

Ko-WADiS – Kompetenzmodell der Erkenntnisgewinnung Kompetenzmodellierung und -erfassung über naturwissenschaftliche Arbeits- und Denkweisen

Philipp Straube, Volkhard Nordmeier

Freie Universität Berlin, Fachbereich Physik, Arnimallee 14, 14195 - Berlin
philipp.straube@fu-berlin.de, volkhard.nordmeier@fu-berlin.de

Kurzfassung

Im Projekt „Ko-WADiS“ (**K**ompetenzmodellierung und -erfassung zum **W**issenschaftsverständnis über naturwissenschaftliche **A**rbeits- und **D**enkweisen bei Studierenden (Lehramt) in den drei naturwissenschaftlichen Fächern Biologie, Chemie und Physik) soll ein Kompetenztest über naturwissenschaftliche Erkenntnisgewinnung in den drei beteiligten Fächern entwickelt werden. Zu diesem Zweck wurde ein Kompetenzmodell adaptiert und Aufgaben für dieses Modell entwickelt, die zunächst im offenen Format konstruiert und anschließend bei Studierenden der beteiligten Fächer eingesetzt wurden. Aus den daraus erhaltenen Antworten wurden Multiple-Choice Items konstruiert. Nachdem diese mehrere Qualitätssicherungsschritte durchlaufen haben, werden sie erstmalig im Sommersemester 2013 an Studierenden pilotiert. Im Beitrag werden der Konstruktions- und Validierungsprozess der Items vorgestellt.

1. Einleitung

Das Projekt „Ko-WADiS“ [1; 2] ist ein Kooperationsprojekt zwischen der Didaktik der Biologie, der Didaktik der Physik (beide Freie Universität Berlin), der Didaktik der Biologie und der Didaktik der Chemie (beide Humboldt Universität zu Berlin). Das erklärte Ziel ist die Entwicklung eines Testinstruments zur Erhebung der Kompetenzen Lehramtsstudierender im Bereich der naturwissenschaftlichen Erkenntnisgewinnung. Zentrale Fragestellungen des Projekts umfassen auch die Struktur und die Entwicklung der naturwissenschaftlichen Erkenntnisgewinnung.

Im diesem Beitrag soll die Entwicklung des Kompetenztests dargestellt werden. Insbesondere wird auf die mehrere Schritte umfassende Aufgabenkonstruktion eingegangen.

2. Kompetenzmodell

Unter- suchungen	Frage- stellung	Hypo- these	Planung und Durch- führung	Aus- wertung und Reflexion
Modelle nutzen	Zweck von Modellen	Testen von Modellen	Ändern von Modellen	

Tab. 1: Kompetenzmodell nach Upmeier zu Belzen & Krüger [3] und Mayer [4]

Grundlage eines Kompetenztests ist ein theoretisches Kompetenzstrukturmodell. Dieses wird häufig

aus der inhaltlichen Struktur des Gegenstandsbereiches hergeleitet [5]. Diese Modelle müssen „durch empirische fachdidaktische und lernpsychologische Forschung“ [5] abgesichert werden. Dieser Forderung soll in diesem Projekt nachgekommen werden.

Das im Projekt genutzte Kompetenzmodell wurde auf Grundlage bestehender Modelle [4; 3] adaptiert. Es umfasst die Bereiche „Untersuchungen“ und „Modelle Nutzen“ mit jeweiligen Teilkompetenzen (siehe Tabelle 1).

3. Aufgabenkonstruktion

Ausgehend von diesem Kompetenzstrukturmodell wurden Testaufgaben, sog. Items, entwickelt, die auf eine bestimmte Teilkompetenz des Modells abzielen. Schecker & Parchmann [5] stellen dazu fest, dass „ein psychometrisch optimales Item [...] sich einer ganz bestimmten Zelle der Kompetenzmatrix zuordnen“ lässt. Entsprechend abgegrenzt mussten Itemstamm und -impuls formuliert sein. Darauf ist ohnehin zu achten, um auszuschließen, dass andere Lösungswege als die intendierten möglich sind [5]. Zusätzlich mussten Aufgaben so formuliert sein, dass eine Konfundierung mit anderen Kompetenzausprägungen ausgeschlossen ist [5]. Das ist gerade in diesem Projekt eine Herausforderung, da hier eine fächerübergreifende naturwissenschaftlichen Erkenntnisgewinnungskompetenz angenommen wird. Die Aufgaben mit einem physikalischen Kontext müssen daher auch für Biologie- und Chemielehramtsstudierende lösbar sein. Eine Konfundierung mit universitärem Fachwissen musste daher ausgeschlossen werden, weshalb bei allen Aufgaben die zur Lösung notwendigen Informationen im Stamm mitangegeben werden mussten.

Die Items des Kompetenztests sollten im Multiple-Choice-Format entwickelt werden. Multiple-Choice Items sind in large-scale-Studien ökonomischer [6]. Außerdem weisen sie eine höhere Auswertungsobjektivität auf [7]. Herausforderung bei der Konstruktion von Multiple-Choice-Items ist nicht zwangsläufig die korrekte Antwort (Attraktor), sondern vielmehr die Auswahl geeigneter falscher Antworten (der sog. Distraktoren). Diese müssen wie richtige Antworten aussehen, aber inhaltliche Fehler aufweisen [7]. Die drei Hauptaspekte, die bei der Auswahl der Distraktoren beachtet werden müssen, sind ihre Auswahlwahrscheinlichkeit (Attraktivität), die Ähnlichkeit mit dem Attraktor und ihre eigene Plausibilität [7]. Um diese Punkte zu erfüllen, sind als Quelle der falschen Antwortmöglichkeiten echte Antworten der Probanden der ‚Fantasie‘ des Aufgabenstellers vorzuziehen. Echte, aber falsche Antworten sind meistens authentischer und verhindern dadurch die Möglichkeit, die Aufgaben durch Ausschluss von Antworten zu lösen [8]. Um derartige Antworten zu erhalten, wird von verschiedenen Autoren [7; 8] vorgeschlagen, die Aufgaben zunächst im halboffenen Format zu entwickeln und einzusetzen. Die Antworten werden anschließend analysiert und ausgewertet. Häufig auftretende falsche Antworten können dann als Distraktoren genutzt werden.

Im Projekt Ko-WADiS wurde diese Möglichkeit in abgewandelter Form zur Itemkonstruktion genutzt. Im Gegensatz zum oben genannten Verfahren wurden die Aufgaben zunächst im offenen Format entwickelt. Um trotz der größeren Freiheit in der Gestaltung Antworten zu erhalten, die der intendierten Richtung entsprechen, mussten die Impulse entsprechend eng formuliert sein. Teilweise wurden diese mit vorgeschalteten, dichotomen Ja/Nein Fragen ergänzt. Dieses Vorgehen sollte sicherstellen, dass die Antworten der Studierenden inhaltlich breit gestreut sind. Hinzu kommt eine schlechte Eignung bestimmter Itemformate (z.B. Beschreibungen von Experimentiersettings) für das halboffene Format. Die fertigen Aufgaben müssen durch Experten bewertet werden, „die sich sowohl mit dem Modell als auch mit der unterrichtlichen Praxis auskennen.“ [5] Übertragen auf dieses Projekt sind demnach projektinterne Hochschullehrer geeignete Experten. Begleitend zur Konstruktion erfolgte daher eine projektinterne Einschätzung der Aufgaben durch einen Experten des eigenen Faches, sowie durch mindestens einen Experten eines der anderen beiden beteiligten Fächer. Durch das Hinzuziehen von fachfremden Experten sollte sichergestellt werden, dass die Aufgaben tatsächlich unabhängig vom spezifischen Fachwissen lösbar sind. Zusätzlich wurden die Aufgaben durch einen psychometrisch ausgewiesenen Experten bewertet und anschließend im Rahmen einer Prä-Pilotierung an Studierenden diverser Semester eingesetzt.

Die in der Prä-Pilotierung erhaltenen Antworten liegen selbstverständlich in sprachlich völlig ver-

schiedenen Formaten vor, weshalb diese nicht wortwörtlich in die Multiple-Choice-Items zu übernehmen sind. Vorher mussten Sprache und Format vereinheitlicht werden. Als Grundlage für die sprachliche Anpassung der Items diente im Wesentlichen [9]. Die aus Sicht der Autoren wichtigsten Richtlinien waren dabei

- eine sprachlich parallele Form,
- eine ähnliche Länge,
- eine gleichmäßige Verwendung bzw. Auslassung von Schlüsselwörtern und
- die Vermeidung von Fachwörtern.

Durch die Parallelisierung der Antwortmöglichkeiten waren die Satzanfänge häufig gleich, weshalb diese nur einmal genannt wurden und der zweite Satzteil durch Ankreuzen entsprechend ergänzt werden musste (siehe auch [9]). Die in der Prä-Pilotierung teilweise enthaltenen Ja/Nein-Fragen wurden in die Antwortmöglichkeiten integriert. Jede Aufgabe bot vier Antwortalternativen an, von denen aber jeweils nur eine richtig war.

Die fertig generierten Items wurden anschließend abermals einer Bewertung durch die oben genannten Experten zugeführt und ggf. erneut überarbeitet.

4. Vorgehen

Der vorgenannte Konstruktionsprozess wird nun im Folgenden an einem Beispiel verdeutlicht:

Eine Aufgabe soll im Bereich „Untersuchungen“ für die Kategorie „Fragestellung“ entwickelt werden. Fragestellungen definieren [10] als “asking relevant questions concerning the phenomena that they have observed.”. Im Sinne der Popper’schen Forschungslogik [11] gehen wir vom Paradigma der grundsätzlichen Falsifizierbarkeit einer wissenschaftlichen These aus. Dementsprechend dürfen Fragestellungen nicht so gestellt werden, dass eine Falsifizierbarkeit von vornerein ausgeschlossen ist. Außerdem dürfen bewährte Thesen nicht ohne Grund fallen gelassen werden [11]. Dementsprechend müssen Fragestellungen konform mit der entsprechenden Theorie aufgestellt werden [12].

Im folgenden Beispiel werden den Studierenden die theoretischen physikalischen Grundlagen für einen Bereich geliefert. Die Theorie ist so ausführlich geschrieben, dass sich auch Studierende der anderen Fächer mit einem rudimentären schulischen Grundwissen in Physik in das Thema hineindenken können. Anschließend wird ein Phänomen präsentiert, das im scheinbaren Widerspruch zu der Theorie steht. Die Studierenden sind aufgefordert, eine naturwissenschaftliche Fragestellung zu benennen, die das Phänomen erklären könnte.

Die hier gezeigten Antworten sind exemplarisch ausgewählt. Weder bilden Sie das vollständige Spektrum der gegebenen Antworten ab, noch handelt es sich bei den vier Antworten zwangsläufig um die später genutzten Multiple-Choice-Auswahlmöglichkeiten.

Aufgabenbeispiel:

<i>Theorie</i>	Die Erdbeschleunigung ist für alle Körper unabhängig von ihrer Masse gleich. Sie beträgt in guter Näherung $g = 9,81 \frac{m}{s^2}$. Wirkt also nur die Erdanziehungskraft, werden alle Körper gleichermaßen beschleunigt. Fallen zwei Körper gleichzeitig aus gleicher Höhe, erreichen sie demnach gleichzeitig und mit gleicher Geschwindigkeit den Boden.
<i>Phänomen</i>	Im Herbst fallen Blätter und Früchte der Kastanie auf den Boden. Während die Früchte schnell auf den Boden fallen, segeln die Blätter langsam nach unten.
<i>Impuls</i>	Formulieren Sie eine naturwissenschaftliche Fragestellung, die sich aus diesem Phänomen ableiten lässt.
<i>Antworten der Studierenden</i>	„Inwieweit beeinflussen andere Faktoren wie z.B. Luftwiderstand und Form der Gegenstände, die Erdbeschleunigung?“ „Wird die Frucht schneller beschleunigt von der Erde?“ „Wie/Wo wird Auftrieb in der Natur deutlich/nachweisbar?“ „Welche Kraft bewirkt, das [sic!] Blätter nicht genauso schnell runter fallen wie Früchte?“

Tab. 2: Aufgabenbeispiel

Die Fragestellungen in den Antworten 1 und 2 stellen eine Änderung der Erdbeschleunigung in den Fokus der Untersuchung. Dies widerspricht der im Stamm genannten Theorie, da die Konstanz der Erdbeschleunigung für alle Körper im Stamm deutlich gemacht wird. Insofern sind diese beiden Antworten als falsch einzustufen. Antwort 3 fokussiert falsch, denn durch eine Beantwortung dieser Frage würde das in der Aufgabe beschriebene Phänomen nicht beantwortet werden. Selbst unter der – fachlich falschen – Annahme, dass der Auftrieb für das unterschiedlich schnelle Fallen verantwortlich wäre, könnte durch die Frage „Wie/Wo wird Auftrieb in der Natur deutlich/nachweisbar?“ nicht geklärt werden, wieso Blätter und Früchte unterschiedlich schnell fallen. Durch die Beantwortung der Frage 4 kann das Phänomen erklärt werden, weshalb diese als richtig einzustufen ist.

Bei der Analyse der Antworten zeigten sich zwei Sorten von geeigneten Distraktoren: Zum einen jene, die die theoretische Basis missachten und zum anderen solche, die eine Frage stellen, deren Beantwortung nicht das beschriebene Phänomen klären wür-

de. Dieses grundsätzliche Unterscheidungsmerkmal zeigte sich auch bei anderen Aufgaben zur Teilkompetenz „Fragestellung“.

5. Begleitfragen

Um die erstellten Items einschätzen zu können, wurden die Studierenden aufgefordert die Aufgaben unter bestimmten Aspekten zu bewerten. Die Einschätzung wurde jeweils auf einer sechs-stufigen Likert-Skala erbeten, wobei 1 wenig bzw. gering bedeutet und 6 viel bzw. hoch. Die einzelnen Aspekte der Befragung und die Ergebnisse für die eben genannte Beispielaufgabe sind in der folgenden Tabelle dargestellt. Auf eine Interpretation der Ergebnisse wird an dieser Stelle bewusst verzichtet:

Frage	M
Wie sicher waren Sie beim Lösen dieser Aufgabe?	4.6 (1.1)
Wie sehr mussten Sie sich anstrengen, um zur Lösung zu gelangen?	2.9 (1.3)
Inwiefern waren hilfreich...	
Fachwissen?	4.1 (1.4)
Methodenkenntnisse?	3.6 (1.6)
logisches Schlussfolgern	4.1 (1.6)

Tab. 3: Studierendeneinschätzung zur Aufgabe mit Mittelwert und Standardabweichung

Des Weiteren wurden die Studierenden befragt, ob Sie weitere Hinweise zu der Aufgabe hatten. Erhofft wurde hierbei, Erkenntnisse über unverständliche Formulierungen oder Fachwörter zu erhalten. Für diese Aufgabe wurden keine Anmerkungen gemacht.

6. Ausblick

Die fertigen Items wurden im Sommersemester 2013 pilotiert, anschließend selektiert und ggf. überarbeitet. In der darauf folgenden längsschnittlichen Studie werden die Studierenden jeweils im 1. und 4. Semester des Bachelors und des Master befragt. Als Kontrollgruppe werden zusätzlich an zwei Messzeitpunkten außerdem die Mono-Bachelor- bzw. Master-Studierenden der beteiligten Fächer befragt. Bereits nach zwei Messzeitpunkten liegen ausreichend Daten für eine quasilängsschnittliche Auswertung vor. Nach 5 Jahren hat eine Kohorte alle Messzeitpunkte durchlaufen.

Das Projekt wird im Rahmen des Programms „Kompetenzen im Hochschulsektor“ (KoKoHs) durch das BMBF gefördert.

7. Literatur

- [1] Straube, Philipp; Nordmeier, Volkhard (2013): Längsschnitt zur SI-Kompetenzerfassung Physik(Lehramts)-Studierender. In: Bernholt, Sascha (Hrsg.): Inquiry-based Learning – Forschendes Lernen, 539–541.
- [2] Stiller, Jurik; Tiemann, Rüdiger (2013): Scientific Inquiry im Studium Längsschnitt zur Kompetenzerfassung im Lehramt Chemie. In: Bernholt, Sascha (Hrsg.): Inquiry-based Learning – Forschendes Lernen, 641–643.
- [3] Upmeyer zu Belzen, Annette; Krüger, Dirk (2010): Modellkompetenz im Biologieunterricht. In: Zeitschrift für Didaktik der Naturwissenschaften 16, 41–57.
- [4] Mayer, Jürgen (2007): Erkenntnisgewinnung als wissenschaftliches Problemlösen. In: Krüger, Dirk; Vogt, Helmut (Hrsg.): Theorien in der biologiedidaktischen Forschung, 177–184.
- [5] Schecker, Horst; Parchmann, Ilka (2006): Modellierung naturwissenschaftlicher Kompetenz. In: Zeitschrift für Didaktik der Naturwissenschaften 12, 45–66.
- [6] Stecher, Bryan M.; Klein, Stephen P. (1997): The Cost of Science Performance Assessments in Large-Scale Testing Programs. In: Educational Evaluation and Policy 19, 1–14.
- [7] Moosbrugger, Helfried; Kelava, Augustin (2012): Testtheorie und Fragebogenkonstruktion.
- [8] Bortz, Jürgen; Döring, Nicola (2006): Forschungsmethoden und Evaluation.
- [9] Burton, Steven; Sudweeks, Richard; Merrill, Paul; Wood, Bud (1991): How to Prepare Better Multiple-Choice Test Items: Guidelines for University Faculty.
- [10] Hofstein, Avi; Navon, Oshrit; Kipnis, Mira; Mamluk-Naaman, Rachel (2005): Developing students' ability to ask more and better questions resulting from inquiry-type chemistry laboratories. In: Journal of Research in Science Teaching 42, 7, 791–806.
- [11] Popper, Karl R. (1984): Logik der Forschung.
- [12] Raithel, Jürgen (2008): Quantitative Forschung.