

NICOLE WELLNITZ, HANS E. FISCHER, ALEXANDER KAUERTZ, JÜRGEN MAYER,
IRENE NEUMANN, HANS ANAND PANT, ELKE SUMFLETH UND MAIK WALPUSKI

Evaluation der Bildungsstandards – eine fächerübergreifende Testkonzeption für den Kompetenzbereich Erkenntnisgewinnung

Evaluation of National Educational Standards – an interdisciplinary test
design for the competence area acquirement of knowledge

ZUSAMMENFASSUNG

Ab 2012 wird in Ländervergleichen das Erreichen der Bildungsstandards in den naturwissenschaftlichen Fächern mit dem Ziel der Weiterentwicklung von Unterricht regelmäßig überprüft. Grundlage des Vergleichs sind empirische Daten, die systematisch darüber Auskunft geben, über welche naturwissenschaftlichen Kompetenzen die Schülerinnen und Schüler zum Zeitpunkt der Erhebung verfügen. In diesem Beitrag wird die standardbezogene Kompetenzmodellierung im Rahmen des Projektes „Evaluation der Standards in den Naturwissenschaften für die Sekundarstufe I“ für den Kompetenzbereich Erkenntnisgewinnung vorgestellt. Ausgehend von nationalen und internationalen Befunden wird ein übergreifendes Kompetenzmodell für die Fächer Biologie, Chemie und Physik abgeleitet, das aus den Dimensionen *Kompetenzbereiche, kognitive Prozesse und Komplexität* besteht. Neben der dreidimensionalen Struktur des Modells werden ausdifferenzierte Niveaumerkmale zur kriteriumsorientierten Leistungsinterpretation beschrieben. Abschließend werden die Kompetenzanforderungen durch Aufgabenbeispiele konkretisiert, das Design der Pilotierungsstudie zur Erprobung aller Items und erste Ergebnisse vorgestellt.

Schlüsselwörter: Bildungsstandards, Erkenntnisgewinnung, Kompetenzmodell, Naturwissenschaften, Pilotierung

ABSTRACT

Starting in 2012, the achievement of the National Standards for science education is assessed through a comparison between the federal states aiming at a continuous enhancement of education in schools. The comparison is based on empirical data that systematically informs about the competencies students possess at the time of the survey. This article presents the standard-based modelling of competencies within the project „Evaluation of the National Educational Standards for Natural Sciences at the Lower Secondary Level“. The focus is on the competence area of acquirement of knowledge. Based on national and international results, an interdisciplinary competence model is derived for the science subjects Biology, Chemistry and Physics. The model consists of the dimensions *areas of competence, cognitive processes and complexity*. The paper details the three dimensions and describes level features of criterion-oriented

assessment. Sample items illustrate which competencies are needed to solve the items. Finally, the design of the pilot study as well as first results are presented.

Keywords: Educational Standards, Acquirement of Knowledge, Competence Model, Science, Pilot Study

1 Einleitung

Die problemorientierte Anwendung und Überprüfung naturwissenschaftlicher Wissensbestände einschließlich der kritischen Reflexion methodischer Herangehensweisen gehören zum Kern naturwissenschaftlicher Grundbildung (Bybee, 2002). Angestoßen durch internationale Schulleistungsstudien wie TIMSS 1995 und PISA 2000, die deutschen Schülerinnen und Schülern unterdurchschnittliche Ergebnisse bei der problemorientierten Bearbeitung naturwissenschaftlicher Fragestellungen bescheinigt haben (Baumert et al., 1997; Baumert et al., 2001), wird den wissenschaftsmethodischen Kompetenzen und ihrer eigenständigen Bildungsfunktion mit Einführung nationaler Bildungsstandards und länderspezifischer Kerncurricula ein höherer Stellenwert eingeräumt. Wissenschaftsmethodische Fähigkeiten und Fertigkeiten werden explizit als Bildungsanspruch formuliert und in den Bildungsstandards im Kompetenzbereich Erkenntnisgewinnung ausgewiesen (Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland [KMK], 2005a, b, c). Naturwissenschaftliche Erkenntnisgewinnung ist damit selbst Unterrichtsinhalt

und als gleichberechtigtes Bildungsziel neben dem traditionell im Mittelpunkt stehenden Fachwissen verankert.

Um die in den Standards benannten Anforderungen einer empirischen Überprüfung zugänglich zu machen, müssen Kompetenzmodelle formuliert werden, die das Gefüge der Fähigkeiten von Schülerinnen und Schülern beschreiben (Schecker & Parchmann, 2006). Erst die Operationalisierung von Kompetenz erlaubt die systematische Entwicklung eines Testinstrumentes zur präzisen und trennscharfen Messung einzelner Teilfähigkeiten. Wie und auf welcher theoretischen Basis die in den Bildungsstandards formulierten Erwartungen als Teilkompetenzen fundiert und graduiert werden, soll im Folgenden dargestellt und somit die Überführung der normativen Standards in empirisch prüfbare Konstrukte aufgezeigt werden. Testentwicklung und Erhebung wurden vom Institut zur Qualitätsentwicklung im Bildungswesen (IQB) koordiniert und in Kooperation mit den Fachdidaktiken der Fächer Biologie, Chemie und Physik unter Beteiligung von Lehrkräften aus allen Bundesländern im Projekt „Evaluation der Standards in den Naturwissenschaften für die Sekundarstufe I (ESNaS)“ durchgeführt (Walpuski, Kampa, Kauertz & Wellnitz, 2008).

2 Naturwissenschaftliche Erkenntnisgewinnung

Mit der Ausweisung des Kompetenzbereichs Erkenntnisgewinnung haben die deutschen Bildungsstandards Anschluss an internationale Standards und Curricula gefunden, die sich besonders im Unterricht des angelsächsischen Sprachraums seit fast drei Jahrzehnten an naturwissenschaftlichen Denk- und Arbeitsweisen orientieren (Department for Education and Skills & Qualification and Curriculum Authority [DfES & QCA], 2004; National Research Council [NRC], 1996). Über das Lernen naturwissenschaftlicher Inhalte hinaus werden die Urteils- und Problemlösefähigkeit sowie das Verstehen der Naturwissenschaften als Prozess mit in den naturwissenschaftlichen Unterricht einbezogen (Shamos, 2002). In den amerikanischen National Science Education Standards werden zwei Bereiche – „Abilities necessary to do scientific inquiry“ und „Understanding about scientific inquiry“ – beschrieben, die auf der einen Seite Fähigkeiten zur konkreten Durchführung naturwissenschaftlicher Untersuchungen und auf der anderen Seite ein grundlegendes Verständnis über die Prozesse naturwissenschaftlicher Erkenntnisgewinnung als Bildungsziele benennen (NRC, 1996, S. 105; NRC, 2000, S. 18). Darüber hinaus wird ein „Understanding of the nature of science“, also die kritische Reflexion wissensgenerierender Handlungsweisen sowie ein Verständnis für die Eigenschaften naturwissenschaftlichen Wissens auf einer Metaebene gefordert (NRC, 1996, S. 105).

Basierend auf der Weinert'schen Kompetenzdefinition (Weinert, 2001) werden in den deutschen Bildungsstandards die problemlösende Prüfung naturwissenschaftlicher Hypothesen und das Aufzeigen möglicher Auswirkungen neuer Erkenntnisse auf Gesellschaft und Umwelt als Kompetenzerwartungen formuliert (KMK, 2005a, b, c). Daneben liegt ein weiterer Schwerpunkt auf der Förderung von Kompetenzen, die das Denken in und mit Modellen verlangen. Während in den deutschen Bildungsstandards Modellkompetenz explizit als Erwartung formuliert wird, werden in den amerikanischen Standards eher implizit der Umgang mit gegenständlichen und mentalen Modellen, deren Entwicklung und Testung gefordert (NRC, 1996).

Durchführung, Verständnis und Reflexion naturwissenschaftlicher Erkenntnisprozesse werden international unter den Begrifflichkeiten *Scientific Inquiry* und *Nature of Science* gefasst, deren Konstrukte gemeinsame Anteile aufweisen, aber auf konzeptioneller Ebene trennende Merkmale beinhalten (Lederman, 2007). Das Modellieren beziehungsweise das kritische Reflektieren über Modelle als Bestandteil naturwissenschaftlicher Denk- und Arbeitsweisen wird in diesem Zusammenhang als „Tür-Öffner“ zum Wissenschaftsverständnis diskutiert und entsprechend über die mediale Perspektive hinausgehend nicht nur als Medium zur Veranschaulichung sondern auch als Mittel der Erkenntnisgewinnung verstanden (Leisner, 2005; Upmeyer zu Belzen & Krüger, 2010).

2.1 Naturwissenschaftliche Untersuchungen

Die Naturwissenschaften zeichnen sich durch die systematische Erforschung der belebten und unbelebten Natur aus. Der erfolgreiche Verlauf des hypothetico-deduktiven Erkenntniswegs zur Generierung naturwissenschaftlicher Erkenntnisse erfordert von den Schülerinnen und Schülern verschiedene Teilfähigkeiten, die in der Regel neben einer iterativen und rekursiven Abfolge von auszuführenden wissenschaftlichen Schritten auch ein Verständnis über die Tragfähigkeit des Untersuchungsdesigns verlangen (Hammann, Phan & Bayrhuber, 2007; Klahr, 2000; Wellnitz & Mayer, 2008). Durchführung und Verständnis von Untersuchungen zur Lösung naturwissenschaftlicher Probleme werden im internationalen Sprachraum unter dem Begriff *Scientific Inquiry* beschrieben. Aus der lerntheoreti-

schen Perspektive werden die Prozesse zur Erkenntnisgewinnung als komplexe Problemlöseprozesse verstanden, die durch grundlegende Prozeduren charakterisiert und von Situations- und Personenvariablen beeinflusst werden (Klahr, 2000; Mayer, 2007). Je nach zugrunde liegender Konzeption werden die zu erwerbenden wissenschaftsmethodischen Teilfähigkeiten im Hinblick auf Anzahl und Details in unterschiedlichem Maße ausdifferenziert (Tab. 1).

In nationalen und internationalen Standards und Curricula wird übereinstimmend auf die sinnvolle Verknüpfung der Prozessvariablen „Fragestellungen und Hypothesen formulieren“, „Untersuchungen planen“, „Untersuchungen durchführen“, „Daten auswerten“ und „Daten interpretieren“ fokussiert (DfES & QCA, 2004; KMK, 2005a, b, c; NRC, 1996). Welche Teilfähigkeiten den naturwissenschaftlichen Erkenntnisprozess hauptsächlich prägen und

Tab. 1: Modellierung naturwissenschaftlicher Erkenntnisprozesse

Hammann et al., 2007; Klahr, 2000; Klos, Henke, Kieren, Walpuski & Sumfleth, 2008	Chinn & Malhotra, 2002	Hofstein, Navon, Kipnis & Mamlok-Naaman, 2005	Grube, 2011; Mayer, Grube & Möller, 2008
Hypothesenbildung und Hypothesenerweiterung	Forschungsfragen generieren	Fragen und Hypothesen aufstellen	Fragestellung formulieren
			Hypothesen generieren
Testen von Hypothesen	Untersuchungen planen	Experimente planen	Untersuchungen planen
	Beobachtungen machen	Experimente ausführen	
Analyse von Evidenzen	Ergebnisse erklären	Ergebnisse analysieren	Daten analysieren und Schlussfolgerungen ziehen
	Theorien entwickeln		

als eigenständige Konstrukte empirisch zu trennen sind, wurde bisher nur in wenigen Studien untersucht. Die von Klahr (2000) postulierten Teilkompetenzen des *Scientific Discovery as Dual Search*-Modells – *Hypothesenbildung*, *Testen von Hypothesen* und *Analyse von Evidenzen* – konnten für das Fach Chemie empirisch teilweise nachgewiesen und das Gesamtkonstrukt vom chemischen Fachwissen und von kognitiven Fähigkeiten abgegrenzt werden (Klos et al., 2008; Mannel, 2011). Für das Fach Biologie ergab eine Untersuchung von Hammann et al. (2007), dass das *Testen von Hypothesen* von den Dimensionen *Hypothesenbildung* und *Analyse von Evidenzen* psychometrisch unterschieden und somit die Existenz zweier Dimensionen belegt werden kann. Mayer et al. (2008) benennen als zentrale Kompetenzen naturwissenschaftlicher Erkenntnisgewinnung vier Prozessvariablen, die als eigenständige Teilkompetenzen beim Experimentieren empirisch identifiziert werden konnten und einen integrativen Bestandteil der Bildungsstandards zum Kompetenzbereich Erkenntnisgewinnung darstellen: (1) Naturwissenschaftliche Fragestellung formulieren, (2) Hypothesen generieren, (3) Untersuchungen planen sowie (4) Daten analysieren und Schlussfolgerungen ziehen. Zwischen den Teilkompetenzen liegen signifikante Leistungsunterschiede vor (Grube & Mayer, 2010; Walpuski & Schulz, 2011). Dabei fällt es Schülerinnen und Schülern leichter, Ergebnisse einer Untersuchung zu deuten als zu untersuchende naturwissenschaftliche Fragestellungen zu erkennen oder selbstständig zu formulieren (Grube, 2011; Prenzel et al.,

2007). Für jede Teilkompetenz lassen sich fünf Kompetenzniveaus beschreiben, die sich hinsichtlich Komplexität und qualitativer Abstufung eines Problemlöseprozesses unterscheiden und deren empirische Überprüfung für drei der fünf Kompetenzniveaus – *Untersuchung eines Faktors*, *Untersuchung von Zusammenhängen* und *Kontrollierte Untersuchung auf Basis von Konzeptverständnis* – die Annahme einer ansteigenden Schwierigkeit bestätigen (Mayer et al., 2008). Inwiefern die genannten Prozeduren nicht nur für das Experimentieren, sondern auch als eigenständige Konstrukte für weitere charakteristische Erkenntnismethoden in den Naturwissenschaften empirisch zu trennen sind, z. B. für das Beobachten und Vergleichen in der Biologie oder die theoretische Deduktion in der Physik, wird im Rahmen von Begleitstudien geprüft (Wellnitz, 2012; Wellnitz & Mayer, 2012).

2.2 Naturwissenschaftliche Modellbildung

Die Auseinandersetzung mit naturwissenschaftlichen Problemen schließt den Umgang mit Modellen auf verschiedenen Repräsentationsebenen ein. Dabei werden theoretische Konstrukte (z. B. Hypothesen, Vorstellungen, Teile von Theorien) und gegenständliche, vereinfachte Repräsentationen gedanklicher oder materieller Realität als Modelle bezeichnet (Kircher, 2010). Theoretische Modelle der Naturwissenschaften versuchen, wesentliche Parameter der meist natürlichen Phänomene zu erfassen. Sie folgen einer Logik und sie

sind überprüfbar. Gegenständliche Modelle umfassen sowohl dreidimensionale als auch zweidimensionale bildhafte Repräsentationen, die reale oder fiktive Beziehungen zum Ausgangsobjekt (Entität, Prozess, Idee) aufweisen (Kircher, 2010). Theoretische und gegenständliche Modelle sind das Ergebnis eines hypothesengeleiteten Modellierungsprozesses und erlauben die Ableitung weiterer naturwissenschaftlicher, zu prüfender Fragestellungen und Hypothesen (Upmeier zu Belzen & Krüger, 2010).

Das Denken in und mit Modellen und ein damit verbundenes Verständnis für den Modellcharakter naturwissenschaftlicher Theorien werden als grundlegend für eine adäquate Entwicklung erkenntnistheoretischer Sichtweisen angesehen (Mikelskis-Seifert & Fischler, 2003). Naturwissenschaftliche Modellkompetenz zeigt sich nach Upmeier zu Belzen und Krüger (2010) in der Fähigkeit, mit Modellen zweckbezogen Erkenntnisse gewinnen, über Modelle zweckbezogen urteilen und den Prozess der Erkenntnisgewinnung modellbezogen reflektieren zu können. Dazu gehört die Bereitschaft, alle Fähigkeiten zusammengenommen in problemhaltigen Situationen anzuwenden (Upmeier zu Belzen & Krüger, 2010). Die Autoren beschreiben zwei Dimensionen der Modellkompetenz, *Kenntnisse über Modelle* und *Modellbildung*, die in weitere Teilfähigkeiten mit verschiedenen Qualitäten ausdifferenziert werden. Die Dimension *Kenntnisse über Modelle* umfasst kognitive Konzepte über Modelle. Die Teilkompetenz *Eigenschaften von Modellen* drückt sich in einem Verständ-

nis für verschiedene Repräsentationsebenen zwischen Modell und Ausgangsobjekt aus. Die Abstraktion beginnt bei naturgetreuen Abbildern und erreicht den höchsten Abstraktionsgrad in theoretischen oder mathematischen Rekonstruktionen (Mikelskis-Seifert & Fischler, 2003). Unter der Teilkompetenz *Alternative Modelle* werden Vorstellungen über alternative Rekonstruktionen zum selben Ausgangsobjekt subsumiert, die als Resultat hypothesengeleiteter Herstellungsprozesse in Verbindung mit einer vom Subjekt individuell bestimmten Anwendungssituation verstanden werden (Upmeier zu Belzen & Krüger, 2010). Die Dimension *Modellbildung* umfasst den Umgang mit Modellen im Modellierungsprozess, d. h. kognitive Fähigkeiten, die eine Reflexion der Schritte im Modellbildungsprozess und ein Urteil über Modelle einschließen. Im *Zweck von Modellen* werden Modelle als zweckbezogene Rekonstruktionen und als Ausgangsobjekt für eine bestimmte Anwendung verstanden. Die Teilkompetenz *Testen von Modellen* benennt die Fähigkeit, die Eignung von Modellen zum Überprüfen von Fragestellungen und Hypothesen beurteilen zu können, während die Teilkompetenz *Ändern von Modellen* ein Verständnis für zu überprüfende Unzulänglichkeiten von Modellen einschließlich deren Revidierung verlangt (Upmeier zu Belzen & Krüger, 2010).

Bisherige Studien zeigen, dass Schülerinnen und Schüler die Rolle von theoretischen Modellen im naturwissenschaftlichen Erkenntnisprozess in der Regel nicht wahrnehmen. Als Modelle werden vorrangig gegenständliche Re-

präsentationen verstanden, von denen zusätzlich angenommen wird, dass sie Realität naturgetreu abbilden (Meisert, 2008; Terzer & Upmeyer zu Belzen, 2007).

2.3 Wissenschaftstheoretische Reflexion

Nach Bybee (1997) geht die höchste Stufe einer Scientific Literacy über naturwissenschaftliches Wissen und naturwissenschaftliche Prozeduren hinaus und schließt andere Ansichten über die Naturwissenschaften ein. Diese Metaebene beinhaltet unter anderem ein Verständnis der sogenannten Nature of Science (Natur der Naturwissenschaften, Wissenschaftsverständnis). Eine derartige Wertschätzung von Nature of Science als zentrales Ziel naturwissenschaftlichen Unterrichts zeigt sich insbesondere in der angloamerikanischen (Lederman, 2007; McComas & Olsen, 1998; Osborne, Collins, Ratcliffe, Millar & Duschl, 2003), aber auch in der deutschsprachigen naturwissenschaftsdidaktischen Literatur (Höttecke, 2001b; Kircher & Dittmer, 2004; Neumann, 2011; Urhahne, Kremer & Mayer, 2008).

Während in den amerikanischen Standards die Entwicklung eines adäquaten Verständnisses von Nature of Science eine starke Akzentuierung erfährt, ist dieses Themengebiet in den deutschen Bildungsstandards dagegen nur implizit berücksichtigt. In jeweils einen Standard pro Fach spielt auch ein Verständnis von Nature of Science hinein. So werden im Kompetenzbereich Erkenntnisgewinnung die Anforderungen formuliert, „die Gültigkeit empirischer Ergebnisse und deren Verallgemeinerung“ beurteilen (KMK, 2005c, S. 11), die „Tragweite und Grenzen von Untersuchungsanlage, -schritten und -ergebnissen“ erörtern (KMK, 2005a, S. 14) oder „exemplarisch Verknüpfungen zwischen gesellschaftlichen Entwicklungen und Erkenntnissen“ aufzeigen zu können (KMK, 2005b, S. 12). Dabei ist kritisch anzumerken, dass der zuletzt genannte Bildungsstandard aus dem Fach Chemie eine inhaltliche Nähe zum Kompetenzbereich Bewertung aus dem Fach Physik aufweist: Die Schülerinnen und Schüler „benennen Auswirkungen physikalischer Erkenntnisse in historischen und gesellschaftlichen Zusammenhängen“ (KMK, 2005c, S. 12).

Im Sinne einer Orientierung an internationalen Standards ist es sinnvoll, die Ausbildung eines adäquaten Verständnisses von Nature of Science als Bestandteil naturwissenschaftlicher Bildung mitzudenken (Neumann, 2011). Da die Evaluation der Bildungsstandards auch eine Grundlage für deren spätere Überarbeitungen bilden soll, wird dieser Teilbereich bewusst in das Kompetenzmodell aufgenommen, um so eine empirische Datenbasis zur Verfügung stellen zu können. Trotz der oft konstatierten verschiedenen wissenschaftsphilosophischen Ansichten über Nature of Science konnte unter Naturwissenschaftsdidaktikern ein weitgehender Konsens hinsichtlich der Einbettung als Thema im naturwissenschaftlichen Unterricht gefunden werden. So schlagen Lederman, Abd-El-Khalick, Bell und Schwartz (2002) einen Fokus auf die für Schülerinnen und Schüler relevanten Aspekte vor und benennen sieben Punkte: (a) die Vorläufigkeit und

Veränderlichkeit naturwissenschaftlichen Wissens, (b) der empirische und objektive Charakter der Wissensgenerierung durch Beobachtungen und Experimente, (c) die subjektive Herangehensweise an Prozesse der Erkenntnisgewinnung, (d) die Unterscheidung zwischen Beobachtungen und Interpretationen, (e) die Unterscheidung zwischen Gesetzmäßigkeiten und Theorien, (f) der Anteil von Kreativität und Vorstellungskraft und (g) der Einfluss von sozialen, kulturellen und technologischen Aspekten. Eine von Osborne et al. (2003) durchgeführte Delphi-Studie, die Experten aus den Naturwissenschaften, der Wissenschaftsphilosophie, -soziologie und Naturwissenschaftsdidaktik einschloss, kam zu vergleichbaren Aspekten, die drei Bereichen zugeordnet werden. *Nature of Scientific Knowledge* umfasst wie bei Ledermann et al. (2002) die Vorläufigkeit, die Veränderlichkeit und den empirischen Charakter des naturwissenschaftlichen Wissens. Der Bereich *Methods of Science* beinhaltet die Aspekte, dass naturwissenschaftliche Erkenntnisgewinnung von Kreativität und Vorstellungskraft geleitet wird, und dass je nach zugrundeliegender Problemstellung unterschiedliche Erkenntnismethoden zur Generierung von Erklärungszusammenhängen erforderlich sind. *Institutions and Social Practices in Science* beinhalten den Einfluss von ethischen, moralischen, technologischen und sozialen Komponenten auf das naturwissenschaftliche Wissen (Osborne et al., 2003).

Die Vorstellungen der Schülerinnen und Schüler von Nature of Science werden traditionell mit Fragebögen oder Interviews

erfasst (Lederman et al., 2002; Urhahne et al., 2008). Ausgehend von den Arbeiten von Lederman et al. (2002) stellte Neumann (2011) ein Kompetenzmodell für Nature of Science vor, das sich an der Grundstruktur am ESNaS-Modell orientiert (Kauertz, Fischer, Mayer, Sumfleth & Walpuski, 2010). Zur Validierung dieses Modells wurden Testaufgaben mit geschlossenem Antwortformat entwickelt und eingesetzt.

3 Das ESNaS-Kompetenzmodell zur naturwissenschaftlichen Erkenntnisgewinnung

Die Überführung der Bildungsstandards des Kompetenzbereichs Erkenntnisgewinnung in ein gemeinsames Kompetenzstrukturmodell aller drei naturwissenschaftlichen Fächer erfolgt unter zweifacher Zielsetzung: Zum einen soll in Analogie zu den Basiskonzepten des Kompetenzbereichs Fachwissen und unter Berücksichtigung curricularer Vorgaben den in den Standards benannten Anforderungen eine explizite, theoriebasierte Kompetenzstruktur zugrunde gelegt werden; zum anderen sollen strukturbezogene Kompetenzniveaus differenziert dargestellt werden, um spezifische Anforderungen für Lernende mit beispielsweise hoher oder niedriger Kompetenz formulieren und messen zu können.

Die normativen Kompetenzbeschreibungen der nationalen Bildungsstandards differenzieren die Kompetenzen naturwissenschaftlicher Erkenntnisgewinnung in drei Anforderungsbereiche, für deren Gra-

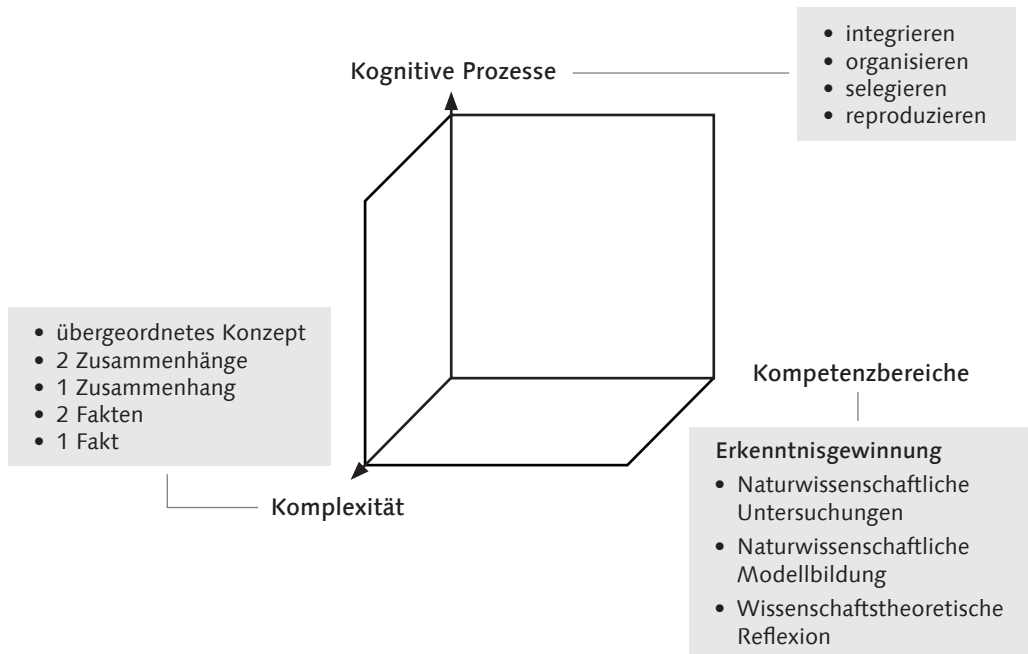


Abb. 1: Kompetenzmodell für den Kompetenzbereich Erkenntnisgewinnung.

duierung auf Kriterien der Einheitlichen Prüfungsanforderungen in der Abiturprüfung (EPA) zurückgegriffen wurde. Für eine Leistungsmessung ist die dort vorgenommene gleichzeitige Variation mehrerer Schwierigkeit erzeugender Merkmale jedoch ungeeignet. Ferner sind die Anforderungsbereiche explizit nicht hierarchisch gestuft (Kauertz et al., 2010). Aus diesem Grund wurden die in den Anforderungsbereichen enthaltenen Merkmale in die beiden Dimensionen *Komplexität* und *kognitive Prozesse* überführt, die es einerseits erlauben, postulierte Schwierigkeitsbestimmende Merkmale gezielt zu variieren und die andererseits den psychometrischen Qualitätsanforderungen der zu entwickelnden Testaufgaben genügen (Kauertz & Fischer, 2006). Eine dritte Di-

mension *Kompetenzbereiche* repräsentiert die Fachinhalte. Damit liegt eine dreidimensionale Kompetenzstruktur vor: zwei Dimensionen, die als schwierigkeits erzeugend angenommen und implizit in den Anforderungsbereichen der Bildungsstandards enthalten sind, sowie eine inhaltliche Dimension (Abb.1).

Das dreidimensionale Kompetenzmodell bildet die Grundlage für die Itemkonstruktion, indem die Dimensionen *Komplexität* und *kognitive Prozesse* durch gezielte Variation mit der Dimension *Erkenntnisgewinnung* in Beziehung gesetzt werden. Die nach diesem Modell entwickelten Testaufgaben werden fachspezifisch konstruiert, d. h. dass in den Aufgaben gemäß der Bildungsstandards biologische, chemische oder physikalische Phänomene,

Modelle etc. als zu bearbeitende Ausgangssituationen vorgegeben werden. Somit kann der Einfluss des Fachinhaltes auf das Zustandekommen der Testleistungen zur Erfassung wissenschaftsmethodischer Kompetenz überprüft werden. Die verschiedenen Kombinationsmöglichkeiten der drei Dimensionen erlauben, so die Annahme, Testaufgaben mit unterschiedlichen Schwierigkeitsgraden zu konstruieren, die mit zunehmendem Anspruch der Aufgabe eine höhere notwendige Kompetenzausprägung auf Seiten der Schülerinnen und Schüler zur erfolgreichen Bewältigung erwarten lassen.

3.1 Dimension kognitive Prozesse

Die in den Aufgaben vorgegebenen Informationen müssen zur Lösung einer Testaufgabe verarbeitet werden. Die kognitiven Prozesse – *Reproduzieren, Selegieren, Organisieren* und *Integrieren* – werden bei der erfolgreichen Bearbeitung der Aufgaben erwartet. Dabei wurden zur Erhöhung des Textverständnisses die drei Verständlichkeits-Maximen der Einfachheit, Eindeutigkeit und Einheitlichkeit bei der Itemkonstruktion berücksichtigt (Groeben, 1982; Langer, Schulz von Thun & Tausch, 2011). Die zur Verfügung gestellte Informationsdichte wurde zudem möglichst gering gehalten. Zur Lösung notwendige Fachinformationen sind im Aufgabenstamm enthalten. Damit soll zum einen dem fehlenden verbindlichen Inhaltskanon zwischen Schulformen und Bundesländern zumindest teilweise Rechnung getragen werden; zum ande-

ren soll auf diese Weise der Einfluss des Fachwissens kontrolliert werden. Die im Modell berücksichtigten kognitiven Prozesse werden im Kompetenzbereich Erkenntnisgewinnung genauso operationalisiert wie im Kompetenzbereich Fachwissen (für eine detaillierte Beschreibung der Operationalisierung siehe Kauertz et al., 2010).

3.2 Dimension Komplexität

Die Komplexitätsstufen beschreiben, welche Informationsmenge für die richtige Lösung einer Aufgabe von den Schülerinnen und Schülern verarbeitet werden muss (Commons et al., 2007; Kauertz et al., 2010). Ein Ansteigen der Komplexität geht in der Regel mit einer Zunahme von Lösungsschritten einher, die von der Nennung einzelner Elemente bis hin zur Verknüpfung dieser Elemente reicht. Mit nur geringen Einschränkungen lassen die verschiedenen Komplexitätsstufen Vorhersagen zur Schwierigkeit von Testaufgaben zu (Kauertz, 2008; Kauertz & Fischer, 2006; Mannel, 2011; Walpuski, Ropohl & Sumfleth, 2011; Wellnitz, 2012). Es werden fünf allgemeine Komplexitätsstufen beschrieben: *1 Fakt, 2 Fakten, 1 Zusammenhang, 2 Zusammenhänge* und *übergeordnetes Konzept*. Die Komplexitätsstufen werden fachspezifisch und wissenschaftsmethodisch ausdifferenziert und im Folgenden für die drei Kompetenzteilbereiche einzeln beschrieben.

3.3 Dimension Kompetenzbereiche

Die inhaltliche Dimension *Kompetenzbereiche* wird für den Kompetenzbereich Erkenntnisgewinnung analog zu den Basiskonzepten beim Fachwissen durch die drei Kompetenzteilbereiche *Naturwissenschaftliche Untersuchungen*, *Naturwissenschaftliche Modellbildung* und *Wissenschaftstheoretische Reflexion* beschrieben. Diese Aufteilung bietet die Möglichkeit, die präskriptiv formulierten Standards theorie- und evidenzbasierten Konstrukten zuzuordnen (Tab. 2).

Die abgeleiteten Teilbereiche bilden die Bildungsstandards der Erkenntnisgewinnung aller drei naturwissenschaftlichen Fächer systematisch ab und sind darüber hinaus anschlussfähig an internationale Standards und Assessments (DfES & QCA, 2004; Mullis, Martin, Ruddock, O'Sullivan & Preuschof, 2009; NRC, 1996). Die Teilbereiche werden wiederum durch verschiedene Kompetenzaspekte ausdifferenziert, die einerseits die in den Bildungsstandards benannten Anforderungen direkt als Teilfähigkeiten benennen und andererseits nationale und

Tab. 2: Rahmenkonzept nach Mayer (2007) (* Anzahl der Bildungsstandards nach Fächern getrennt; B: Biologie, C: Chemie, P: Physik; Tabelle im Anhang)

KMK-Standards (exemplarisch)	Anzahl*			Teilbereiche
	B	C	P	
Die Schülerinnen und Schüler...				
<ul style="list-style-type: none"> ...erkennen und entwickeln Fragestellungen, die [...] insbesondere durch chemische Experimente, zu beantworten sind. (C/E1) ...stellen an einfachen Beispielen Hypothesen auf. (P/E6) ...planen geeignete Untersuchungen zur Überprüfung von Vermutungen und Hypothesen. (C/E2) ...planen einfache Experimente, führen die Experimente durch und/oder werten sie aus. (B/E6) 	7	6	6	Naturwissenschaftliche Untersuchungen (Scientific Inquiry)
<ul style="list-style-type: none"> ...wenden Modelle zur Veranschaulichung [...] an. (B/E9) ...nutzen geeignete Modelle [...] um chemische Fragestellungen zu bearbeiten. (C/E7) ...verwenden Analogien und Modellvorstellungen zur Wissensgenerierung (P/E3). 	5	1	3	Naturwissenschaftliche Modellbildung (Scientific Modelling)
<ul style="list-style-type: none"> ...zeigen exemplarisch Verknüpfungen zwischen gesellschaftlichen Entwicklungen und Erkenntnissen der Chemie auf. (C/E8) ...erörtern Tragweite und Grenzen von Untersuchungsanlage, -schritten und -ergebnissen. (B/E8) ...beurteilen die Gültigkeit empirischer Ergebnisse und deren Verallgemeinerung. (P/E10) 	1	1	1	Wissenschaftstheoretische Reflexion (Nature of Science)

Erkenntnisgewinnung			
Teilbereiche	Naturwissenschaftliche Untersuchungen	Naturwissenschaftliche Modellbildung	Wissenschaftstheoretische Reflexion
Aspekte	Fragestellung	Funktionalität	Eigenschaften
	Hypothese	Modellanwendung	Entwicklung
	Untersuchungsdesign	Grenzen	
	Datenauswertung		

Abb. 2: Operationalisierung des Kompetenzbereichs Erkenntnisgewinnung.

internationale Forschungsbefunde integrieren (Abb. 2).

Die Anzahl der den Kompetenzteilbereichen zugewiesenen Bildungsstandards (Tab. 2, Tabelle im Anhang) repräsentiert die für jeden Bereich konstruierte Itemmenge. Das bedeutet, dass ca. 60 % der Testaufgaben dem Bereich *Naturwissenschaftliche Untersuchungen*, ca. 30 % dem Bereich *Naturwissenschaftliche Modellbildung* und ca. 10 % der Testaufgaben dem Bereich *Wissenschaftstheoretische Reflexion* angehören. Einige wenige, auf wissenschaftliche Arbeitstechniken oder praktische Fähigkeiten bezogene Standards, z. B. „Die Schülerinnen und Schüler mikroskopieren Zellen und stellen sie in einer Zeichnung dar“ (KMK, 2005a, S. 14), werden im Rahmen der Evaluation der Bildungsstandards nicht überprüft.

Teilbereich

Naturwissenschaftliche Untersuchungen

Im Kompetenzteilbereich *Naturwissenschaftliche Untersuchungen* werden in An-

lehnung an Mayer et al. (2008) und Klos et al. (2008) die wissenschaftsmethodischen Aspekte *Fragestellung*, *Hypothese*, *Untersuchungsdesign* und *Datenauswertung* des hypothetico-deduktiven Vorgehens subsumiert. Aufgaben zum Aspekt *Fragestellung* beschreiben im Stimulus ein naturwissenschaftliches Problem oder ein Untersuchungsdesign auf deren Basis die Schülerinnen und Schüler eine naturwissenschaftliche Frage, die zur Problemlösung beiträgt, erarbeiten müssen. Passende naturwissenschaftliche Fragen müssen erkannt und wiedergegeben (reproduzieren), ausgewählt (selegieren), abgeleitet (organisieren) oder eigenständig unter Berücksichtigung des im Stimulus zur Verfügung gestellten Fachwissens formuliert werden (integrieren). Aufgaben zum Aspekt *Hypothese* nehmen Bezug auf eine gegebene naturwissenschaftliche Fragestellung, eine theoretische oder modellhafte Annahme oder auf ein beschriebenes Untersuchungsdesign. Auf der Basis dieser Ausgangssituation haben die Schülerinnen und Schüler

Aufgabe zum Aspekt <i>Hypothese</i>		
Komplexitätsstufe: III / kognitiver Prozess: organisieren		
Fleißiges Lieschen		
Max will die Wachstumsbedingungen von Pflanzen untersuchen. Er verwendet als Versuchspflanze das „Fleißige Lieschen“, eine pflegeleichte und widerstandsfähige Zimmerpflanze.		
Er führt folgendes Experiment durch:		
	Pflanze 1	Pflanze 2
Temperatur	25 °C	25 °C
Ort	am hellen Fenster	im dunklen Schrank
Gießen	1 mal täglich	1 mal täglich
Düngen	2 mal wöchentlich	2 mal wöchentlich
Nenne die Hypothese (Vermutung), die dem Experiment von Max zugrunde liegt.		

Abb. 3: Aufgabenbeispiel für das Fach Biologie: Kompetenzteilbereich *Naturwissenschaftliche Untersuchungen*.

eine Hypothese zu erkennen (reproduzieren oder selektieren), abzuleiten (organisieren) beziehungsweise unter Darbietung komplexer Untersuchungsdesigns zu formulieren und ggf. zu beurteilen (integrieren) (Abb. 3).

Unter dem *Aspekt Untersuchungsdesign* wird sowohl die Planung als auch die Durchführung einer Untersuchung gefasst. Naturwissenschaftliche Erkenntnisgewinnung ist in der Regel an die praktische Durchführung von Untersuchungen zur Datengewinnung gebunden. Im Rahmen der Evaluation der Bildungsstandards werden die den praktischen Durchführungen zugrunde liegenden Kompetenzen nicht erhoben. Dies ist testökonomischen Gründen geschuldet, da praktische Tests sehr aufwendig beziehungsweise kostspielig und in der Regel weniger reliabel sind (Stecher & Klein, 1997). Inwieweit mit schriftlicher

Leistungsmessung oder virtuellen Experimentierumgebungen praktische Fähigkeiten erfasst werden können, konnte in bisherigen Studien noch nicht eindeutig geklärt werden (Emden, 2011; Hammann, Phan, Ehmer & Grimm, 2008; Schreiber, Theyßen & Schecker, 2009; Shavelson et al., 1991). Die in den Bildungsstandards beschriebenen praktischen Kompetenzen werden deshalb zunächst als prozedurales Wissen („Wissen, wie“) interpretiert und entsprechend getestet. Das bedeutet, dass in der Aufgabenstellung zwar keine praktische Umsetzung gefordert wird, sie kann jedoch beschrieben und von den Schülerinnen und Schülern bezüglich der Angemessenheit beurteilt werden. Aufgaben zum Aspekt Untersuchungsdesign nehmen dementsprechend Bezug auf eine gegebene Hypothese, die empirisch überprüft werden soll, oder auf ein zu erläuterndes Untersuchungsdesign. Beim

Reproduzieren und Selegieren wird beispielsweise von den Schülerinnen und Schülern die Wiedergabe oder die Auswahl von Variablen, beim Organisieren die Ableitung eines passenden Untersuchungsdesigns und beim Integrieren die selbstständige Beschreibung eines experimentellen Designs oder eines Beobachtungsdesigns verlangt. Aufgaben zum Aspekt Datenauswertung nehmen Bezug auf vorgegebene Ergebnisse einer Untersuchung, deren Daten beispielsweise hypothesenbezogen und fehlerbezogen ausgewertet und interpretiert werden sollen. Die Ergebnisse einer Untersuchung werden wiedergegeben (reproduzieren), ausgewählt (selegieren), eigene Schlussfolgerungen abgeleitet (organisieren) oder die Ergebnisse einer Untersuchung angewendet, indem diese auf eine andere Situationen übertragen werden (integrieren).

Für die vier Aspekte werden jeweils fünf qualitative Abstufungen beschrieben. Diese graduieren die Komplexität des wissenschaftsmethodischen Anspruchs (Tab. 3).

Die fünf Komplexitätsstufen gehen auf ein Kompetenzmodell zur Stufung der Schwierigkeit von Aufgaben im Kompetenzbereich Erkenntnisgewinnung (Mayer et al., 2008) und einem analogen Modell für den Kompetenzbereich Fachwissen (Kauertz, 2010) zurück. Beide Modelle basieren auf einer Zunahme relevanter Elemente des jeweiligen Kompetenzbereichs und deren Verknüpfungen.

Teilbereich

Naturwissenschaftliche Modellbildung

Im Kompetenzteilbereich *Naturwissenschaftliche Modellbildung* werden, in Anlehnung an Upmeyer zu Belzen und Krüger (2010), die drei Aspekte *Funktionalität*, *Modellanwendung* und *Grenzen* ausdifferenziert. Unter dem Aspekt Funktionalität wird ein Verständnis für die Funktion von Modellen und damit für den inhaltlichen Zweck der Modellbildung gefordert. Aufgaben zu diesem Aspekt beschreiben beispielsweise im Stimulus ein Modell oder mehrere Modelle, die hinsichtlich ihrer Funktion beurteilt werden sollen (Abb. 4). Der Aspekt Modellanwendung umfasst das Prüfen von gegenständlichen oder theoretischen Modellen. Modelle sollen nicht vorrangig als Repräsentanten materieller oder gedanklicher Realität sondern ihre Bedeutung im Forschungsprozess selbst verstanden werden. Ein Modell kann Gegenstand einer Testung sein und zur Überprüfung von Hypothesen zur Generierung neuer Erkenntnisse herangezogen werden. Im Aufgabenstamm wird beispielsweise eine Untersuchung beschrieben, der ein geeignetes Modell zur Überprüfung von Hypothesen zugewiesen werden muss. Der Aspekt Grenzen verlangt ein Verständnis für verschiedene Arten der Repräsentation von Modellen, deren höchster Abstraktionsgrad in theoretischen Modellen als symbolische Darstellung gegeben ist. Alternative Repräsentanten zum selben Ausgangsobjekt beziehungsweise zur selben Ausgangsidee sollen als mögliche Rekonstruktionen kognitiv verarbeitet werden. Im Aufgabenstamm werden beispielsweise


Tab. 3: Theoretisch formulierte Komplexitätsstufen für die Aspekte **(F)** Fragestellung, **(H)** Hypothese, **(U)** Untersuchungsdesign und **(D)** Datenauswertung

I	1 Fakt	<p>F einfache naturwissenschaftliche Frage nach einer Variablen oder einer Eigenschaft</p> <p>H einfache, testbare Hypothese unter Einbeziehung einer Variablen oder einer Eigenschaft</p> <p>U eine Variable, eine Eigenschaft, eine Veränderung eines Systems oder einen Unterschied/eine Gemeinsamkeit zweier Systeme</p> <p>D ein Datum oder ein Ergebnis</p>
II	2 Fakten	<p>F einfache naturwissenschaftliche Frage nach zwei Variablen oder zwei Eigenschaften</p> <p>H einfache, testbare Hypothese unter Einbeziehung von zwei Variablen oder zwei Eigenschaften</p> <p>U zwei Variablen, zwei Eigenschaften, zwei Veränderungen eines Systems oder zwei Unterschiede/Gemeinsamkeiten zweier Systeme</p> <p>D zwei Daten oder zwei Ergebnisse</p>
III	1 Zusammenhang	<p>F naturwissenschaftliche Frage zu einem qualitativen oder quantitativen Zusammenhang</p> <p>H einfache Zusammenhangshypothese oder Unterschiedshypothese</p> <p>U zu verändernde und zu messende Variable, ein System mit seinen Merkmalen und Veränderungen oder mit einem anderen System in Beziehung setzen</p> <p>D ein Zusammenhang zwischen Rohdaten oder eine Schlussfolgerung aus den Daten</p>
IV	2 Zusammenhänge	<p>F naturwissenschaftliche Frage zu zwei Zusammenhängen oder zwei naturwissenschaftliche Fragen</p> <p>H zwei verbundene, unverbundene Hypothesen oder Null- und Alternativhypothese</p> <p>U Kontrollvariablen, Merkmalsausprägungen oder mehrere Vergleichskriterien zusätzlich berücksichtigen</p> <p>D zwei Zusammenhänge zwischen Rohdaten oder zwei Schlussfolgerungen aus den Daten</p>
V	übergeordnetes Konzept	<p>F Eigenschaften naturwissenschaftlicher Fragestellungen (Nachvollziehbarkeit, Beitrag zur Problemlösung)</p> <p>H Eigenschaften naturwissenschaftlicher Hypothesen (empirische Überprüfbarkeit, Falsifizierbarkeit, theoriegeleitet)</p> <p>U Eigenschaften naturwissenschaftlicher Untersuchungen (Notwendigkeit von Messwiederholung, Versuchsdauer, Stichprobengröße, Genauigkeit, Variablenkontrolle, Kriterienstetigkeit, Merkmalsvariation)</p> <p>D Eigenschaften naturwissenschaftlicher Datenauswertungen (theorie- und fehlerbezogene Evaluation der Daten, alternative Deutungen, Sicherheit und Reichweite)</p>

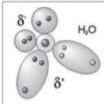
Aufgabe zum Aspekt *Funktionalität*
Komplexitätsstufe: V / kognitiver Prozess: integrieren

Modelle des Wassermoleküls

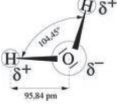
Jutta sucht im Internet nach einem Modell für das Wassermolekül und findet folgende Abbildungen:




www.wikipedia.de



www.atommodell.de



www.uni-münster.de



www.uni-bonn.de

Sie wundert sich, dass die Abbildungen so verschieden aussehen. Welche Erklärung dafür ist richtig? Kreuze an:

- Nur eines der Modelle ist das richtige Modell für das Wassermolekül, die anderen enthalten Fehler.
- Die Wissenschaftler sind noch gar nicht sicher, wie das Wassermolekül genau aufgebaut ist, daher gibt es verschiedene Modelle.
- Die Modelle zeigen verschiedene Arten von Wassermolekülen, daher sehen sie alle unterschiedlich aus.
- Die Modelle auf den Bildern sollen verschiedene Eigenschaften des Wassermoleküls verdeutlichen und unterscheiden sich daher.

Abb. 4: Aufgabenbeispiel für das Fach Chemie: Kompetenzteilbereich *Naturwissenschaftliche Modellbildung*.

gegenständliche oder theoretische Modelle vorgegeben, die hinsichtlich ihrer Idealisierungen oder Analogien im Vergleich zur Ausgangsidee hinterfragt werden müssen. Tabelle 4 verdeutlicht die Operationalisierung der zugehörigen Komplexitätsstufen.

Teilbereich *Wissenschaftstheoretische Reflexion*

Im Kompetenzteilbereich *Wissenschaftstheoretische Reflexion* wird im Gegensatz zu den anderen beiden Teilbereichen nicht die konkrete Ebene problemlösender Handlungsweisen zur Modell- und Theoriebildung abgebildet, sondern die Kompetenz zur Reflexion über den na-

turwissenschaftlichen Erkenntnisprozess und der Charakteristika der Naturwissenschaften als wissenschaftliche Disziplin. Es werden zwei Aspekte modelliert: *Entwicklung naturwissenschaftlichen Wissens* und *Eigenschaften naturwissenschaftlichen Wissens*. Zur Operationalisierung der Inhalte beider Aspekte werden die von Lederman (2007) vorgeschlagenen Eingrenzungen auf die im Schüleralltag relevanten Gesichtspunkte herangezogen und zu Schwerpunkten zusammengefasst (Neumann, 2011; Zilker, Kauertz & Fischer, 2009). Der *Aspekt Entwicklung naturwissenschaftlichen Wissens* bezieht sich auf Eigenschaften naturwissenschaftlicher Erkenntnisgewinnungsprozesse (Under-

Tab. 4: Theoretisch formulierte Komplexitätsstufen für die Aspekte (**F**) Funktionalität, (**M**) Modellanwendung und (**G**) Grenzen

I	1 Fakt	<p>F eine (Anschauungs-) Funktion eines Modells</p> <p>M eine Eigenschaft eines Modells</p> <p>G ein Unterschied zum Original</p>
II	2 Fakten	<p>F zwei (Anschauungs-) Funktionen eines Modells</p> <p>M zwei Eigenschaften eines Modells</p> <p>G ein Unterschied und eine Gemeinsamkeit zum Original</p>
III	1 Zusammen- hang	<p>F Bezug zwischen Modell und Verwendung</p> <p>M Modell, z. B. für Hypothese oder Schlussfolgerung</p> <p>G Grund für Modellgrenzen</p>
IV	2 Zusammen- hänge	<p>F Funktionen und Verwendung verschiedener Modelle</p> <p>M mehrere Modelle, z. B. für Hypothese oder Schlussfolgerung oder Modell für verschiedene Hypothesen oder Schlussfolgerungen</p> <p>G Begrenzungen des Modells und Optimierungsmöglichkeiten</p>
V	übergeordnetes Konzept	<p>F Eigenschaften des Modellierungsprozesses (Modellvorstellung, Problemlösung)</p> <p>M Testen von Modellen (alternative Rekonstruktionen, Eignung)</p> <p>G Modellkritik (Begrenzung und Vorläufigkeit, Optimierungen und Veränderungen, Entsprechungen, Verkürzungen und Beiwerk)</p>

standing about Scientific Inquiry) und gliedert sich in drei Gesichtspunkte: (a) Verfolgen einer naturwissenschaftlichen Fragestellung, (b) Nutzung vielfältiger Methoden und (c) Mannigfaltigkeit der Interpretation von Daten. Diese Gesichtspunkte beinhalten ein grundlegendes Verständnis darüber, dass naturwissenschaftliche Fragestellungen den Ursprung von naturwissenschaftlichen Untersuchungen bilden und diese beeinflussen, dass es nicht „die eine Methode“ gibt, die zu naturwissenschaftlicher Erkenntnis führt und dass gewonnene Daten verschiedenartig, z. B. in

Abhängigkeit vom Vorwissen oder dem vorherrschenden Paradigma, interpretiert werden können.

Der *Aspekt Eigenschaften naturwissenschaftlichen Wissens* bezieht sich auf das Produkt naturwissenschaftlicher Erkenntnisgewinnung und wird ebenfalls durch drei Gesichtspunkte operationalisiert: (a) Subjektivität naturwissenschaftlichen Wissens, (b) Daten und Interpretation als Grundlage naturwissenschaftlichen Wissens und (c) Vorläufigkeit naturwissenschaftlichen Wissens. Damit umfasst dieser Aspekt ein grundlegendes Verständ-

nis darüber, dass naturwissenschaftliches Wissen nicht objektiv ist, sondern durch Persönlichkeitsmerkmale eines Naturwissenschaftlers, wie beispielsweise durch seine Kreativität und durch soziale Faktoren, wie beispielsweise durch das vorherrschende Weltbild, beeinflusst wird, ohne dabei jedoch beliebig zu sein. Dass empirische Daten erst durch ihre Interpretation zu Wissen werden und gleichermaßen ein Gültigkeitskriterium für naturwissenschaftliche Theorien und Modelle darstel-

len, sowie dass naturwissenschaftliches Wissen nicht absolut wahr ist, sondern Veränderungen, z. B. bedingt durch neu entdeckte Phänomene, verbesserte Messinstrumente oder neue Theorien, unterliegt. Als Aufgabenkontexte werden fächerübergreifend für beide Aspekte zum Teil historische Fallbeispiele verwendet, da diese im Unterricht zur Thematisierung von Nature of Science als geeignet angesehen werden (Höttecke, 2001a; McComas, 2008) (Abb. 5).

Aufgabe zum Aspekt *Eigenschaften naturwissenschaftlichen Wissens*

Komplexitätsstufe: V / kognitiver Prozess: integrieren

Das Universum

Zu Beginn des 20. Jahrhunderts gab es unterschiedliche naturwissenschaftliche Modelle, die das Universum beschrieben. Albert Einstein und Willem de Sitter stellten sich vor, dass das Universum statisch ist und in sich ruht. Georges Lemaître stellte sich vor, dass das Universum expandiert und mit der Zeit größer wird. Von diesem Modell leitete Lemaître ein Gesetz für die Bewegung von benachbarten Galaxien her. Unabhängig davon machte der Physiker Edwin Hubble Beobachtungen und Berechnungen, aus denen er das gleiche Gesetz herleitete. Die Kombination der Arbeiten von Lemaître und Hubble ergab eine Vorstellung von der Entstehung und der Entwicklung des Weltalls. Damit konnten Hubbles Beobachtungen erklärt werden und sie ist heute noch anerkannt.

Inwiefern förderte Hubbles und/oder Lemaîtres Arbeit die Weiterentwicklung des naturwissenschaftlichen Wissens? Kreuze an.

- Weil in den Naturwissenschaften nur Daten zu neuem Wissen führen, trug nur Hubbles Arbeit zu einer Weiterentwicklung bei. Obwohl Lemaître das gleiche Gesetz entdeckte, war seine Arbeit nicht wichtig.
- Weder Lemaîtres noch Hubbles Arbeit trugen zu einer Weiterentwicklung bei, weil es schon zwei Modelle gab, an die viele Naturwissenschaftler glaubten. In den Naturwissenschaften müssen Gesetze und Beobachtungen zu bewährten Theorien passen.
- Weil in den Naturwissenschaften Beobachtungen nur durch Modelle erklärt werden können, trug nur Lemaîtres Arbeit zum Fortschritt in den Naturwissenschaften bei. Hubbles Beobachtungen waren nicht nötig, um Lemaîtres Arbeit zu unterstützen.
- Lemaîtres Theorie wurde durch Hubbles Beobachtungen und Berechnungen untermauert. In den Naturwissenschaften werden Theorien für die Erklärung von Beobachtungen gebraucht und sie müssen von Daten untermauert sein.

Abb. 5: Aufgabenbeispiel für das Fach Physik: Kompetenzteilbereich *Wissenschaftstheoretische Reflexion* nach Neumann (2011). Modellkonformes Item.

Tab. 5: Theoretisch formulierte Komplexitätsstufen für die Aspekte (**Ent**) Entwicklung naturwissenschaftlichen Wissens und (**Eig**) Eigenschaften naturwissenschaftlichen Wissens

I	1 Fakt	<p>Ent Beschreibung eines einfachen Sachverhaltes, Zuschreibung einer Eigenschaft, Begriffs o.ä., die sich auf den Prozess naturwissenschaftlicher Erkenntnisgewinnung beziehen</p> <p>Eig Beschreibung eines einfachen Sachverhaltes, Zuschreibung einer Eigenschaft, Begriffs o.ä., die sich auf naturwissenschaftliches Wissen beziehen</p>
II	2 Fakten	<p>Ent zwei Beschreibungen einfacher Sachverhalte, Zuschreibungen von Eigenschaften, Begriffen o.ä., die sich auf den Prozess naturwissenschaftlicher Erkenntnisgewinnung beziehen</p> <p>Eig zwei Beschreibungen einfacher Sachverhalte, Zuschreibungen von Eigenschaften, Begriffen o.ä., die sich auf naturwissenschaftliches Wissen beziehen</p>
III	1 Zusammenhänge	<p>Ent explizite Abhängigkeit, gegenseitige Bedingung, Relation, Korrelation, erklärte Kausalität oder Wechselwirkung, die sich auf den Prozess naturwissenschaftlicher Erkenntnisgewinnung beziehen</p> <p>Eig explizite Abhängigkeit, gegenseitige Bedingung, Relation, Korrelation, erklärte Kausalität oder Wechselwirkung, die sich auf Eigenschaften naturwissenschaftlichen Wissens beziehen</p>
IV	2 Zusammenhänge	<p>Ent zwei explizite Abhängigkeiten, gegenseitige Bedingungen, Relationen, Korrelationen, erklärte Kausalitäten oder Wechselwirkungen, die sich auf den Prozess naturwissenschaftlicher Erkenntnisgewinnung beziehen</p> <p>Eig zwei explizite Abhängigkeiten, gegenseitige Bedingungen, Relationen, Korrelationen, erklärte Kausalitäten oder Wechselwirkungen, die sich auf Eigenschaften naturwissenschaftlichen Wissens beziehen</p>
V	übergeordnetes Konzept	<p>Ent Anwendung eines reflektierenden, konzeptuellen Verständnisses über die Eigenschaften naturwissenschaftlicher Erkenntnisgewinnung (bezogen auf die drei Aspekte Verfolgen einer naturwissenschaftlichen Fragestellung, Nutzung vielfältiger Methoden, Mannigfaltigkeit der Interpretation von Daten)</p> <p>Eig Anwendung eines reflektierenden, konzeptuellen Verständnisses über die Eigenschaften naturwissenschaftlichen Wissens (bezogen auf die drei Aspekte Subjektivität, Daten und Interpretation als Grundlage naturwissenschaftlichen Wissens, Vorläufigkeit)</p>

Zusammenfassend ergeben sich die in Tabelle 5 dargelegten Operationalisierungen.

Bezogen auf die kognitiven Prozesse bedeutet das für die Schülerinnen und Schüler, dass die Eigenschaften des natur-

wissenschaftlichen Erkenntnisprozesses oder Produkte naturwissenschaftlicher Erkenntnisgewinnung wiedergegeben (reproduzieren), ausgewählt (selegieren), in Beziehung gesetzt (organisieren) oder angewendet werden (integrieren.)

4 Pilotierungsstudie

Die von abgeordneten Lehrkräften, sogenannten Aufgabenentwicklern, mehrschrittig konstruierten Testaufgaben durchliefen unter Koordination durch das IQB das im Projekt übliche Verfahren zur Qualitätssicherung (Walpuski et al., 2008; 2010). Im Anschluss an den umfangreichen Bewertungsprozess wurden die entwickelten Testitems hinsichtlich der zuvor festgelegten Kriterien überprüft, anhand kleinerer Schülerstichproben mehrmals präpilotiert und wiederum überarbeitet. Abschließend wurden die als geeignet eingestuft Items in den Gesamtpool der Testaufgaben aufgenommen und pilotiert. Insgesamt wurden zu den oben beschriebenen Kompetenzteilbereichen für alle naturwissenschaftlichen Fächer 492 Testaufgaben im offenen, halb-offenen und geschlossenen Aufgaben-

format entwickelt, die vom IQB in acht Bundesländern (Brandenburg, Bremen, Hamburg, Mecklenburg-Vorpommern, Niedersachsen, Rheinland-Pfalz, Saarland und Sachsen-Anhalt) an insgesamt 160 Schulen mit 6854 Schülerinnen und Schülern der 10. Jahrgangsstufe erprobt wurden (Tab. 6, 7 und 8).

Für die Testadministration wurde vom IQB ein komplexes Spiral-Design entwickelt, welches aus einer Kombination von 20-minütigen Aufgabenblöcken besteht und die verschiedenen Testhefte systematisch zueinander in Beziehung setzt. Die Auswertung erfolgte am IQB und in den Fachdidaktiken mit Methoden der klassischen und probabilistischen Testtheorie. Ziel der Pilotierung war die Analyse aller Items, auf deren Basis die Auswahl und Überarbeitung aller Aufgaben für die bundesweit repräsentative Normierungsstudie 2011 erfolgte.

Tab. 6: Verteilung der Items aller Fächer auf die Komplexitätsstufen und kognitiven Prozesse (rep: reproduzieren, sel: selektieren, org: organisieren, int: integrieren)

	alle Fächer					Biologie					Chemie					Physik				
	rep	sel	org	int	Σ	rep	sel	org	int	Σ	rep	sel	org	int	Σ	rep	sel	org	int	Σ
I	37	33	–	–	70	16	7	–	–	23	8	10	–	–	18	13	16	–	–	29
II	4	16	14	–	34	2	6	1	–	9	0	2	2	–	4	2	8	11	–	21
III	31	53	65	84	233	19	19	21	31	90	4	27	10	24	65	8	7	34	29	78
IV	9	16	35	40	100	4	4	11	21	40	0	10	10	12	32	5	2	14	7	28
V	2	1	14	38	55	1	1	5	20	27	0	0	0	1	1	1	0	9	17	27
Σ	83	119	128	162	492	42	37	38	72	189	12	49	22	37	120	29	33	68	53	183

Tab. 7: Verteilung der Items aller Fächer auf die Kompetenzteilbereiche und kognitiven Prozesse (NU: Naturwissenschaftliche Untersuchungen, F: Fragestellung, H: Hypothese, U: Untersuchungsdesign, D: Datenauswertung; NM.: Naturwissenschaftliche Modellbildung, F: Funktionalität, M: Modellanwendung, G: Grenzen; WR: Wissenschaftstheoretische Reflexion, Eig: Eigenschaften, Ent: Entwicklung)

		alle Fächer					Biologie					Chemie					Physik				
		rep	sel	org	int	Σ	rep	sel	org	int	Σ	rep	sel	org	int	Σ	rep	sel	org	int	Σ
NU	F	11	15	9	6	41	7	8	6	4	25	3	5	0	1	9	1	2	3	1	7
	H	6	11	10	5	32	5	8	5	5	23	0	3	1	0	4	1	0	4	0	5
	U	12	44	29	40	125	3	10	4	21	38	5	21	7	10	43	4	13	18	9	44
	D	8	14	45	35	102	4	5	18	10	37	2	9	9	10	30	2	0	18	15	35
NM	F	8	8	3	10	29	8	1	0	10	19	0	4	0	0	4	0	3	3	0	6
	M	16	11	19	32	78	3	2	0	3	8	2	3	5	13	23	11	6	14	16	47
	G	3	4	1	9	17	3	1	0	7	11	0	3	0	2	5	0	0	1	0	1
WR	Eig	12	6	6	10	34	5	1	2	4	12	0	0	0	0	0	7	5	4	6	22
	Ent	7	6	6	15	34	4	1	3	8	16	0	1	0	1	2	3	4	3	6	16
Σ		83	119	128	162	492	42	37	38	72	189	12	49	22	37	120	29	33	68	53	183

Tab. 8: Verteilung der Items aller Fächer auf die Kompetenzteilbereiche und Komplexitätsstufen

		alle Fächer						Biologie						Chemie						Physik					
		I	II	III	IV	V	Σ	I	II	III	IV	V	Σ	I	II	III	IV	V	Σ	I	II	III	IV	V	Σ
NU	F	6	0	33	2	0	41	1	0	22	2	0	25	4	0	5	0	0	9	1	0	6	0	0	7
	H	5	1	21	5	0	32	4	0	15	4	0	23	1	0	3	0	0	4	0	1	3	1	0	5
	U	15	17	44	30	19	125	5	4	10	8	11	38	6	2	19	16	0	43	4	11	15	6	8	44
	D	4	4	53	33	8	102	0	3	15	16	3	37	2	1	19	8	0	30	2	0	19	9	5	35
NM	F	9	1	13	2	4	29	6	0	8	1	4	19	1	0	3	0	0	4	2	1	2	1	0	6
	M	12	5	41	17	3	78	0	0	4	4	0	8	2	1	12	8	0	23	10	4	25	5	3	47
	G	2	1	8	2	4	17	0	1	6	2	2	11	2	0	2	0	1	5	0	0	0	0	1	1
WR	Eig	12	1	8	3	10	34	6	0	3	0	3	12	0	0	0	0	0	0	6	1	5	3	7	22
	Ent	5	4	12	6	7	34	1	1	7	3	4	16	0	0	2	0	0	2	4	3	3	3	3	16
Σ		70	34	233	100	55	492	23	9	90	40	27	189	18	4	65	32	1	120	29	21	78	28	27	183

Für die Itemanalyse wurden, nach Fächern getrennt, eindimensionale Raschskalierungen anhand von Teilstichproben durchgeführt und zur Beurteilung der Modellpassung einzelne Modellparameter (Item-Fits), die gewichteten Abweichungsquadrate (MNSQ) und die zugehörigen T-Werte herangezogen. Die Entscheidung über Elimination oder Verbleib von Items im Gesamtpool wurde außerdem unter Berücksichtigung der klassischen Trennschärfe und der inhaltlichen Begutachtung derjenigen Items vollzogen, die besonders hohe beziehungsweise besonders niedrige Itemschwierigkeiten aufwiesen.

Die IRT-basierte EAP/PV-Reliabilität als Skalenhomogenitätsmaß der Personenparameterschätzung (Wu, Adams, Wilson & Haldane, 2007) kann zum jetzigen Zeitpunkt für alle eindimensionalen Skalierungen als ausreichend angesehen werden (Biologie: $Rel_{EAP/PV} = .58$; Chemie: $Rel_{EAP/PV} = .67$; Physik: $Rel_{EAP/PV} = .63$). Der mit einer einfaktoriellem ANOVA durchgeführte Vergleich aller Gruppennittelwerte der Itemschwierigkeiten zeigt für die fünf Komplexitätsstufen (Biologie: $F(4, 184) = 6.79, p < .001, \eta_p^2 = .13$; Chemie: $F(4, 115) = 3.61, p < .01, \eta_p^2 = .11$; Physik: $F(4, 178) = 9.48, p < .001, \eta_p^2 = .18$) und die vier kognitiven Prozesse (Biologie: $F(3, 185) = 7.52, p < .001, \eta_p^2 = .11$; Chemie: $F(3, 116) = 5.83, p < .01, \eta_p^2 = .13$; Physik: $F(3, 179) = 6.12, p < .01, \eta_p^2 = .09$), dass zwischen den Stufen beider Dimensionen signifikante Unterschiede mit mittleren bis starken Effekten in den Aufgabenschwierigkeiten vorliegen. Lineare Trendanalysen bestätigen, dass mit zunehmender Komple-

xitätsstufe (Biologie: $F(1, 184) = 21.86, p < .001, \eta_p^2 = .10$; Physik: $F(1, 178) = 34.72, p < .001, \eta_p^2 = .16$) die Itemschwierigkeiten in der Grundgesamtheit mit einer Ausnahme (Chemie: $F(1, 115) = 1.81, p = .18, \eta_p^2 = .01$) signifikant linear ansteigen. Signifikante lineare Trends mit mittleren bis starken Effekten liegen ebenfalls für die kognitiven Prozesse vor (Biologie: $F(1, 185) = 18.81, p < .001, \eta_p^2 = .09$; Chemie: $F(1, 116) = 16.80, p < .001, \eta_p^2 = .13$; Physik: $F(1, 179) = 16.74, p < .001, \eta_p^2 = .08$). Die Abbildungen 6, 7 und 8 verdeutlichen die Zusammenhänge zwischen den Aufgabenmerkmalen beider Dimensionen und den Itemschwierigkeiten.

Die Abbildungen zeigen, dass der Anstieg der mittleren Itemschwierigkeiten nicht für alle Fächer durchgängig von einem Komplexitätsniveau beziehungsweise von einem kognitiven Prozess zum nächsten verläuft. Während die Itemschwierigkeiten von *1 Fakt* (Chemie: $M = -1.37, SD = 1.19$; Physik: $M = -1.57, SD = 1.34$) zu *2 Fakten* (Chemie: $M = -1.25, SD = 1.72$; Physik: $M = -.64, SD = .85$) beziehungsweise vom *Reproduzieren* (Chemie: $M = -1.69, SD = .91$; Physik: $M = -1.16, SD = 1.60$) zum *Selegieren* (Chemie: $M = -.74, SD = 1.29$; Physik: $M = -.98, SD = .98$) für chemische und physikalische Testaufgaben modellkonform ansteigen, sind die biologischen Testaufgaben zu *1 Fakt* ($M = -.62, SD = 1.29$) beziehungsweise zum *Reproduzieren* ($M = -.58, SD = 1.30$) im Mittel für die Schülerinnen und Schüler etwas schwerer zu lösen als Testaufgaben zu *2 Fakten* ($M = -1.01, SD = 1.72$) oder zum *Selegieren* ($M = -.67, SD = 1.78$). Als Ergebnis der Pilotierungsstudie kann

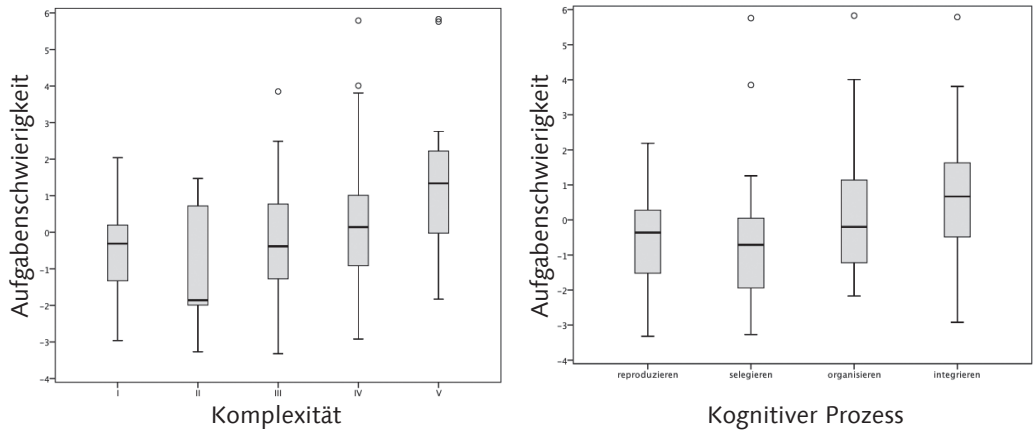


Abb. 6: Zusammenhänge zwischen den Aufgabenmerkmalen Komplexität / kognitiver Prozess und der Aufgabenschwierigkeit im Fach Biologie.

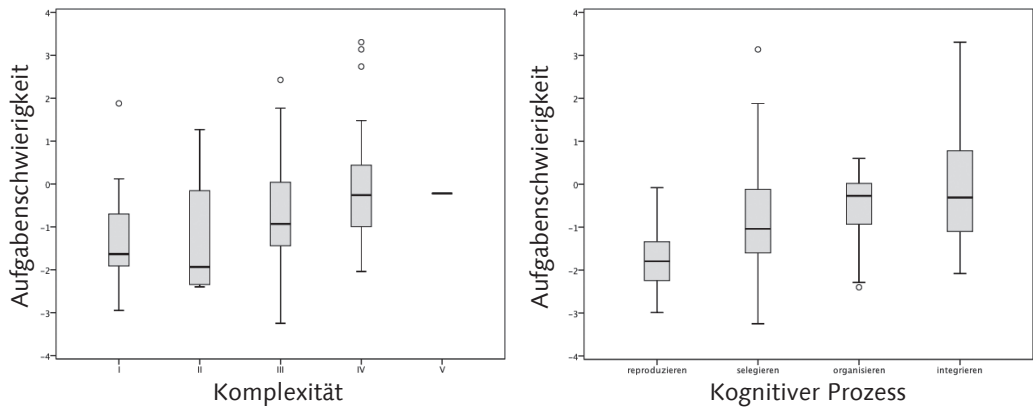


Abb. 7: Zusammenhänge zwischen den Aufgabenmerkmalen Komplexität / kognitiver Prozess und der Aufgabenschwierigkeit im Fach Chemie.

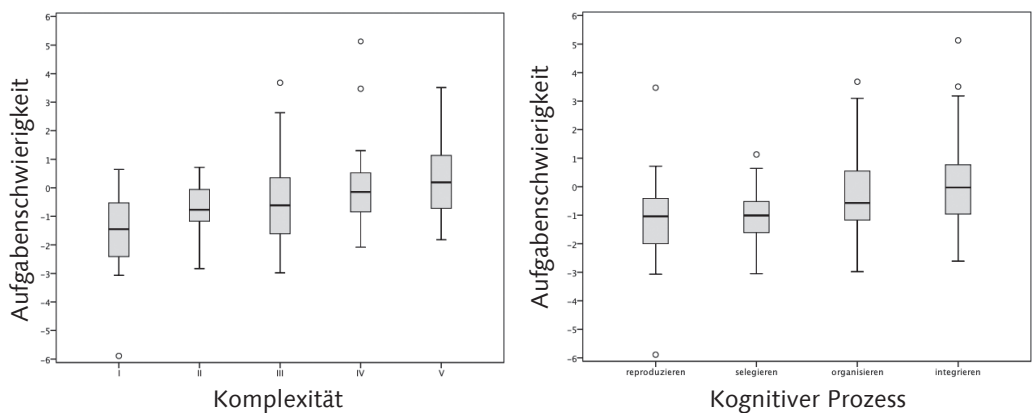


Abb. 8: Zusammenhänge zwischen den Aufgabenmerkmalen Komplexität / kognitiver Prozess und der Aufgabenschwierigkeit im Fach Physik.

ten 506 Testitems für die Normierung 2011 bereitgestellt werden. Die höhere Zahl der Items ergibt sich aus der Anzahl der Items, die die Kriterien der Pilotierung erfüllt haben zuzüglich einer Anzahl von nachentwickelten Items mit besonders hohen Anforderungen. Insgesamt konnten aus der Biologie 89 %, aus der Chemie 79 % und aus der Physik 85 % der Items unverändert in die Normierungsstudie überführt werden ($MNSQ < 1.15$; Trennschärfe > 0.25 ; Itemschwierigkeit zwischen -3 und 3). Weitere Ergebnisse werden im Rahmen von Begleitstudien veröffentlicht (Walpuski et al., 2010; Wellnitz, 2012).

5 Ausblick

Mit dem dargestellten Kompetenzmodell kann die inhaltliche Differentialität des Kompetenzbereichs Erkenntnisgewinnung detailliert beschrieben und durch theoretisch begründete a priori Annahmen die Aufgabenschwierigkeit zu verschiedenen Teilfähigkeiten gezielt variiert werden. Auf Basis der Pilotierungsdaten wurden aus den explizit gewählten und kontrollierten (Komplexität, kognitive Prozesse) sowie den implizit vorhandenen schwierigkeitsbestimmenden Merkmalen (Aufgabenformat, Kontext, Fachinhalt) Kompetenzstufendeskriptoren generiert. Dazu wurden in einem iterativen Standard-Setting Verfahren nach der Bookmark-Methode Expertenurteile herangezogen, um Schwellenwerte (Cut-Scores) festzulegen, die eine eindimensionale Kompetenzskala in sukzessive

Kompetenzstufen einteilt (Pant, Tiffin-Richards & Köller, 2010). Die Daten der Normierungsstudie ($N = 12.500$) dienen der kriteriumsorientierten Beschreibung der erfassten Kompetenzen zur Festlegung nationaler Kompetenzskalen, um für zukünftige Ländervergleiche eine kriteriumsorientierte Leistungsinterpretation zu ermöglichen. 2012 ist der erste IQB-Ländervergleich für die Naturwissenschaften in Anknüpfung an PISA vorgesehen. Die Berichterstattung zum Ländervergleich erfolgt 2013. Es werden Unterschiede in der Kompetenzausprägung von Schülerinnen und Schülern zwischen den Ländern, zwischen sozialen und ethnischen Gruppen und zwischen Bildungsgängen, die denselben Bildungsabschluss vergeben, ermittelt (Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland & Institut zur Qualitätsentwicklung im Bildungswesen, 2006).

Literatur

- Baumert, J., Klieme, E., Neubrand, M., Prenzel, M., Schiefele, U. & Schneider, W. (2001). *PISA 2000: Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich*. Opladen: Leske + Budrich.
- Baumert, J., Lehmann, R., Lehrke, M., Schmitz, B., Clausen, M. & Hosenfeld, I. (1997). *TIMSS – mathematisch-naturwissenschaftlicher Unterricht im internationalen Vergleich. Deskriptive Befunde*. Opladen: Leske + Budrich.
- Bybee, R. W. (1997). Toward an Understanding of Scientific Literacy. In W. Gräber & C. Bolte (Hrsg.), *Scientific Literacy* (S. 37–68). Kiel: IPN.
- Bybee, R. W. (2002). Scientific Literacy – Mythos oder Realität? In W. Gräber, P. Nentwig, T. Koballa & R. Ewans (Hrsg.), *Scientific Literacy – Der Beitrag der Naturwissenschaften zur Allgemeinen Bildung* (S. 21–43). Opladen: Leske + Budrich.
- Chinn, C. A. & Malhotra, B. A. (2002). Epistemologically Authentic Inquiry in Schools: A Theoretical Framework for Evaluating Inquiry Tasks. *Science Education*, 86(2), 175–218.
- Commons, M. L., Goodheart, E. A., Pekker, A., Dawson, T. L., Draney, K. & Adams, K. M. (2007). Using Rasch Scaled Stage Scores to Validate Orders of Hierarchical Complexity of Balance Beam Task Sequences. In E. V. Smith & R. M. Smith (Hrsg.), *Rasch Measurement: Advanced and Specialized Applications* (S. 121–147). Maple Grove: JAM.
- Department for Education and Skills & Qualification and Curriculum Authority. (2004). *Science – The National Curriculum for England*. London: HMSO.
- Emden, M. (2011). *Prozessorientierte Leistungsmessung des naturwissenschaftlich-experimentellen Arbeitens. Eine vergleichende Studie zu Diagnoseinstrumenten zu Beginn der Sekundarstufe I*. Berlin: Logos.
- Groeben, N. (1982). *Leserpsychologie. Textverständnis – Textverständlichkeit*. Münster: Aschendorff.
- Grube, C. (2011). *Kompetenzen naturwissenschaftlicher Erkenntnisgewinnung. Untersuchung der Struktur und Entwicklung des wissenschaftlichen Denkens bei Schülerinnen und Schülern der Sekundarstufe I*. Verfügbar unter <http://kobra.bibliothek.uni-kassel.de/handle/urn:nbn:de:hebis:34-2011041537247>. [17.06.2011].
- Grube, C. & Mayer, J. (2010). Wissenschaftsmethodische Kompetenzen in der Sekundarstufe I: eine Untersuchung zur Entwicklung des wissenschaftlichen Denkens. In U. Harms & I. Mackensen-Friedrichs (Hrsg.), *Lehr- und Lernforschung in der Biologiedidaktik. Band 4*, (S. 155–168). Innsbruck: StudienVerlag.
- Hammann, M., Phan, T. T. H. & Bayrhuber, H. (2007). Experimentieren als Problemlösen: Lässt sich das SDDS-Modell nutzen, um unterschiedliche Dimensionen beim Experimentieren zu messen? *Zeitschrift für Erziehungswissenschaften*, Sonderheft 8, 33–49.
- Hammann, M., Phan, T. T. H., Ehmer, M. & Grimm, T. (2008). Assessing Pupils' Skills in Experimentation. *Journal of Biological Education*, 42(2), 66–72.
- Hofstein, A., Navon, O., Kipnis, M. & Mamlok-Naaman, R. (2005). Developing Students' Ability to Ask More and Better Questions Resulting from Inquiry-Type Chemistry Laboratories. *Journal of Research in Science Teaching*, 42(7), 791–806.
- Höttecke, D. (2001a). *Die Natur der Naturwissenschaften historisch verstehen. Fachdidaktische und wissenschaftshistorische Untersuchungen*. Berlin: Logos.
- Höttecke, D. (2001b). Die Vorstellungen von Schülern und Schülerinnen von der "Natur der Naturwissenschaften". *Zeitschrift für Didaktik der Naturwissenschaften*, 7, 7–23.
- Kauertz, A. (2008). *Schwierigkeitserzeugende Merkmale physikalischer Leistungstestaufgaben*. Berlin: Logos.
- Kauertz, A. & Fischer, H. E. (2006). Assessing Students' Level of Knowledge and Analysing the Reasons for Learning Difficulties in Physics by Rasch Analysis. In X. Liu & W. J. Boone (Hrsg.), *Applications of Rasch Measurement in Science Education* (S. 212–246). Maple Grove, USA: JAM.

- 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.
- Kircher, E. (2010). Modellbegriff und Modellbildung in der Physikdidaktik. In E. Kircher, R. Girwitz & P. Häußler (Hrsg.), *Physikdidaktik: Theorie und Praxis* (S. 735–761). Berlin, Heidelberg: Springer.
- Kircher, E. & Dittmer, A. (2004). Lehren und lernen über die Natur der Naturwissenschaften – ein Überblick. In C. Hößle, D. Höttecke & E. Kircher (Hrsg.), *Lehren und lernen über die Natur der Naturwissenschaften* (S. 1–22). Kronach: Schneider.
- Klahr, D. (2000). *Exploring Science: The Cognition and Development of Discovery Processes*. Cambridge: MIT.
- Klos, S., Henke, C., Kieren, C., Walpuski, M. & Sumfleth, E. (2008). Naturwissenschaftliches Experimentieren und chemisches Fachwissen – zwei verschiedene Kompetenzen. *Zeitschrift für Pädagogik*, 54(3), 304–321.
- Langer, I., Schulz von Thun, F. & Tausch, R. (2011). *Sich verständlich ausdrücken* (9. Aufl.). München: Reinhardt.
- Lederman, N. G. (2007). Nature of Science: Past, Present, and Future. In S. K. Abell & N. G. Lederman (Hrsg.), *Handbook of research on science education* (S. 831–879). Mahwah, NJ: Erlbaum.
- Lederman, N. G., Abd-El-Khalick, F., Bell, R. L. & Schwartz, R. S. (2002). Views of Nature of Science Questionnaire: Toward Valid and Meaningful Assessment of Learners' Conceptions of Nature of Science. *Journal of Research in Science Teaching*, 36(6), 497–521.
- Leisner, A. (2005). Modellkompetenz im Physikunterricht. In H. Giest (Hrsg.), *Lern- und Lehr-Forschung: LLF-Berichte* (S. 35–49). Potsdam: Universitätsverlag.
- Mannel, S. (2011). *Assessing Scientific Inquiry. Development and Evaluation of a Test for the Low-Performing Stage*. Berlin: Logos.
- Mayer, J. (2007). Erkenntnisgewinnung als wissenschaftliches Problemlösen. In D. Krüger & H. Vogt (Hrsg.), *Handbuch der Theorien in der biologiedidaktischen Forschung – Ein Handbuch für Lehramtsstudenten und Doktoranden* (S. 178–186). Berlin: Springer.
- Mayer, J., Grube, C. & Möller, A. (2008). Kompetenzmodell naturwissenschaftlicher Erkenntnisgewinnung. In U. Harms & A. Sandmann (Hrsg.), *Lehr- und Lernforschung in der Biologiedidaktik. Band 3. Ausbildung und Professionalisierung von Lehrkräften* (S. 63–79). Innsbruck: StudienVerlag.
- McComas, W. F. (2008). Seeking Historical Examples to Illustrate Key Aspects of the Nature of Science. *Science & Education*, 17(2), 249–263.
- McComas, W. F. & Olson, J. K. (1998). The Nature of Science in International Science Education Standards Documents. In W. F. McComas (Hrsg.), *The Nature of Science in Science Education: Rationales and Strategies* (S. 41–52). Dordrecht: Kluwer.
- Meisert, A. (2008). Vom Modellwissen zum Modellverständnis – Elemente einer umfassenden Modellkompetenz und deren Fundierung durch lernerseitige Kriterien zur Klassifikation von Modellen. *Zeitschrift für Didaktik der Naturwissenschaften*, 14, 243–261.
- Mikelskis-Seifert, S. & Fischler, H. (2003). Die Bedeutung des Denkens in Modellen bei der Entwicklung von Teilchenvorstellungen – Stand der Forschung und Entwurf einer Unterrichtskonzeption. *Zeitschrift für Didaktik der Naturwissenschaften*, 9, 75–88.
- Mullis, I. V. S., Martin, M. O., Ruddock, G. J., O'Sullivan, C. Y. & Preuschof, C. (2009). *TIMSS 2011 Assessment Frameworks*. Boston: TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College.
- Neumann, I. (2011). *Beyond Physics Content Knowledge. Modeling Competence Regarding Nature of Scientific Inquiry and Nature of Scientific Knowledge*. Berlin: Logos.
- National Research Council. (1996). *National Science Education Standards*. Washington DC: National Academy Press.

- National Research Council. (2000). *Inquiry and the National Science Education Standards*. Washington, DC: National Academy Press.
- Osborne, J., Collins, S., Ratcliffe, M., Millar, R. & Duschl, R. (2003). What "Ideas-about-Science" Should Be Taught in School Science? A Delphi Study of the Expert Community. *Journal of Research in Science Teaching*, 40(7), 692–720.
- Pant, H. A., Tiffin-Richards, S. P. & Köller, O. (2010). Standard-Setting für Kompetenztests im Large-Scale-Assessment. Projekt Standardsetting. *Zeitschrift für Pädagogik*, 56, 175–188.
- Prenzel, M., Schöps, K., Rönnebeck, S., Senkbeil, M., Walter, O., Carstensen, C. H. & Hammann, M. (2007). Naturwissenschaftliche Kompetenz im internationalen Vergleich. In M. Prenzel, C. Artelt, J. Baumert, W. Blum, M. Hammann, E. Klieme, R. Pekrun (Hrsg.), *PISA 2006: Die Ergebnisse der dritten internationalen Vergleichsstudie* (S. 63–105). Münster: Waxmann.
- Schecker, H. & Parchmann, I. (2006). Modellierung naturwissenschaftlicher Kompetenz *Zeitschrift für Didaktik der Naturwissenschaften*, 12, 45–66.
- Schreiber, N., Theyßen, H. & Schecker, H. (2009). Experimentelle Kompetenz messen?! *Physik und Didaktik in Schule und Hochschule*, 8(3), 92–101.
- Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland (Hrsg.) (2005a). *Beschlüsse der Kultusministerkonferenz – Bildungsstandards im Fach Biologie für den Mittleren Schulabschluss. Beschluss vom 16.12.2004*. München, Neuwied: Luchterhand.
- Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland (Hrsg.) (2005b). *Beschlüsse der Kultusministerkonferenz – Bildungsstandard im Fach Chemie für den Mittleren Schulabschluss. Beschluss vom 16.12.2004*. München, Neuwied: Luchterhand.
- Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland (Hrsg.) (2005c). *Beschlüsse der Kultusministerkonferenz – Bildungsstandards im Fach Physik für den Mittleren Schulabschluss. Beschluss vom 16.12.2004*. München, Neuwied: Luchterhand.
- Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland & Institut zur Qualitätsentwicklung im Bildungswesen (Hrsg.). (2006). *Gesamtstrategie der Kultusministerkonferenz zum Bildungsmonitoring*. Berlin: LinkLuchterhand.
- Shamos, M. H. (2002). Durch Prozesse ein Bewusstsein für die Naturwissenschaften entwickeln. In W. Gräber, P. Nentwig, T. Koballa & R. Evans (Hrsg.), *Scientific Literacy: Der Beitrag der Naturwissenschaften zur Allgemeinen Bildung* (S. 45–68). Opladen: Leske + Budrich.
- Shavelson, R. J., Baxter, G. P., Pine, J., Yuré, J., Goldman, S. R. & Smith, B. (1991). Alternative Technologies for Large Scale Science Assessment: Instrument of Education Reform. *School Effectiveness and School Improvement: An International Journal of Research, Policy and Practice*, 2(2), 97–114.
- Stecher, B. M. & Klein, S. P. (1997). The Cost of Science Performance Assessments in Large-Scale Testing Programs. *Educational Evaluation and Policy Analysis*, 19(1), 1–14.
- Terzer, E. & Upmeier zu Belzen, A. (2007). Naturwissenschaftliche Erkenntnisgewinnung durch Modelle – Modellverständnis als Grundlage für Modellkompetenz. *Berichte des Institutes für Didaktik der Biologie der Westfälischen Wilhelms-Universität Münster*, 16, 33–56.
- Upmeier zu Belzen, A. & Krüger, D. (2010). Modellkompetenz im Biologieunterricht. *Zeitschrift für Didaktik der Naturwissenschaften*, 16, 41–57.

- Urhahne, D., Kremer, K. & Mayer, J. (2008). Welches Verständnis haben Jugendliche von der Natur der Naturwissenschaften? Entwicklung und erste Schritte zur Validierung eines Fragebogens. *Unterrichtswissenschaft*, 36(1), 71–93.
- Walpuski, M., Kampa, N., Kauertz, A. & Wellnitz, N. (2008). Evaluation der Bildungsstandards in den Naturwissenschaften. *Der mathematische und naturwissenschaftliche Unterricht*, 61(6), 323–326.
- Walpuski, M., Kauertz, A., Kampa, N., Fischer, H. E., Mayer, J., Sumfleth, E. & Wellnitz, N. (2010). ESNaS – Evaluation der Standards für die Naturwissenschaften in der Sekundarstufe. In A. Gehrmann, U. Hericks & M. Lüders (Hrsg.), *Bildungsstandards und Kompetenzmodelle – Beiträge zu einer aktuellen Diskussion über Schule, Lehrerbildung und Unterricht* (S. 171–184). Bad Heilbrunn: Klinkhardt.
- Walpuski, M., Ropohl, M. & Sumfleth, E. (2011). Students' Knowledge about Chemical Reactions – Development and Analysis of Standard-Based Test Items. *Chemistry Education Research and Practice*, 12, 174–183.
- Walpuski, M. & Schulz, A. (2011). Erkenntnisgewinnung durch Experimente – Stärken und Schwächen deutscher Schülerinnen und Schüler im Fach Chemie. *Chimica didactica*, 37(104), 6–27.
- Weinert, F. E. (2001). Vergleichende Leistungsmessung in Schulen – eine umstrittene Selbstverständlichkeit. In F. E. Weinert (Hrsg.), *Leistungsmessungen in Schulen* (S. 17–32). Weinheim: Beltz.
- Wellnitz, N. (2012). *Kompetenzstruktur und -niveaus von Methoden naturwissenschaftlicher Erkenntnisgewinnung*. Berlin: Logos.
- Wellnitz, N. & Mayer, J. (2008). Evaluation von Kompetenzstruktur und -niveaus zum Beobachten, Vergleichen, Ordnen und Experimentieren. In D. Krüger, A. Upmeyer zu Belzen, T. Riemeier & K. Niebert (Hrsg.), *Erkenntnisweg Biologiedidaktik 7* (S. 129–143). Hannover: Universität Kassel.
- Wellnitz, N. & Mayer, J. (2012). Beobachten, Vergleichen und Experimentieren: Wege der Erkenntnisgewinnung. In U. Harms, F. X. Bogner (Hrsg.), *Lehr- und Lernforschung in der Biologiedidaktik*. Band 5 (S. 63–79). Innsbruck: StudienVerlag.
- Wu, M. L., Adams, R. J., Wilson, M. R. & Haldane, S. A. (2007). *ConQuest – Generalised Item Response Modelling Software* (Version 2). Camberwell: Australian Council for Educational Research.
- Zilker, I., Kauertz, A. & Fischer, H. E. (2009). Kompetenzdiagnose im Bereich Nature of Science und Scientific Inquiry. In D. Höttecke (Hrsg.), *Chemie- und Physikdidaktik für die Lehramtsausbildung* (S. 392–394). Münster: Lit.

KONTAKT

Dr. Nicole Wellnitz
Universität Kassel, Didaktik der Biologie
Heinrich-Plett-Str. 40
34132 Kassel
nicole.wellnitz@uni-kassel.de

AUTORENINFORMATION

Dr. rer. nat. Nicole Wellnitz arbeitet als wissenschaftliche Mitarbeiterin in der Abteilung Didaktik der Biologie an der Universität Kassel. Ihre Forschungsschwerpunkte sind Kompetenzmodellierung und -messung naturwissenschaftlicher Erkenntnismethoden.

Dr. rer. nat. Hans Ernst Fischer ist Professor für Didaktik der Physik in der Fakultät für Physik der Universität Duisburg-Essen und Sprecher der DFG-Forschergruppe naturwissenschaftlicher Unterricht. Seine Forschungsschwerpunkte liegen im Bereich der fachspezifischen Unterrichtsforschung, der fachspezifischen Analysen zur Lehrerprofessionalisierung unter Berücksichtigung von Lernvoraussetzungen und Lernergebnissen.

Dr. rer. nat. Alexander Kauertz ist Professor für naturwissenschaftliches Lernen mit Schwerpunkt Physik an der Universität Koblenz-Landau (Campus Landau). Seine Forschungsschwerpunkte sind Physikkompetenzmodellierung und -diagnose in der Sekundarstufe sowie Unterrichtsanalysen im naturwissenschaftlichen Sachunterricht der Primarstufe.

Dr. rer. nat. Jürgen Mayer ist Professor für Didaktik der Biologie an der Universität Kassel. Seine Forschungsschwerpunkte sind die Kompetenzmodellierung im Biologieunterricht, die Entwicklung und Analyse von Lernumgebungen, insbesondere des Forschenden Lernens, sowie Lehrerprofessionalisierung.

Dr. phil. nat. Irene Neumann ist Juniorprofessorin für Physikdidaktik an der Ruhr-Universität Bochum. Ihr Forschungsschwerpunkt liegt im Bereich von Nature of Science. Darüber hinaus beschäftigt sie sich mit Aspekten der Mathematik im Physikunterricht.

Dr. Hans Anand Pant ist Professor für Methoden der Erziehungswissenschaft und seit Februar 2010 Direktor des Instituts zur Qualitätsentwicklung im Bildungswesen (IQB) an der Humboldt-Universität zu Berlin.

Dr. Elke Sumfleth ist Professorin für Didaktik der Chemie und Sprecherin des DFG-Graduiertenkollegs „Naturwissenschaftlicher Unterricht“ an der Universität Duisburg-Essen. Ihre Forschungsschwerpunkte sind Empirische Lehr-Lern-Forschung in Chemie mit unterschiedlichen Schwerpunkten und Entwicklung und Evaluation von Materialien für den Chemieunterricht.

Dr. rer. nat. Maik Walpuski ist Professor für Didaktik der Chemie an der Universität Duisburg-Essen und Mitglied der DFG-Forschergruppe Naturwissenschaftlicher Unterricht. Seine Forschungsschwerpunkte sind die Kompetenzmodellierung und Kompetenzmessung für das Fach Chemie sowie die prozessanalytische Untersuchung des Chemieunterrichts, insbesondere in Experimentierphasen.

Anhang zu Tabelle 2

KMK-Standards	Teilbereiche
Die Schülerinnen und Schüler ...	
<p>Biologie</p> <ul style="list-style-type: none"> • mikroskopieren Zellen und stellen sie in einer Zeichnung dar (E1). • beschreiben und vergleichen Anatomie und Morphologie von Organismen (E2). • analysieren die stammesgeschichtliche Verwandtschaft bzw. ökologisch bedingte Ähnlichkeit bei Organismen durch kriteriengeleitetes Vergleichen (E3). • ermitteln mithilfe geeigneter Bestimmungsliteratur im Ökosystem häufig vorkommende Arten (E4). • führen Untersuchungen mit geeigneten qualifizierenden oder quantifizierenden Verfahren durch (E5). • planen einfache Experimente, führen die Experimente durch und/oder werten sie aus (E6). • wenden Schritte aus dem experimentellen Weg der Erkenntnisgewinnung zur Erklärung an (E7). 	<p>Naturwissenschaftliche Untersuchungen (<i>Scientific Inquiry</i>)</p>
<p>Chemie</p> <ul style="list-style-type: none"> • erkennen und entwickeln Fragestellungen, die mit Hilfe chemischer Kenntnisse und Untersuchungen, insbesondere durch chemische Experimente, zu beantworten sind (E1). • planen geeignete Untersuchungen zur Überprüfung von Vermutungen und Hypothesen (E2). • führen qualitative und einfache quantitative experimentelle und andere Untersuchungen durch und protokollieren diese (E3). • beachten beim Experimentieren Sicherheits- und Umweltaspekte (E4). • erheben bei Untersuchungen, insbesondere in chemischen Experimenten, relevante Daten oder recherchieren sie (E5). • finden in erhobenen oder recherchierten Daten, Trends, Strukturen und Beziehungen, erklären diese und ziehen geeignete Schlussfolgerungen (E6). 	
<p>Physik</p> <ul style="list-style-type: none"> • beschreiben Phänomene und führen sie auf bekannte physikalische Zusammenhänge zurück (E1). • wählen Daten und Informationen aus verschiedenen Quellen zur Bearbeitung von Aufgaben und Problemen aus, prüfen sie auf Relevanz und ordnen sie (E2). • stellen an einfachen Beispielen Hypothesen auf (E6). • führen einfache Experimente nach Anleitung durch und werten sie aus (E7). • planen einfache Experimente, führen sie durch und dokumentieren die Ergebnisse (E8). • werten gewonnene Daten aus, ggf. auch durch einfache Mathematisierungen (E9). 	

Fortsetzung Tabelle 2 Anhang nächste Seite ...

... Fortsetzung Tabelle 2 Anhang

KMK-Standards	Teilbereiche
Die Schülerinnen und Schüler ...	
Biologie <ul style="list-style-type: none"> • wenden Modelle zur Veranschaulichung von Struktur und Funktion an (E9). • analysieren Wechselwirkungen mit Hilfe von Modellen (E10). • beschreiben Speicherung und Weitergabe genetischer Information auch unter Anwendung geeigneter Modelle (E11). • erklären dynamische Prozesse in Ökosystemen mithilfe von Modellvorstellungen (E12). • beurteilen die Aussagekraft eines Modells (E13). 	Naturwissenschaftliche Modellbildung <i>(Scientific Modelling)</i>
Chemie <ul style="list-style-type: none"> • nutzen geeignete Modelle (z.B. Atommodelle, Periodensystem der Elemente) um chemische Fragestellungen zu bearbeiten (E7). 	
Physik <ul style="list-style-type: none"> • verwenden Analogien und Modellvorstellungen zur Wissensgenerierung (E3). • wenden einfache Formen der Mathematisierung an (E4). • nehmen einfache Idealisierungen vor (E5). 	
Biologie <ul style="list-style-type: none"> • erörtern Tragweite und Grenzen von Untersuchungsanlage, -schritten und -ergebnissen (E8). 	Wissenschaftstheoretische Reflexion <i>(Nature of Science)</i>
Chemie <ul style="list-style-type: none"> • zeigen exemplarisch Verknüpfungen zwischen gesellschaftlichen Entwicklungen und Erkenntnissen der Chemie auf (E8). 	
Physik <ul style="list-style-type: none"> • beurteilen die Gültigkeit empirischer Ergebnisse und deren Verallgemeinerung (E10). 	

Die grau unterlegten Bildungsstandards werden im Rahmen der Evaluation nicht als praktische Fähigkeiten, sondern ebenfalls als prozedurales Wissen („Wissen, wie“) interpretiert und getestet.

