sich zeigte, dass Anforderungen in diesem Bereich sehr schnell hohe Komplexitäten aufweisen und so einige Anforderungen bei niedrigen Komplexitäten nicht ohne weiteres zu operationalisieren sind. Die Einordnung der Items in das Modell wird aktuell durch einen externen Experten wiederholt, um höchstmögliche Validität herzustellen.

Die entwickelten Items wurden zu vier Testheften zusammengestellt, welche in den ersten 21 Items identisch sind. Diese „Ankeritems“ decken das Modell bereits recht breit ab. Darüber hinaus enthält jedes Testheft 23 weitere Items, die in keinem anderen Testheft vorkommen. Jedes Testheft enthält somit 44 Items, von denen jeweils 32 der *Nutzung von Fachwissen* und 12 der *Nutzung von Experimenten und Modellen* zuzuordnen sind. Dieses Design ermöglicht den Einsatz möglichst vieler verschiedener Testitems.

Neben den Testitems enthält jeder Test einen Abschnitt mit demographischen Daten: Studiendauer, Studiengang, Umfang der besuchten Lehrveranstaltungen, Vorwissen (Kurswahl, letzte Physik- und Mathematiknoten). Die Testdauer beträgt 90 Minuten, so dass der Test in den üblichen universitären Lehrveranstaltungen eingesetzt werden kann.

Um eine valide und reliable Kodierung der Antworten zu ermöglichen, wurde zu jedem Item eine Musterlösung erarbeitet und in einem Kodiermanual festgehalten. Bei Items, die eine Rechnung verlangen, wird dabei jeweils auch mitkodiert, ob ein richtiger Ansatz zur (evtl. fehlerhaften) Rechnung gezeigt wurde.

4. Stichprobe der Pilotierung

Der entwickelte Test wurde im Sommersemester 2011 und Wintersemester 2011/2012 an insgesamt N = 100 Studenten der Universität Paderborn pilotiert. Die Stichprobe der Pilotierung (vgl. Tab. 1) deckt dabei zu jeweils etwa einem Viertel Studenten der Fach-Physik, untere Lehrämter Physik (v.a. Haupt- und Realschul-Lehramt), obere Lehrämter Physik (v.a. Gymnasial-Lehramt) und Studierende der Elektrotechnik mit Physik als Anteilsfach ab. Es wurden keine Studenten in Lehramts-Bachelor-Studiengängen erfasst, da die Lehramts-Studiengänge in

Paderborn erst während des Pilotierungszeitraumes auf Bachelor-Studienordnungen umgestellt wurden.

Die Probanden streuen über alle Semester des Grundstudiums. Lediglich die E-Techniker wurden ausschließlich im ersten Semester erfasst. Für die Pilotierung wurde von verschiedenen Dozenten der Physik Vorlesungs- oder Übungs-Stunden zur Verfügung gestellt.

5. Empirische Erkenntnisse

Die Antworten der Probanden wurden im Rahmen einer Rasch-Analyse verwendet, um jedem Probanden einen Fähigkeitswert und jedem Item einen Schwierigkeitswert zuordnen zu können. Bevor diese Werte genutzt werden können, müssen zunächst mehrere Rasch-Modelle, die sich in der Anzahl ihrer Dimensionen unterscheiden, auf ihre Passung auf die Test-Daten hin untersucht werden.

5.1 Modellvergleiche und Gütekriterien

In der Frage der Trennbarkeit der beiden Kompetenzkonstrukte *Nutzung von Fachwissen* und *Nutzung von Experimenten und Modellen* ist zunächst von Interesse, ob ein 2D-Rasch-Modell (Modell M2 in Tab. 2), bei dem die Items in diese beiden Dimensionen aufgeteilt werden, einem 1D-Modell (Modell M1), bei dem alle Items auf dieselbe Dimension laden, in einem χ^2 -Anpassungstest überlegen ist. Ist dies der Fall, kann für beide Teilkonstrukte jeweils getestet werden, ob eine Aufteilung in weitere Unterbereiche wiederum einen Vorteil bringt. Hier wurde für die *Nutzung von Fachwissen* die Aufteilung nach *Fach-Stufe* (Modell M3) und bei der *Nutzung von Experimenten und Modellen* die Aufteilung nach *Suchraum* (Modell M4) getestet. Einen Überblick über die jeweiligen Modellpassungen zeigt Tab. 2.

Es zeigt sich, dass ein Modell, das zwischen den Dimensionen *Nutzung von Fachwissen* und *Nutzung*

	N	Abitur-Note	Letzte Physiknote	Fach-Semester	Anteil Weiblich
Untere Lehrämter¹	28	2,7	2,1	3,8	35,7%
Obere Lehrämter²	27	2,1	1,8	4,0	33,3%
Fach Physik (B.Sc.)	21	2,1	1,7	3,3	14,3%
E-Technik³	24	2,5	1,9	1,0	4,2%
Gesamt	100	2,4	1,9	3,1	23,0%

Tab. 1: Überblick über die Stichprobe der Pilotierung. Für die Noten und Semesterzahlen sind jeweils die Mittelwerte der Teilstichproben angegeben. 1) 25 Studierende LA Haupt-/Realschule; 3 Studierende LA Grundschule 2) 25 Studierende LA Gymnasium; 2 Studierende LA Berufskolleg 3) Ausschließlich erstes Semester, Veranstaltung „Physik für E-Technik“

	M1	M2	M3	M4
Dimensionen	1D	FW + EM	3*FW + EM	FW + 2*EM
Final Deviance	4763,26	4734,13	4717,18	4731,40
Geschätzte Parameter	126	128	135	131
Modellvergleich (χ^2-Anpassungstest)		*** →	** →	n.s. →

Tab. 2: Übersicht über die getesteten Rasch-Modelle. ***: p<0,001; **: p<0,01

von Experimenten und Modellen unterscheidet (M2) höchstsignifikant besser passt, als ein Modell, dass diese Unterscheidung nicht trifft (M1). Teilt man weiter die Nutzung von Fachwissen in die drei Dimensionen Schulwissen, vertieftes Wissen und universitäres Wissen auf, erhält man zwar eine weiter verbesserte Modellpassung, allerdings korrelieren die drei Teildimensionen latent mit Koeffizienten von jeweils über 0,9 bis hin zu 0,95, so dass in Frage gestellt werden muss, ob es sich hierbei um empirisch klar unterscheidbare Merkmale handelt. Die Aufteilung der Dimension Nutzung von Experimenten und Modellen nach Suchraum liefert im χ^2 -Test keine bessere Modellanpassung. Aus diesen Gründen und weil die Einstufung der Items auf diesem Detailniveau zum jetzigen Zeitpunkt (Stand Mai 2012) noch nicht vollständig validiert wurde, wird im Folgenden zunächst mit dem zweidimensionalen Modell M2 weitergearbeitet.

Der Item-Fit ist bei der Mehrzahl der Items gut. Lediglich 5 Items liegen außerhalb der Grenzen von $MNSQ < 1,25$ und $T < 1,96$.

	FW	EM
Varianz	1,309	0,724
EAP-Reliabilität	0,891	0,726
WLE-Reliabilität	0,878	0,553
Latente Korrelation	0,784	

Tab. 3: Gütekriterien der Rasch-Analyse (Modell M2).

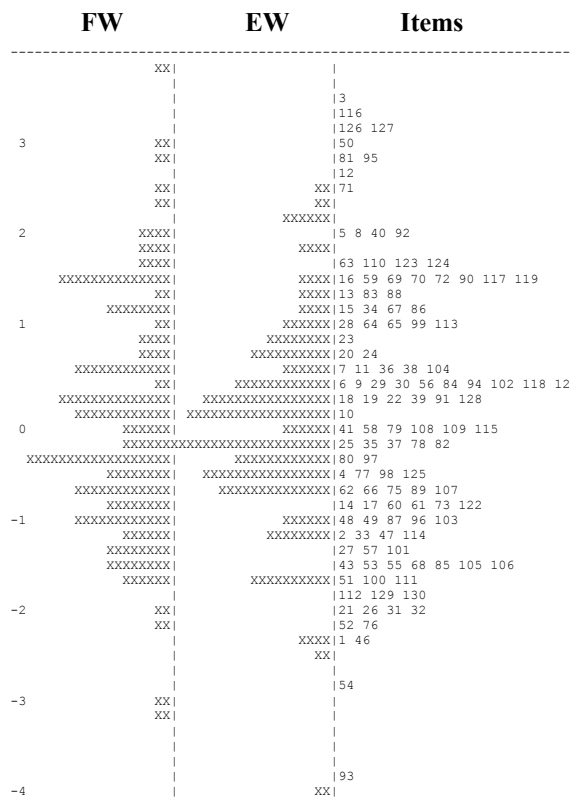


Abb. 4: Wright-Map der Raschanalyse (Modell M2) Jedes 'X' repräsentiert 0.5 Fälle. Erstellt mit ConQuest.

Betrachtet man die Gütekriterien (vgl. Tab. 3) der beiden so entstandenen Skalen, so zeigt die Skala FW (Nutzung von Fachwissen) eine große Varianz und gute Reliabilitäts-Werte. Bei der Skala EM (Nutzung von Experimenten und Modellen) ergeben sich Reliabilitäten am Rande des akzeptablen Bereichs, die aber möglicherweise mit einer geringen Varianz zusammenhängen. Eine latente Korrelation von 0,784 bestätigt das Bild von zwei empirisch gut trennbaren Kompetenzdimensionen.

In der Wright-Map (Abb. 4) zeigt sich für die FW-Dimension eine gute Passung der Itemschwierigkeiten mit den Probandenfähigkeiten. Besonders im oberen (d.h. schwierigeren) Bereich kann der Item-Bestand aber noch ausgedünnt werden. Bei der EM-Dimension finden sich die Probanden-Fähigkeiten vor allem in der oberen Hälfte des vom Test abgedeckten Bereiches, was mit dem Befund einer zu geringen Varianz korrespondiert.

5.2 Gruppenvergleiche

Die Stichprobe enthält mehrere Gruppen, deren Fähigkeiten im Folgenden miteinander verglichen werden sollen. Dabei muss beachtet werden, dass die Daten aus der Pilotierung stammen (in welcher naturgemäß noch nicht die endgültige Version der Testhefte verwendet worden ist) und einige der Teilgruppen mit $N \approx 10$ für verallgemeinerbare Aussagen zu klein sind. Dementsprechend sollen an dieser Stelle lediglich erste Ausblicke auf möglicherweise bei der Hauptstudie zu erwartende Effekte gezeigt werden.

5.2.1 Erstes Semester Fach-Physik – E-Technik

In der Stichprobe finden sich zwei Gruppen, die jeweils zum Ende ihres ersten Semesters getestet wurden. Das sind einerseits die Teilnehmer der Physik A (WS 2011/12; Fach-Physik und obere Lehrämter; 4 SWS; $N = 14$) und die Studierenden der E-Technik (Physik für E-Techniker, WS 2011/12; 6 SWS; $N = 24$).

Abb. 5 zeigt die durchschnittlichen Fähigkeitswerte der beiden Gruppen. Es ergeben sich keine signifikanten Unterschiede zwischen den Gruppen. Die Effektstärke (Cohens d) zeigt beim Fachwissen einen mittleren ($d = .35$) und bei der Nutzung von Experi-

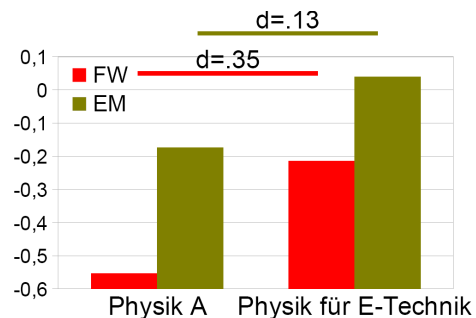


Abb. 5: Vergleich der durchschnittlichen Probanden-Fähigkeiten (WLE-Schätzwerte) in den Erstsemester-Veranstaltungen Physik A (Fach-Physik und obere Lehrämter) und Physik für E-Techniker.

menten und Modellen lediglich einen schwachen Effekt ($d = .13$). Diese leichten Vorteile lassen sich möglicherweise durch die höhere Time on Task (also die Zeit der tatsächlichen Auseinandersetzung mit den Lerninhalten) in der Veranstaltung für E-Techniker erklären.

5.2.2 Vergleich erstes, drittes, fünftes Semester

Für die Gruppe der Fach-Studenten und der oberen Lehrämter lässt sich ein Vergleich der Teilstichproben der Besucher der Lehrveranstaltungen für das erste ($N = 11$), dritte ($N = 17$) und fünfte ($N = 10$) Semester (WS 2011/12) anstellen. In den ersten vier Semestern belegen diese Studierenden in der Physik dieselben Lehrveranstaltungen. Die Festkörperphysik im fünften Semester ist für die Lehramts-Studenten eine Wahl- und für die Fach-Studenten eine Pflichtveranstaltung.

In Abb. 6 zeigt sich zunächst ein deutlicher Vorteil (5%-Signifikanzniveau, hohe Effektstärke) der Teilnehmer der Physik C (drittes Semester) gegenüber der Physik A (erstes Semester).

Der Vergleich des dritten mit dem fünften Semester zeigt allerdings einen gegenteiligen Effekt. Während sich für das Fachwissen noch nicht-signifikanter Effekt mittlerer Stärker zeigt, zeigen gerade die Lehramts-Studenten im Bereich der Nutzung von Experimenten und Modellen einen signifikanten Effekt mit Cohens d von 1.54. Möglicherweise handelt es sich

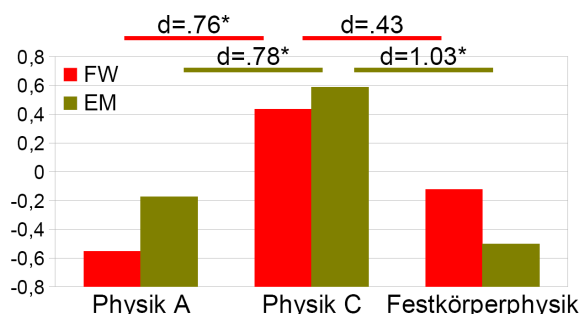


Abb. 6: Vergleich der durchschnittlichen Probanden-Fähigkeiten (WLE-Schätzwerte) in den drei Veranstaltungen für Fach-Studenten und obere Lehrämter im ersten (Physik A), dritten (Physik C) und fünften (Festkörperphysik, Wahlpflicht) Semester.

hier um den Effekt einer Negativauswahl, der im Rahmen einer Regressionsanalyse geprüft werden kann. Alternativ wäre es möglich, dass der Inhaltsbereich Mechanik für diese Stichprobe nicht im notwendigen Maße geeignet ist, allgemeine physikalische Lernfortschritte abzubilden.

5.3 Regressionsanalyse – Einflussfaktoren

Im Rahmen einer Regressionsanalyse kann geprüft werden, welche äußeren Faktoren wie Studiengang, Vorwissen oder Demographische Daten, statistisch mit den erreichten Fähigkeitswerten zusammenhängen.

In Tab. 4 ist eine Übersicht über drei für die Fähigkeiten zur *Nutzung von Fachwissen* geprüften linearen Regressionsmodelle. Bereits ein recht einfaches Modell (FW-1) aus den Semesterwochenstunden (als Maß für die Time on Task), der Abiturnote (als Indikator für die allgemeine kognitive Leistungsfähigkeit) und dem Geschlecht (hier ergeben sich Zusammenhänge mit dem Selbstkonzept, das im Rahmen dieser Studie nicht eigens erhoben wurde, vgl. [9]) liefert eine Varianzaufklärung von $R^2 = 31,5\%$. Nimmt man Indikatoren des Vorwissen dazu (Modell FW-2), so sind nur die letzte Physiknote (aber nicht die Leistungs- oder Grundkurswahl) und die Wahl eines Mathematik-Leistungskurses (aber nicht die erreichte Abschlussnote) signifikante Einflussfaktoren. Zuletzt wurde geprüft, ob sich ein signifikanter Unterschied zwischen Physik-Studenten (Fach- oder Lehramt) und E-Technikern ergibt (Modell FW-3). Dies ist nicht der Fall. Im Rahmen der Pilotierung (es wurden *ausschließlich* E-Techniker im ersten Semester befragt!) ist hier also noch kein Unterschied im Fachwissen-Erwerb zwischen den Studiengängen zu sehen.

Insgesamt können somit die Ergebnisse von Riese [9] reproduziert werden, wobei weitere Einflussfaktoren wie etwa das fachbezogene Selbstkonzept, Beliefs und motivationale Orientierungen mit den vorhandenen Daten nicht geprüft werden konnten.

Bei der *Nutzung von Experimenten und Modellen* (Tab. 5) zeigt sich ein sehr viel unklareres Bild. Nur wenige der im Test erfassten Demographischen Faktoren haben einen Einfluss, der in die Nähe der 5%-Signifikanzschwelle kommt. Diese Skala ist insge-

Modell	FW-1		FW-2		FW-3	
	β	p	β	p	β	p
1 SWS im Fach	.226	.013	.215	.013	.220	.017
Abiturnote	-.430	.000	-.336	.001	-.340	.001
Geschlecht	.223	.013	.211	.011	.206	.020
2 Letzte Physiknote			-.266	.008	-.262	.001
Mathe LK gewählt			.282	.001	.281	.001
3 E-Techniker					.016	.858
p Modellverbesserung	.000		.000		.000	
R² korrigiert	.315		.427		.420	

Tab. 4: Regressionsanalyse für die Fähigkeitswerte der Skala *Nutzung von Fachwissen*.

samt vom Fachwissen-Fähigkeitswert dominiert. Insbesondere sieht man hier keinen Geschlechter-Effekt, keinen signifikanten Zusammenhang mit der Dauer der absolvierten Laborpraktika (auch nicht, wenn man diese nach Studiengängen auftrennt) und keinen Einfluss von Fach-SWS und Abiturnote, der über die Konfundierung mit der Fachwissen-Skala hinausgeht. Demnach wäre es hier sinnvoll, weitere Faktoren in die Analysen einzubeziehen. Vor allem die epistemologischen Beliefs, Ansichten über die Natur der Naturwissenschaften (NOS und NOSI) und die explizite Thematisierung des wissenschaftlichen Erkenntnisprozesses in der bisherigen Ausbildung wären hier von Interesse. Auch eine differenziertere Aufklärung der Zusammenhänge einzelner Teile des Fachwissens mit dieser Skala ist notwendig, sofern die Ergebnisse der laufenden Validierungsprozesse eine differenziertere Betrachtung zulassen.

6. Fazit und Ausblick

Aufbauend auf einem Strukturmodell fachlichen Wissens von Physikern, welches die beiden Bereiche *Nutzung von Fachwissen* und *Nutzung von Experimenten und Modellen* unterscheidet, wurde ein umfangreicher Test entwickelt und modelliert. Bei guten bis zufriedenstellenden Gütekriterien konnten die beiden Teilbereiche auch empirisch getrennt werden, so dass jedem Probanden ein FW- und ein EM-Fähigkeitswert zugewiesen wurde.

Erste Analysen konnten die aus [9] bekannten Zusammenhänge des Fachwissenserwerbs für die FW-Skala reproduzieren. Die EM-Skala zeigt einerseits in diesem Stadium der Instrumententwicklung noch Schwächen in den statistischen Kennwerten und andererseits eine Konfundierung mit den FW-Fähigkeiten, die durch weitere Differenzierungen und das Hinzuziehen weiterer Informationen wie Beliefs und konkreten Studieninhalten in diesem Bereich weiter aufgeklärt werden muss.

Erste Gruppenvergleiche zeigen jedoch, dass der Test sinnvoll bei Lehramts-, Fach- und Nebenfach-Studenten der Physik eingesetzt werden kann.

Das Testinstrument wird in den nächsten Schritten durch Expertenbefragungen einerseits und dem Einsatz lauten Denkens zur Aufklärung der Prozesse bei der Aufgabenlösung andererseits validiert. Anhand von inhaltlichen Erwägungen, Aspekten der Model-

labdeckung und statistischen Gütekriterien werden einzelne Items aus dem Test entfernt. Ein weiterer Testteil zu Beliefs, NOS-/NOSI-Verständnis und zum Selbstkonzept wird ergänzt. Mit dem so überarbeiteten Instrument wird eine Erhebung von mindestens $N = 300$ Studenten der Physik angestrebt.

7. Literatur

- [1] Shulman, L. (1986). Those who Understand: Knowledge Growth in Teaching. In *Educational Researcher*, 15, S. 4–14.
- [2] Baumert, J. & Kunter, M. (2006). Stichwort: Professionelle Kompetenz von Lehrkräften. In *Zeitschrift für Erziehungswissenschaft*, 9 (4), S. 469–520.
- [3] 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.
- [4] 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.
- [5] 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. In *American Educational Research Journal*, pp 133-180.
- [6] Anderson, C.W. (1979). An observational study of classroom management and information structuring in elementary school science lessons. In *Dissertation Abstracts International*, 40(11), 5810A (UMI no. 8009823).
- [7] Terhart, E. (2002). Standards für die Lehrerbildung. Eine Expertise für die Kultusministerkonferenz (ZKL-Texte Nr. 23). Universität Münster: Zentrale Koordination Lehrerbildung.
- [8] 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.
- [9] Riese, J. (2009). Professionelles Wissen und professionelle Handlungskompetenz von (ange-

Modell	EM-1		EM-2		EM-3	
	β	p	β	p	β	p
1 SWS im Fach	.210	.078	.274	.047	.054	.647
Abiturnote	-.243	.042	-.250	.037	.043	.697
2 SWS Labor-Praktika			-.127	.349	-.038	.738
3 FW-Fähigkeit					.653	.000
p Modellverbesserung	.014		.025		.000	
R ² korrigiert	.095		.094		.386	

Tab. 5: Regressionsanalyse für die Fähigkeitswerte der Skala *Nutzung von Experimenten und Modellen*.

- henden) Physiklehrkräften. Dissertation. Berlin: Logos Verlag.
- [10] 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.
- [11] 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.
- [12] 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.
- [13] Woitkowski, D., Riese, J., Reinhold, P. (2011). Modellierung fachwissenschaftlicher Kompetenz angehender Physiklehrkräfte. In *Zeitschrift für Didaktik der Naturwissenschaften*, 17, S. 289–313.
- [14] 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.
- [15] 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.
- [16] 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.
- [17] Schecker, H. & Parchmann, I. (2006). Modellierung naturwissenschaftlicher Kompetenz. In *Zeitschrift für Didaktik der Naturwissenschaften*, 12, S. 45–66.
- [18] Metzger, S. & Labudde, P. (2007). HarmoS Naturwissenschaften+. Bildungsstandards für die Schweiz. In *Praxis der Naturwissenschaften – Physik in der Schule*, 56 (6), S. 14–18.
- [19] 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.
- [20] Friege, G. & Lind, G. (2004). Leistungsmessung im Leistungskurs. In: *Der Mathematische und Naturwissenschaftliche Unterricht – MNU*, 57(5), S. 259–265.
- [21] Kauertz, A. (2008). *Schwierigkeitserzeugende Merkmale physikalischer Leistungstestaufgaben*. Dissertation. Berlin: Logos Verlag.
- [22] Bernholt, S. (2010). *Kompetenzmodellierung in der Chemie – Theoretische und empirische Reflexion am Beispiel des Modells hierarchischer Komplexität*. Dissertation. Berlin: Logos Verlag.
- [23] Klein, F. (1933). *Elementarmathematik vom höheren Standpunkte aus*. 4. Auflage, Nachdruck 1968, Berlin: Springer.
- [24] Kirsch, A. (1987). *Mathematik wirklich verstehen*. Köln: Aulis Verlag Deubner.
- [25] Mayer, J. (2007). Erkenntnisgewinnung als wissenschaftliches Problemlösen. In: Krüger, D., Vogt, H.: *Theorien in der biologiedidaktischen Forschung*. Berlin/Heidelberg: Springer, S. 177–186.
- [26] Klahr, D. (2000) *Exploring Science: The Cognition and Development of Discovery Processes*. Cambridge, MA: MIT Press.
- [27] 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