

Ansätze zur Modellierung informatischer
Anforderungen und Fähigkeiten von Schülern
auf der Basis von Aufgabenmerkmalen

D i s s e r t a t i o n

zur Erlangung des akademischen Grades eines
Doktors der Naturwissenschaften
(Dr. rer. nat.)

durch die Fakultät für Wirtschaftswissenschaften der
Universität Duisburg-Essen
Campus Essen

vorgelegt von

Kirsten Schlüter
geboren in Köln

Essen im März 2015

Tag der mündlichen Prüfung: 4. März 2015

Erstgutachter: Prof. Dr. Torsten Brinda

Zweitgutachter: Prof. Dr. Peter Hubwieser

Zusammenfassung

Für das junge Schulfach Informatik existieren meist nicht empirisch sondern fachlich begründete Kompetenzmodelle, die einen Sollzustand beschreiben. Bislang fehlen Ansätze, die den Istzustand modellieren und einen Bezugsrahmen liefern, in dem informatikspezifische Anforderungen von Aufgaben und die korrespondierenden Fähigkeiten von Schülern darstellbar sind. In dieser Arbeit wird eine Methodik erprobt, ein solches Strukturmodell auf empirisch-analytischem Weg zu gewinnen. Das Modell soll sichtbar machen, welche Dimensionen das System kognitiver Anforderungen und Fähigkeiten aufspannen, die charakteristisch für die Informatik in der Sekundarstufe sind.

Als umfangreiches Untersuchungsmaterial stehen die Aufgaben und Schülerantworten des Online-Wettbewerbs Informatik-Biber zur Verfügung, der jährlich mehrere Tausend Teilnehmer aller Schularten und Jahrgangsstufen der Sekundarstufe verzeichnet, Mädchen wie Jungen. Davon ausgehend, dass neben dem Fachinhalt kognitive Kriterien wie der Abstraktionsgrad oder die Lernzielstufe die Aufgabenschwierigkeit bestimmen, wird ein Katalog informatikspezifischer Merkmale erstellt. In einer Expertenbefragung werden die Wettbewerbsaufgaben nach den Merkmalen klassifiziert.

Eine Clusteranalyse der Aufgaben arbeitet vier Typen heraus: Wiedergabe, Verständnis, Anwendung und Problemlösung, die als *Anforderungsdimensionen* interpretiert werden. Weiter führt die Idee, dass ein Schüler, der etwa abstrakte Aufgaben löst, die Fähigkeit besitzt, mit Abstraktion umzugehen. Die Aufgabenmerkmale werden mit den Antwortmustern der Teilnehmer zu Merkmalsmustern verknüpft. Faktorenanalysen dieser Muster berechnen Faktoren, die als *Fähigkeitsdimensionen* interpretiert werden. Hier tritt die Datenproblematik zutage, die darin besteht, dass die Merkmale bereits im Aufgabensatz nur in bestimmten Kombinationen vorkommen. Sie sind konfundiert.

So ist das Ergebnis der Aufgabenanalyse ein fundierter Entwurf eines mehrdimensionalen Anforderungsmodells. Der Ertrag aus der Analyse der Schülerantworten und Merkmalsmuster ist zunächst ein Vorgehensbericht, der auch die Konfundierung thematisiert, die die Gültigkeit des Fähigkeitsmodells in Frage stellt. Um die Fähigkeitsdimensionen zu überprüfen, wird ein Leitfaden für weitere Iterationen der Modellentwicklung erstellt, der die gezielte Aufgabenkonstruktion empfiehlt.

Abstract

Most competence models for informatics in secondary education are determined according to the structure of the subject. Therefore they represent the *target* state. Up to now, models representing the *actual* state, derived from empirical results, are rarely described. There is a strong need for approaches to describing the actual state in terms of task requirements and student abilities. In this thesis an empirical statistical procedure of developing a structural model is proven. The aim is to visualize the dimensions spanning the system of characteristic cognitive requirements and corresponding abilities relevant in computer science in secondary education.

The German Informatik-Biber is member of an international initiative for the promotion of informatics, addressing students of all types of secondary school, girls and boys. The yearly contest with thousands of participants provides extensive and rich data for the investigation. Assuming that the task difficulty, beneath contents, depends on cognitive characteristics, a catalogue of criteria is specified. Criteria are, for example, the level of abstraction, or Bloom's taxonomic level of learning goals. The tasks of the Informatik-Biber are classified into criteria categories through expert rating.

A cluster analysis of the classified tasks results in four types: reproduction, comprehension, application, and problem solving. The types are interpreted as four dimensions of task requirements. The idea that someone who is able to solve abstract tasks is capable to deal with abstraction, as an example, leads even further. The criteria and the students' answer patterns are combined to students' criteria patterns. Factor analyses result in factors that can be interpreted as dimensions of students' abilities. At this point a problem of the task set emerges: some criteria cannot be separated from each other because they only occur in certain combinations.

So there are two conclusions. Task analysis results in a multidimensional model of task requirements, while pattern analysis of student answers does not result in a valid model. In fact the outcome is a guideline for continuing iterations of model development, placing emphasis on the construction of tasks, eliminating the problem of confounded variables. The intention is to prove the dimensions of student abilities.

Inhaltsverzeichnis

1	Einleitung	11
2	Motivation	15
3	Stand der Forschung	19
3.1	Kompetenzmodellierung	19
3.1.1	Lernziel-Taxonomien	21
3.1.2	Bildungsstandards, Bildungsstudien	25
3.1.3	Methodengrundlagen	25
3.1.4	Mathematik	28
3.1.5	Naturwissenschaften	30
3.1.6	Andere Fächer und Fächerübergreif	34
3.2	Ansätze in der Informatik-Didaktik	36
3.2.1	Lernziel-Taxonomien	37
3.2.2	Computer Literacy	38
3.2.3	Bildungsstandards	39
3.2.4	Kompetenzmodelle	41
3.3	Schwierigkeitsmerkmale	43
3.3.1	Mathematik	43
3.3.2	Naturwissenschaften	46
3.3.3	Andere Fächer und Fächerübergreif	50
3.4	Zusammenschau	52
4	Zielstellung und Methodik	57
4.1	Forschungsfragen	58
4.2	Forschungsmethodik	60
4.2.1	Theoretische Begründung: Aufgabenmerkmale	62
4.2.2	Datenerhebung: Expertenbefragung, Wettbewerb	62
4.2.3	Datenaufbereitung: Klassifizierung, Antwortmuster, Merkmalsquoten	63
4.2.4	Statistische Analysen: Dimensionale Struktur	64
4.2.5	Interpretation: Anforderungen und Fähigkeiten	64
4.2.6	Modellprüfung: Anpassungsgüte	65

5	Aufgabenmerkmale	67
5.1	Erfahrungsweltnähe	69
5.2	Abstraktionsgrad	70
5.3	Formalisierungsgrad	71
5.4	Komplexitätsgrad	73
5.5	Redundanz	74
5.6	Anforderungsbereich	74
5.7	Prozessbereich	75
5.8	Lernzielstufe	76
5.9	Art des Wissens	77
5.10	Inhaltsbereich	78
5.11	Zusammenfassung	78
6	Erhebungen	81
6.1	Schülerwettbewerb Informatik-Biber	82
6.2	Expertenbefragung	85
6.2.1	Auswertungsmodalitäten	87
6.2.2	Aufgabenklassifizierung	92
7	Daten	97
7.1	Aufgabenklassifizierung	97
7.2	Wettbewerbsergebnisse (Antwortmuster)	100
7.3	Personen-Merkmalsquoten	102
7.3.1	Aggregation	102
7.3.2	Konfundierte Effekte	104
8	Analysen: Von den Merkmalen zum Modell	107
8.1	Anforderungsmodell	108
8.1.1	Wechselbeziehungen der Merkmale	109
8.1.2	Clusteranalyse der Aufgaben	112
8.2	Aufgabenkennzahlen	112
8.2.1	Schwierigkeit und Trennschärfe	112
8.2.2	Exkurs: Varianzanalyse der Merkmale	118
8.3	Fähigkeitsmodell A (Antwortmuster)	119
8.3.1	Faktorenanalyse der Antwortmuster	120
8.4	Fähigkeitsmodell B (Merkmalsquoten)	122
8.4.1	Wechselbeziehungen der Merkmale	122
8.4.2	Faktorenanalyse der Merkmalsquoten	124
8.5	Zusammenfassung	126
9	Interpretation: Strukturmodelle	133
9.1	Empirische Aufgabenschwierigkeit	133
9.2	Anforderungsmodell	137

9.3	Fähigkeitsmodell A (Antwortmuster)	142
9.4	Fähigkeitsmodell B (Merkmalsquoten)	146
9.5	Zusammenfassung	150
10	Modellgeltung	155
10.1	Strukturprüfung	155
10.1.1	MultiRa-Analyse	157
10.1.2	MultiRa-Designmatrizen	158
10.1.3	MultiRa-Ergebnisse	159
10.2	Reflexion	163
11	Resümee	167
11.1	Ziel und Ergebnis	167
11.2	Modellbildung	168
11.3	Einschränkungen	169
11.4	Beitrag zur Informatik-Didaktik	169
11.5	Ausblick	171
A	Aufgaben	173
B	Bildungsstandards für die Informatik	197
B.1	Anforderungsbereiche	197
B.2	Inhalts- und Prozessbereiche	198
B.2.1	Inhaltsbereiche	199
B.2.2	Prozessbereiche	200
C	Expertenbefragung	201
C.1	Versendete Unterlagen	201
C.2	Bewertungsverteilung, Merkmalskorrelation	209
	Literaturverzeichnis	217
	Abbildungsverzeichnis	232
	Tabellenverzeichnis	234

Kapitel 1

Einleitung

Angeregt durch die Frage, welche Informatikaufgaben von Schülern leicht oder schwierig empfunden werden, ist diese Forschungsarbeit entstanden. Aufgaben werden im Unterrichtsfach Informatik wie in anderen Fächern zu vielfältigen Zwecken eingesetzt. Sie spielen eine zentrale Rolle bei der Erkundung, Vertiefung, Aktivierung und Übung des Stoffes sowie bei der Lernstandskontrolle, Leistungserhebung oder dem Leistungsvergleich. Angemessene Aufgaben sind dem Unterrichtsgelingen förderlich und wesentlicher Bestandteil der Leistungsbeurteilung. Eine sorgfältige Auswahl von Aufgaben, die dem Zweck und dem Adressaten im Wortsinn angemessen sind, setzt die Einschätzung ihrer kognitiven Anforderungen voraus. In der Praxis wird dazu auf Taxonomien wie die Bloomsche Lernzieltaxonomie zurückgegriffen. Bisher fehlen informatikspezifische Kriterien, um den Aufgabenanspruch in diesem Fach möglichst objektiv zu taxieren. So ist der kognitive Anspruch von Informatikaufgaben Gegenstand dieser Studie. Ziel ist es, Schwierigkeitskriterien zu identifizieren und daraus eine Modellvorstellung der Aufgabenanforderungen und Schülerfähigkeiten zu entwickeln.

Aufgaben werden gemeinhin mit Schwierigkeitsmerkmalen – reproduktiv, schematisch, abstrakt, kreativ, komplex und viele mehr – verbunden, die die Aufgabenstellung oder den erforderlichen Problemlöseprozess charakterisieren. Als Beispiel sei das Merkmal Komplexität genannt, meist als Größe des Lösungsraums verstanden. Werden etwa beim Schachspiel *mehrere* Figuren betrachtet, ist die Komplexität größer als wenn nur ein Spielzug *einer* Figur bedacht wird. Noch komplexer ist, sich die *Abfolge* von Zügen mehrerer Figuren zu vergegenwärtigen. Komplexe Aufgaben gelten als anspruchsvoll, ebenso abstrakte Aufgaben oder solche mit unbekanntem Lösungsansatz, die Kreativität verlangen. Diese Schwierigkeitseinschätzung kann treffend und informativ sein, ist jedoch qualitativer Art. Sie ist eindimensional und gestattet keinen integrierten Vergleich zweier oder mehrerer Kriterien. Ein Bezugssystem fehlt, in dem in mehreren Dimensionen festgehalten wird, was eine leichte und was eine schwierige Aufgabe auszeichnet, und in das jede Aufgabe anhand ihrer Schwierigkeitsmerkmale eingeordnet werden kann.

Die Eingangssituation ist also die, dass informatikspezifische Kriterien und vor allem ein Bezugsrahmen für Aufgabenanforderungen fehlen. Um dem zu begegnen, wird als erster Schritt ein Katalog von Aufgabenmerkmalen erstellt. Dazu werden Publikationen zu klassischen Lernziel-Taxonomien, Forschungsarbeiten der benachbarten Fachdidaktiken, Bildungsstudien und Veröffentlichungen zu inhaltlichen Standardisierungsansätzen in der Schulinformatik ausgewertet. Eine Ausgangsmenge an Kriterien wird identifiziert, von denen ein Effekt auf die Aufgabenschwierigkeit erwartet wird, der teilweise bereits untersucht wurde. Als nächster Schritt werden die Schwierigkeitsmerkmale mit Aufgabenmaterial verknüpft. Schüler bearbeiten die Aufgaben, die Resultate werden erfasst und die empirischen Daten werden statistisch auf Wirkzusammenhänge zwischen Schwierigkeit und Erfolg untersucht. Anhand der Ergebnisse werden die vielzähligen Merkmale zu wenigen, möglichst unabhängigen Dimensionen verdichtet.

Verbildlicht entspricht das entstehende Strukturmodell einem Koordinatensystem mit einer Achse je Dimension, in dem Aufgaben direkt anhand ihrer Merkmale und Personen mittelbar anhand der Aufgaben verortet werden, die sie erfolgreich lösten. Ein Beispiel eines solchen mehrdimensionalen Modells findet sich in der nationalen Ergänzungsuntersuchung zu PISA (das Akronym steht für Programme for International Student Assessment). Darin werden die Typen mathematischen Arbeitens, die „die Spannweite mathematischer Grundbildung markieren“ (Neubrand et al. 2004, S. 39), als Dimensionen vergegenständlicht. Technische Fertigkeiten (1), etwa die Durchführung bekannter Lösungsverfahren, Rechnerisches Modellieren (2), wo erst die Mathematisierung einer Situation den Lösungsansatz ergibt, der sodann rechnerisch durchzuführen ist, und Begriffliches Modellieren (3), wo begriffliche Zusammenhänge im Vordergrund stehen und eine Argumentation zu erbringen ist, charakterisieren die Dimensionen (vgl. S. 30).

In dieser Arbeit wird ein empirisch-analytischer Ansatz erprobt, aus spezifischen Merkmalen von Informatikaufgaben ein Strukturmodell zu konstruieren, das den Aufgabenanspruch und die Fähigkeiten zu deren Lösung beschreibt. Zugrunde liegen die Teilnahmedaten des Online-Schülerwettbewerbs Informatik-Biber¹, der im Jahresturnus ausgerichtet wird. Am Informatik-Biber 2007 nahmen bundesweit über 20.000 Schüler der fünften bis dreizehnten Jahrgangsstufe teil. Im Zuge der Untersuchung werden die Wettbewerbsaufgaben in einer Expertenbefragung nach den Anforderungsmerkmalen eingeschätzt. Von Interesse ist, wie die Bearbeitungserfolge der Schüler im Zusammenhang mit den Merkmalsausprägungen variieren. Die beobachteten Merkmalseffekte werden mit explorativen statistischen Methoden, vor allem Cluster- und Faktorenanalysen, zu wenigen Dimensionen gebündelt, die die Varianz der Schülerergebnisse möglichst gut repräsentieren.

¹Der Informatik-Biber (www.informatik-biber.de) ist Teil der internationalen Initiative Bebras International Contest on Informatics and Computer Fluency (www.bebas.org). In Deutschland bildet er das Einstiegsformat zum Bundeswettbewerb Informatik (www.bundeswettbewerb-informatik.de).

Während das Vorgehen zur strukturellen Modellierung prinzipiell unstrittig ist, erweist sich die Beschaffenheit des Datenmaterials, das für den Informatikwettbewerb und nicht für die Untersuchung erhoben wurde, als ungünstig. Die Anzahl von insgesamt nur neunundzwanzig beziehungsweise fünfzehn Aufgaben je Altersstufe ist problematisch. Für eine Analyse mehrerer Merkmale ist das eine äußerst geringe Fallzahl, denn die Merkmale kommen zwangsläufig nur in bestimmten Kombinationen vor. Diese Merkmalskonfundierung wird im Datenkapitel thematisiert. Positiv steht der eingeschränkten Aussagekraft der Daten in Bezug auf die Merkmale eine große Bandbreite der Aufgabeninhalte und eine beachtliche Größe der Stichprobe gegenüber, die Schüler aller Bundesländer, nahezu aller Schularten und jeder Altersgruppe der Sekundarstufe einschließt.

Der Beitrag verzichtet pragmatisch auf einen aufwändigen eigenen Versuchsaufbau und greift auf den Datenfundus des Informatik-Bibers zurück. Die Arbeiten konzentrieren sich auf einen exemplarischen ersten Durchlauf einer empirisch-statistischen Modellentwicklung. Als Ergebnis erwächst auf der Grundlage der theoretischen Vorüberlegungen zu den Schwierigkeitskriterien eine Modellvorstellung der Dimensionen informatischer Aufgabenanforderungen und Schülerfähigkeiten. Bedingt durch die Konstruktion aus Aufgaben und Schülerdaten bildet das Modell die Realität ab, im Gegensatz zu den verbreiteten normativen Modellen, die eine Erwartung formulieren.

Da sich die empirisch-statistische Herangehensweise an die Modellbildung als praktikabel erweist, ist diese Bestandsaufnahme ein idealer Ausgangspunkt für einen iterativen Entwicklungsprozess, in dem modellgeleitet Aufgaben konstruiert werden, deren Bearbeitungsergebnisse wiederum in die Korrektur, Verfeinerung und Festigung des Modells einfließen.

Gliederung

Die Niederschrift gliedert sich in elf Kapitel. Im Anschluss an die Motivation in Kapitel 2 wird in Kapitel 3 der Forschungsstand zu klassischen Taxonomien, aktuellen Bildungsstudien, Kompetenzmodellierung in der Informatik und in benachbarten Fachwissenschaften sowie schwierigkeits erzeugenden Merkmalen aufgearbeitet. In dem Zusammenhang werden die Begriffe Kompetenzstrukturmodell versus Kompetenzstufenmodell, Aufgabe und Aufgabenschwierigkeit versus Item und Itemschwierigkeit geklärt. In diesen Forschungskontext wird in Kapitel 4 die Zielstellung eingebettet.

Die potentiellen Schwierigkeitsmerkmale werden in Kapitel 5 theoretisch begründet und mit Beispielaufgaben aus dem Informatikwettbewerb verknüpft, der in Abschnitt 6.1 des darauffolgenden Kapitels vorgestellt wird. In Abschnitt 6.2 werden die Modalitäten und Ergebnisse der Expertenbefragung zur Aufgabenklassifizierung beschrieben. Kapitel 7 enthält die Aufbereitung der Rohdaten, so dass für die weiteren Auswertungen Aufgabentabellen mit der Merkmalsklassifizierung und Schülertabellen mit den Wettbewerbsergebnissen bereitstehen.

Thema des Kapitels 8 sind die statistischen Analysen, die durchgeführt werden, um dimensionale Strukturen in den empirischen Daten aufzuspüren. Es werden zwei unterschiedliche Blickwinkel eingenommen, die Aufgabenperspektive und die Personenperspektive. Aus der *Aufgabenperspektive* werden in Abschnitt 8.1 clusteranalytisch vier qualitative Anforderungstypen abgegrenzt, die als Wiedergabe-, Verständnis-, Anwendungs- und Problemlösedimension gedeutet werden. In Abschnitt 8.2 werden ergänzend die Aufgabenkennzahlen empirische Schwierigkeit und Trennschärfe berechnet. Aus der *Personenperspektive* werden Fähigkeitsdimensionen abgegrenzt. Das geschieht in 8.3 durch eine Faktorenanalyse der Antwortmuster (Welcher Schüler löst welche Aufgabe?) und in 8.4 durch eine Faktorenanalyse der Merkmalsquoten (Welcher Schüler löst wie viel Prozent der Aufgaben eines Merkmals?). Die resultierenden Strukturhypothesen werden in Kapitel 9 interpretiert und verglichen.

Abschließend wird die Modellgeltung überprüft. In 10.1 wird jede Strukturhypothese in ein multidimensionales Rasch-Modell eingepasst, in dem anhand einer Bootstrap-Simulation die Anpassungsgüte an die Wettbewerbsdaten geschätzt wird. In 10.2 wird das Vorgehen zur Modellbildung einschließlich Datengrundlage und Merkmalsauswahl reflektiert. Das Fazit ist, das Anforderungsmodell beizubehalten. Das Modell fächert die Kompetenzanforderungen in vier Dimensionen auf, die durch ihre Merkmalsprofile anschaulich charakterisiert werden. Dagegen kann die Modellierung der Personenfähigkeit wegen der Vorbehalte aus der Datenproblematik nicht als abgesichert gelten. Die erprobte Methodik wird als praktikabel und zielführend beurteilt. So mündet Kapitel 11 in der Empfehlung, geleitet durch das Anforderungsmodell neue Aufgaben zu entwickeln, in denen ein überarbeiteter Merkmalsatz systematisch variiert wird, um die bisherigen Ergebnisse zu verfeinern und fortzuentwickeln.

Kapitel 2

Motivation

Das junge Schulfach Informatik nimmt in Deutschland, wo die Bildungshoheit bei den Bundesländern liegt, noch nicht überall einen gleichberechtigten Platz unter den naturwissenschaftlichen Fächern ein. So findet Informatikunterricht in der gymnasialen Mittelstufe in einigen Ländern in einem eigenständigen Pflichtfach statt, in anderen nur in Form einer integrativen informatischen Grundbildung. Eine Gegenüberstellung der Situation des Schulfachs in den Ländern liefert Staruß (2010) in einer Synopse zum Informatikunterricht in Deutschland. Es gibt teilweise länderübergreifende Initiativen von Interessensverbänden und Fachlehrervereinigungen mit dem Zweck, das Profil des Faches in der öffentlichen Wahrnehmung zu schärfen, seine Position in der Schule zu stärken und das Interesse der Schüler zu fördern. Als Beispiel sei das Memorandum der Gesellschaft für Informatik (GI) e. V. zur Stärkung der Schulinformatik genannt (GI 2004).

Die Didaktikforschung leistet ihren Beitrag dazu in Gestalt von Praxisberichten, Vermittlungskonzepten und theoretischen Modellen, die das spezifische Wissen und Können in Informatikdomänen wie Programmieren erklären. Diese Modelle erschließen das komplexe Geflecht aus Fachinhalten und fachtypischen kognitiven Anforderungen, indem sie es auf wenige Komponenten und Niveaustufen reduzieren. Sie werden in der aktuellen Bildungsdiskussion als Kompetenzmodelle bezeichnet und bilden die Grundlage systematischer Evaluationsprozesse, ob zur Qualitätssicherung des Unterrichts, zur Beurteilung individueller Schülerleistungen oder in vergleichenden Schulleistungsstudien.

Die existierenden Kompetenzmodelle für die Informatik in der Sekundarstufe sind rein fachlich begründet, die wenigsten wurden empirisch überprüft. Neben dem normativen Ansatz fehlt bisher eine analytische Vorgehensweise, eine Modellvorstellung der informatischen Kompetenz aus konkreten Erhebungen abzuleiten. Dieser komplementäre Weg wird hier begangen. Mit den Aufgaben und Resultaten des bundesweiten Schülerwettbewerbs Informatik-Biber sind authentische Daten verfügbar. Die Idee ist, einen Katalog beobachtbarer Kriterien zu erstellen, etwa ob Faktenwissen benötigt wird oder ob Formalismen enthalten sind, von denen vermutet wird, dass eine Aufgabe ihretwegen schwieriger ist, also seltener

gelöst wird. Wenn von den Aufgaben bekannt ist, welche Schwierigkeitsmerkmale sie tragen, und von den Versuchspersonen, welche Aufgaben sie gelöst haben, dann können statistische Verfahren eingesetzt werden, die Muster in den Daten sichtbar machen. Es eröffnen sich reiche Forschungsmöglichkeiten. Geleitet von der Frage nach Aufgabentypen können die Aufgaben danach klassifiziert werden, welche Merkmale sie tragen, und von welchen Personen sie gelöst oder nicht bewältigt wurden. Geleitet von der Frage nach Problemlösetypen können die Personen danach klassifiziert werden, welche Aufgaben mit welchen Eigenschaften sie erfolgreich bearbeitet haben. Der Zusammenhang zwischen Aufgabenmerkmalen und Schwierigkeit kann analysiert werden.

Diese Untersuchung konzentriert sich auf die Fragestellung, welche Eigenschaften den Aufgabenanspruch prägen. Um das Schwierigkeitsgefüge aus unüberschaubar vielen Eigenschaften abzubilden, wird ein möglichst sparsames, das heißt geringdimensionales Modell gesucht. Als Koordinatensystem begriffen, wie in Abbildung 2.1 angedeutet, können zum Beispiel bei drei Modelldimensionen in der größten Auflösung acht hypothetische Aufgabentypen unterschieden werden, von der Aufgabe ohne informatikspezifische Anforderungen, im Ursprung A des Koordinatensystems verortet, über die Aufgabe mit Anforderungsschwerpunkten in einer oder zwei Dimensionen, bis zur Aufgabe mit Anforderungsaspekten aller Dimensionen, verortet im Punkt G. Je nach Feinheitgrad des Modells sind auch Abstufungen möglich.

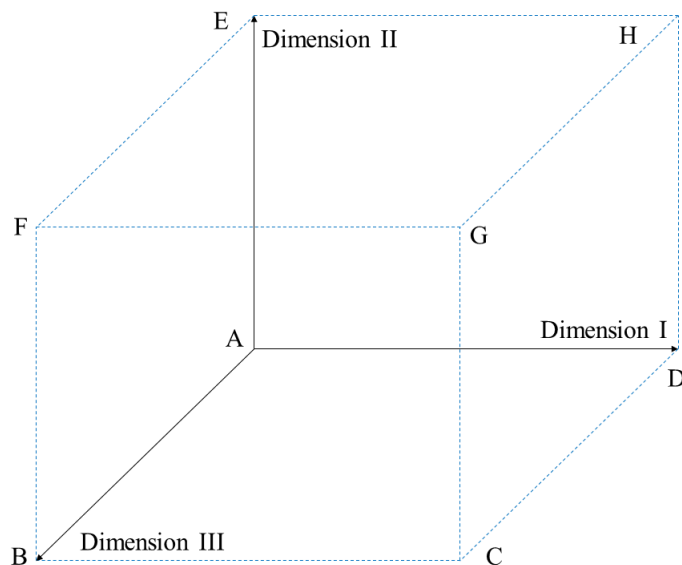


Abbildung 2.1: Schema eines dreidimensionalen Strukturmodells. Die Eckpunkte markieren idealtypische Aufgaben mit Anforderungen in einer Dimension (B, D, E), in zwei (C, F, H) oder drei Dimensionen (G) oder in keiner Dimension (A).

Von großer Wichtigkeit ist die Auswahl der Merkmale, weil das Modell nur diejenigen Teilaspekte abbilden kann, die in die Modellierung eingebracht werden. Ein Vorteil der Konstruktionsweise der Modelldimensionen aus den Aufgabenmerkmalen ist, dass die resultierende Modellvorstellung unmittelbar auf vorhandenes Unterrichtsmaterial im Schulkontext übertragbar ist und neue Aufgaben unkompliziert anhand ihrer Eigenschaften eingeordnet werden können.

Der Prozess der Modellentwicklung wird im Folgenden beschrieben. Ein Weg wird aufgezeigt, auf der Basis empirischer Daten ein Kompetenzmodell „bottom-up“ zu konstruieren. Die Vorgehensweise ist übertragbar auf weitere Fachinhalte und andere Altersgruppen und Schularten. Das resultierende Modell dient der Charakterisierung und vergleichenden Einordnung vorhandener Aufgaben wie der modellgeleiteten Erstellung neuer Aufgaben. Es nimmt auch Personen auf, die anhand ihrer Fähigkeiten, nachgewiesen durch die erfolgreiche Bearbeitung entsprechender Aufgaben, in das gleiche Bezugssystem eingeordnet werden wie die Aufgaben. Das wiederum eröffnet vielfältige weitere Betrachtungen, zum Beispiel Vergleiche zwischen Altersgruppen, Mädchen und Jungen, Unterrichtskonzepten. Nicht zuletzt bedarf es empirisch verankerter Anforderungs- und Leistungsmodelle, wie sie mit geeignetem Aufgabenmaterial nach dem Vorbild dieser Arbeit entstehen können, um als Argumentationsgrundlage für Bildungsentscheidungen den Stand informatischer Bildung zu erfassen.

Kapitel 3

Stand der Forschung

Die Diskussion um Bildungsstandards im Schulfach Informatik gab wichtige Impulse für diese Arbeit. Tragfähige Bildungsstandards, wie im Expertengremium um Klieme et al. (2003) erörtert, benötigen als theoretisches Fundament ein Modell dessen, was Schülerinnen und Schüler zu leisten vermögen. Im weitesten Sinne zielen schon Ordnungssysteme von Lernstufen und kognitiven Prozessen wie die bekannte Taxonomie von Bloom et al. (1972) darauf ab. Die aktuelle Forschung zur Kompetenzmodellierung, die besonders in den sprachlichen, mathematischen und naturwissenschaftlichen Fächern mit der Schulleistungsstudie PISA (Programme for International Student Assessment) intensiviert wurde, brachte eine Reihe von Modellansätzen hervor. Auch die Entwicklungsmethoden rückten seither verstärkt in das Interesse. Hier wird eine Entwicklungsmethodik gewählt, die aus Aufgabencharakteristika als Bausteinen bottom-up ein Kompetenzmodell aufbaut. Das lenkt die Aufmerksamkeit auf potentielle Schwierigkeitsmerkmale, deren Relevanz und Einflussgrad insgesamt sehr uneinheitlich beurteilt werden.

In diesem Kapitel wird ein Überblick über Methoden der Kompetenzmessung und Modellvorschläge aus der Mathematik und den Naturwissenschaften gegeben (Abschnitt 3.1). Ausführlicher dargestellt werden der Forschungsstand in der Informatik (Abschnitt 3.2) und Forschungsauszüge, die sich den Schwierigkeitsmerkmalen von Aufgaben widmen (Abschnitt 3.3).

3.1 Kompetenzmodellierung

Die Forschungsberichterstattung zur Kompetenzmodellierung ist geprägt von der fortdauernden Begriffsdiskussion. Chomsky (1970) grenzte in den 60er-Jahren *theoretische Kompetenz* als unabdingbare Voraussetzung für *aktuelle Performanz* ab. „Eine Untersuchung des [Sprach-] Lernens kann sinnvollerweise erst dann erfolgen, wenn diese vorgängige Aufgabe gelöst ist und man zu einer vernünftigen, gut bestätigten Theorie der zugrunde liegenden Kompetenz gelangt ist – das heißt im Fall der Sprache: zur Formulierung der generativen Grammatik, die

dem beobachteten Sprachgebrauch zugrunde liegt“ (Chomsky 1970, S. 120f). Roth (1971), Mitglied der Bildungskommission des Deutschen Bildungsrates, übernahm den Begriff für die Pädagogik. Er führte sein anthropologisches Konzept von Mündigkeit als Selbst-, Sach-, Methoden- und Sozialkompetenz ein, das Verantwortlichkeit, Urteils- und Handlungsfähigkeit und damit Zuständigkeit einschließt (Roth 1971, S. 180). Auf Mündigkeit gerichtet, ordnet Roth Lernprozesse aufsteigend an, mit problemlösenden und kreativen als höchsten Lernleistungen.

„Dabei muss allerdings einsichtig sein, dass dieses Ziel einen immer erneut zu wiederholenden mühseligen Durchgang durch Lernstufen fordert, bis über die *Reproduktion des Gelernten, die Reorganisation unter neuen Aspekten schließlich die produktive Transformation des Gelernten auf neue Probleme gelingt*“ (Roth 1971, S. 475, Hervorhebung im Original).

Über die Empfehlungen der Bildungskommission hat die Stufung Reproduktion, Reorganisation, Transfer, Problemlösen und Kreativität modellbildenden Eingang in Lehrpläne und Standards aller Fächer gefunden. Bis heute erfährt das Konstrukt Kompetenz mannigfaltige Auslegungen. Weinert (1999) hat in einem Gutachten für die OECD (Organisation for Economic Co-operation and Development) eine umfangreiche Aufstellung verschiedener Kompetenzbegriffe vorgenommen. Weinerts oft zitierte Definition liegt der Mehrheit der aktuellen Forschungsansätze in Deutschland zu Grunde.

„Dabei versteht man unter Kompetenz die bei Individuen verfügbaren oder durch sie erlernbaren kognitiven Fähigkeiten und Fertigkeiten, um bestimmte Probleme zu lösen, sowie die damit verbundenen motivationalen, volitionalen und sozialen Bereitschaften und Fähigkeiten, um die Problemlösungen in variablen Situationen erfolgreich und verantwortungsvoll nutzen zu können“ (Weinert 2001, S. 27f).

Eine breite Auffächerung von Kompetenzauffassungen bietet Schaper (2008) an, der „versucht das Konstrukt ‚Kompetenz‘ aus einer (arbeits-)psychologischen Perspektive zu beleuchten“ (Schaper 2008, S. 91). Er erklärt Kompetenz als

- Persönlichkeitsmerkmal (Intelligenz, Kreativität, emotionale Stabilität, Offenheit für neue Erfahrungen),
- Eignungsmerkmal (Teamfähigkeit, Führungsmotivation, Durchsetzungsfähigkeit),
- Wissens- und Fähigkeitsvoraussetzung (Kenntnisse, Fertigkeiten, Fähigkeiten),
- Fähigkeit zum situationsangemessenen Verhalten (Performanz),
- Befähigung zur handelnden Bewältigung komplexer Anforderungssituationen (kognitive, sozial-kommunikative, emotional-motivationale Voraussetzungen),
- Expertise (intensive Erfahrung und Übung),

- Selbstorganisationsdisposition (lernend und problemlösend in neuartigen Situationen),
- biographisches Konstrukt (durch biographische Erfahrungen und Entwicklungsanforderungen erworbene Fähigkeiten) (Schaper 2008, S. 92f).

Schaper stellt kompetenzorientierte Ansätze der Kompetenzforschung den konventionellen arbeitsanalytischen Ansätzen gegenüber. Weniger Aufgaben als Anforderungssituationen sind die Analyseeinheit, selbstorganisationstheoretische Bezüge bilden die theoretische Fundierung, qualitative Analysemethoden werden bevorzugt, die Auswertung erfolgt mehr interpretativ-generisch als analytisch-vergleichend, die Anforderungscharakterisierung ist eher personen- als aufgabenbezogen und die Ausrichtung der Analyse ist prospektiv-strategisch, also in die Zukunft gerichtet (Schaper 2008, S. 94f).

In den Fachdidaktiken ist zu beobachten, dass sich der Tätigkeitsschwerpunkt von der anfänglichen Ableitung normativer Modelle, die eine Erwartungshaltung widerspiegeln, zum empirischen Abgleich verschiebt. Zum Einsatz kommen klassische statistische Analyseverfahren und mehr und mehr auch Methoden der Item-Response-Theorie (IRT). Allgemein wendet sich das Interesse verstärkt der Kompetenzentwicklung zu. Kompetenz in ihrer Dynamik zu betrachten erfordert einen stärkeren Rückhalt durch psychologische Modelle, der bisher nicht immer gegeben ist, denn Dynamik kann neben Verlauf, Zuwachs, Stagnation oder sogar Rückschritt auch Interaktion mit der Persönlichkeitsentwicklung bedeuten.

Bezug zu dieser Arbeit

Aus der Fülle an Veröffentlichungen zur Kompetenzmodellierung, meist aus den Fachdidaktiken, Pädagogik und Psychologie oder den Bildungsforschungsinstituten, werden im Weiteren diejenigen ausgewählt, die die Informatik oder benachbarte Domänen betreffen und sich mit dem Zusammenhang zwischen Aufgaben, Lösungsmustern und Kompetenzmustern befassen. Der Überblick beginnt mit Lernzieltaxonomien und reicht über die PISA-Rahmenkonzeption bis zu Erprobungen neuer mehrdimensionaler Messmodelle.

Mit den Kompetenzauffassungen verglichen, die Schaper (2008) auffächert, entspricht der Kompetenzbegriff in dieser Arbeit dem klassischen Verständnis von Kompetenz als Wissens- und Fähigkeitsvoraussetzung. Auch der Forschungsansatz ist so gesehen konventionell: Aufgaben, nicht Situationen werden analysiert; die Auswertung erfolgt analytisch-vergleichend mit quantitativen Analysemethoden und orientiert sich am Ist-Zustand.

3.1.1 Lernziel-Taxonomien

Bloom et al. (1972) stellen 1956 mit der Absicht, Lernziele standardisiert zu klassifizieren, eine Taxonomie Kognitiver Lernziele vor. Sie grenzen den Bereich

Kognitive Lernziele gegen die Bereiche Affektive Lernziele und Psychomotorische Fertigkeiten ab. Die Taxonomie mit sechs hierarchischen Stufen findet als Fundament zahlreicher klassifizierender Untersuchungen Eingang in alle Fächer:

1. Wissen (Knowledge)
2. Verständnis (Comprehension)
3. Anwendung (Application)
4. Analyse (Analysis)
5. Synthese (Synthesis)
6. Bewertung (Evaluation)

Anderson und Krathwohl et al. (2001) stellen fest, dass die Stufung empirisch nur teilweise nachvollziehbar und die eindimensionale Struktur allzu stark vereinfachend sei. Sie revidieren die Bloomsche Taxonomie und formen daraus ein zweidimensionales System, in dem die Stufe Wissen zu einer eigenen *Wissens-Dimension* (Knowledge Dimension) erweitert und mit der *Prozess-Dimension* (Cognitive Process Dimension) gekreuzt wird. Die Wissensarten bezeichnen sie mit Substantiven und die modifizierten Prozesskategorien mit Verben, was den Prozesscharakter unterstreicht:

Kognitive Prozesse (The Cognitive Process Dimension)

1. Erinnern (Remember)
2. Verstehen (Understand)
3. Anwenden (Apply)
4. Analysieren (Analyze)
5. Bewerten (Evaluate)
6. Erschaffen (Create)

Wissensarten (The Knowledge Dimension)

- A Faktenwissen (Factual Knowledge)
- B Konzeptwissen (Conceptual Knowledge)
- C Prozedurwissen (Procedural Knowledge)
- D Metakognition (Metacognitive Knowledge)

Das Gerüst des revidierten Modells ist die Originaltaxonomie, erweitert um die Möglichkeit, die Wissensart in die Taxierung einzubringen. Unverändert wird erwartet, dass die Prozesskategorien im Sinne wachsender Komplexität und Abstraktheit hierarchisch sind. Indes wird nicht länger unterstellt, dass die Kategorien Stufen sind, deren untere unbedingt vor den oberen erreicht werden.

„First, although the dimension is believed to be hierarchical in that the ‚higher‘ categories are believed to be more complex and more abstract than the ‚lower‘ ones, the constraint of a ‚cumulative hierarchy‘

was removed. That is, for example, Understand is no longer a necessary prerequisite for Apply. In fact, a student may Apply in part in order to Understand. Second, when multiple cognitive processes are included in an objective (e. g., Understand, Analyze), the objective is classified according to the most complex cognitive process (e. g., Analyze)“ (Anderson 2005, S. 106).

Auch Marzano und Kendall (2007) kritisieren, dass die Annahme einer Schwierigkeitshierarchie der Lernzielstufen theoretisch angreifbar sei. Sie schlagen ein zweidimensionales Modell vor, in dessen Dimension *Verarbeitungsstufen* (Levels of Processing) drei mentale Systeme angeordnet sind, das kognitive System mit vier Stufen, das metakognitive System als fünfte, das Selbstsystem als sechste Stufe. In der Dimension *Wissensgebiete* (Domains of Knowledge) sind die drei Komponenten des Wissens vertreten, mit denen jeder Lernprozess notwendig verbunden ist, Information, mentale und psychomotorische Abläufe. Mit dem Selbstsystem (6) und den psychomotorischen Abläufen (C) werden die Bereiche affektiver und psychomotorischer Lernziele integriert, die Bloom ausklammert:

Verarbeitungsstufen (Levels of Processing)

1. Kenntnisse abrufen (Retrieval, Cognitive System)
2. Verständnis (Comprehension, Cognitive System)
3. Analyse (Analysis, Cognitive System)
4. Kenntnisse verwenden (Knowledge Utilization, Cognitive System)
5. Metakognitives System (Metacognitive System)
6. Selbstsystem (Self-system)

Wissensgebiete (Domains of Knowledge)

- A Information
- B Mentale Abläufe (Mental Procedures)
- C Psychomotorische Abläufe (Psychomotor Procedures)

Die Komplexitätsstufung der Prozesse wird aufgehoben und durch eine geschachtelte Hierarchie ersetzt, die auf dem Kontrollfluss und dem Bewusstseinsgrad fußt (Marzano und Kendall 2007, S. 16). Vom Selbstsystem geht die Motivation aus, eine Aufgabe zu priorisieren und in Angriff zu nehmen. Das Selbstsystem kontrolliert das metakognitive System, das für Ziele und Strategien zuständig ist und mit dem kognitiven System interagiert, welches schließlich die kognitive Aktivität steuert. So sind die drei mentalen Systeme hierarchisch angeordnet (Marzano und Kendall 2007, S. 17). Innerhalb des kognitiven Systems nimmt der Bewusstseinsgrad von der Anwendung (4) bis zum Abruf (1) von Kenntnissen mit zunehmender Routine ab. Unter der Annahme, dass Prozesse des Selbstsystems (6) mehr bewusstes Denken erfordern als Metakognition (5) und Metakognition ihrerseits mehr als Kognition, bilden die sechs Verarbeitungsstufen eine Hierarchie des Bewusstseinsgrades (Marzano und Kendall 2007, S. 60f).

Der Neuentwurf der Taxonomie findet hier als zweidimensionales Modell mit einer Prozess- und einer Wissensdimension Erwähnung. Neu ist die Annahme, dass die Schwierigkeit nicht allein aufgabeninhärent ist, sondern wesentlich von der Vertrautheit des Bearbeiters mit dem Löseprozess abhängt.

„The complexity of a mental process is invariant – the number of steps and their relationship do not change. However, familiarity with a process will change over time. The more familiar one is with a process, the more quickly one executes it, and the easier it becomes“ (Marzano und Kendall 2007, S. 11).

Fuller et al. (2007) diskutieren eine Auswahl der existierenden Taxonomien, darunter die von Bloom et al. (1972), die Revision von Anderson (2001), die Taxonomie von De Block (Fuller et al. 2007, S. 155), in der eine erste Dimension mit den Kategorien Knowledge (1), Understanding (2), Application (3), Integration (4) durch eine zweite Dimension mit den Kategorien Cognitive (C), Affective (A), Psychomotor domain (P) ergänzt wird, und die SOLO-Taxonomie von Biggs und Collis (1982; SOLO steht für Structure of the Observed Learning Outcome), die Prüfungswissen als Prestructural, Unistructural, Multistructural, Relational oder Extended Abstract klassifizieren (Fuller et al. 2007, S. 155). Die Arbeitsgruppe um Fuller stellt Aufgaben heraus, die charakteristisch für die Informatik als Disziplin sind, und schlägt eine informatikspezifische Adaption der Bloomschen Taxonomie vor, die Matrix Taxonomie. Diese wird hier als Lernzieltaxonomie mit besonderer Relevanz für die Informatik vorgestellt (S. 38).

Bezug zu dieser Arbeit

Die Lernziel-Taxonomien bieten theoretisch und empirisch begründete Anforderungskategorien. Die Bloomsche Taxonomie ist etabliert und findet breiten Einsatz in allen Fächern, da die Kategorien direkt und eindeutig anwendbar sind. Die Revision durch Anderson und Krathwohl ist berechtigt in den Punkten, dass die Originaltaxonomie die Wissensart nicht berücksichtigt und dass die Lernzielstufen nicht notwendig aufeinander aufbauen. In dieser Arbeit werden die Prozess- und die Wissenskategorien nach Anderson und Krathwohl in die Startmenge der Schwierigkeitsmerkmale aufgenommen. In einem Exkurs wird unter anderem untersucht, ob für die Informatikaufgaben des Schülerwettbewerbs die Vermutung hierarchischer Kategorien im Sinne wachsender Schwierigkeit zutrifft (S. 118).

Eine Besonderheit der Taxonomie von Marzano und Kendall ist die Abbildung der Geläufigkeit des Lösungsvorgangs, der Routine, die zweifelsohne eine Wirkung auf die subjektive Aufgabenschwierigkeit hat. Diese Studie beschränkt sich explizit auf Merkmale, die den Aufgaben eigen und unabhängig von der Person sind, die die Aufgabe bearbeitet. Aus diesem Grund und auch, weil Marzano und Kendall mit der Verarbeitungsstufe Selbstsystem die Motivation einbeziehen, die nicht zum kognitiven Bereich zählt, wird die Taxonomie nicht berücksichtigt.

Fuller et al. stellen mit der Matrix Taxonomie eine informatikspezifische Adaption der Bloomschen Taxonomie vor. Ihr Artikel ist darüber hinaus von Interesse, weil die Autoren verbreitete und weniger verbreitete Taxonomien aus der Warte der Computer Science diskutieren.

3.1.2 Bildungsstandards, Bildungsstudien

Einen kurz gefassten Überblick über die aktuelle, stark verzweigte Diskussion des Kompetenzkonzepts geben Oelkers und Reusser (2008) im Kontext der Bildungsstandards in einer Expertise im Auftrag des Bundesministeriums für Bildung und Forschung. Sie fordern die Einbindung von Kompetenztheorien in die Erforschung und Überwachung der Prozesse des kumulativen Kompetenzaufbaus. Hartig und Klieme (2006) präzisieren den Kompetenzbegriff in Abgrenzung zu den klassischen Leistungskonzepten Intelligenz, Aufmerksamkeit, Gedächtnis und Kreativität. Sie stellen Konzepte und Methoden der Kompetenzdiagnostik in PISA (Programme for International Student Assessment) vor. PISA ist eine international standardisierte Leistungsmessung, die alle drei Jahre von der OECD (Organisation für wirtschaftliche Zusammenarbeit und Entwicklung) mit fünfzehnjährigen Schülerinnen und Schülern an ihren Schulen durchgeführt wird. Die Inhalte sind Leseverständnis (Reading Literacy), mathematische Grundbildung (Mathematical Literacy), naturwissenschaftliche Grundbildung (Scientific Literacy) und fächerübergreifende Kompetenzen (Cross-Curricular Competencies). Der erste vollständige PISA-Zyklus umfasst die Jahre 2000 mit dem Schwerpunkt Lesen, 2003 mit dem Schwerpunkt Mathematik und 2006 mit dem Schwerpunkt Naturwissenschaften. An einem zweiten Testtag findet jeweils ein nationaler Ergänzungstest statt, der sich in Deutschland am Stoff der Schulfächer orientiert.

Bezug zu dieser Arbeit

Anhand zahlreicher begleitender Veröffentlichungen lässt sich der iterative Prozess nachvollziehen, in dem die Methoden- und Modellgrundlagen im Fortgang erprobt und weiterentwickelt wurden. Für diese Arbeit sind vor allem die Modelle und Methoden aufschlussreich, auf die sich die deutschen PISA-Ergänzungstests in den mathematischen und naturwissenschaftlichen Fächern stützen. Sie kommen in den Abschnitten zur Kompetenzmodellierung in der Mathematik (S. 28) und den Naturwissenschaften (S. 30) zur Sprache.

3.1.3 Methodengrundlagen

Das Spektrum der Entwicklungsmethoden in der Kompetenzmodellierung reicht von explorativen und konfirmatorischen Faktorenanalysen bis zu den Testmodellen der Item-Response-Theorie (IRT). In den Lehrbüchern von Bortz und Döring

(2006), Bühner (2006) und Rost (2004a) werden die Verfahren ausführlich dargestellt. Eine kurz gehaltene, verständliche Darstellung psychometrischer Kompetenzmodelle präsentieren Hartig und Jude (2007) in ihrer Expertise zu den Möglichkeiten und Voraussetzungen technologiebasierter Kompetenzdiagnostik im Auftrag des Bundesministeriums für Bildung und Forschung.

Rost (2004b) vergleicht drei Verfahrensweisen zur Abbildung von Kompetenzen und Kompetenzstufen in einem psychometrischen Modell, verkürzt mit Stufen-, Dimensionen- und Muster-Modellen bezeichnet. In den aus PISA bekannten *quantitativen* eindimensionalen Modellen werden Kompetenzstufen durch die Verteilung der Aufgabenschwierigkeiten beziehungsweise Personenfähigkeiten auf einer Rasch-Skala festgelegt. Das Vorgehen stieß im Naturwissenschaftstest von PISA 2000 an seine Grenzen, denn die postulierten Kompetenzstufen ließen sich empirisch nicht in Form von Schwierigkeitsclustern validieren (Rost 2004b, S. 665). Demgegenüber gehen *qualitative* Itemkomponentenmodelle von einer Charakterisierung der Testaufgaben durch Merkmale aus, verbunden mit bestimmten kognitiven Anforderungen. Die Itemkomponenten drücken je Aufgabe und je Merkmal aus, ob das Merkmal in der Aufgabe vorhanden ist, und gehen als Gewichte in die Modellierung der Lösungswahrscheinlichkeit ein. Rost unterscheidet nun Stufen-Modelle als eindimensionale *Schwierigkeitsmodelle*, die nur die Itemschwierigkeit berücksichtigen, von Dimensionen-Modellen als mehrdimensionalen *Fähigkeitsmodellen*, die die Annahme machen, jedem Aufgabenmerkmal entspreche eine Fähigkeitsdimension. Das multidimensionale Itemkomponentenmodell MultiRa, das als Verallgemeinerung des Rasch-Modells in dieser Arbeit strukturprüfend verwendet wird, ist ein solches Fähigkeitsmodell. Noch weiter gehen Muster-Modelle oder latente Klassenmodelle als *Mischverteilungsmodelle*, die nicht nur die Aufgaben, sondern auch die Schüler nach qualitativen Gesichtspunkten differenzieren, nämlich nach den Antwortmustern anstatt nach dem Gesamtscore. „Die Anwendung von Mischverteilungsmodellen zielt darauf ab, die Gesamtpopulation derart zu *entmischen*, dass in den Mischungskomponenten ein bestimmtes Modell passt. Beim Mischverteilungs-Rasch-Modell (mixed Rasch model) soll das Rasch-Modell in jeder Mischungskomponente (latente Klasse) gelten“ (Rost 2004b, S. 672).

Die DFG (Deutsche Forschungsgemeinschaft) richtete im Jahr 2007 mit sechsjähriger Laufzeit das Schwerpunktprogramm „Kompetenzmodelle zur Erfassung individueller Lernergebnisse und zur Bilanzierung von Bildungsprozessen“ ein. In 23 interdisziplinären Einzelprojekten der Domänen Mathematik, Naturwissenschaften, Sprache und Lesen, Fächerübergreifende Kompetenzen, Lehrerkompetenzen werden unter dem Dach des Schwerpunktprogramms die kognitionspsychologischen und fachdidaktischen Grundlagen von Kompetenz sowie psychometrische Modelle zu ihrer Messung erforscht. Das Programm folgt der Begriffsbestimmung durch Klieme und Leutner (2006), die ihren Begriff der Kompetenz durch die Abhängigkeit vom Kontext explizit gegen die generalisierte, kontextunabhängige Intelligenz abgrenzen.

„Für das SPP [Schwerpunktprogramm Kompetenzmodelle] definieren wir Kompetenzen als kontextspezifische kognitive Leistungsdispositionen, die sich funktional auf Situationen und Anforderungen in bestimmten Domänen beziehen“ (Klieme und Leutner 2006, S. 879).

Im Jahr 2010 erschien ein Zwischenbericht zu den Einzelprojekten, ergänzt durch kritische Gutachten von Experten der oben genannten Domänen (Klieme et al. 2010). Zwei Projekte zu mehrdimensionalen IRT-Modellen werden herausgegriffen, von denen Impulse für dieses Vorhaben ausgingen.

In dem Projekt „Modellierung von Kompetenzen mit mehrdimensionalen IRT-Modellen“ gehen Hartig und Höhler (2010) anhand der Daten aus der DESI-Studie (Deutsch Englisch Schülerleistungen International) der Frage nach, ob die aufwändige mehrdimensionale IRT-Modellierung diagnostisch informativer ist als die einfachere eindimensionale IRT-Modellierung. Zumindest auf der Basis der Anpassung an die empirischen Daten sei keine Entscheidung möglich (Hartig und Höhler 2010, S. 192). Zwei wesentliche Eigenschaften der mehrdimensionalen Modelle werden genannt, die einen Mehrwert versprechen. Erstens könnten diese Modelle eine komplexe Struktur abbilden, bei der ein Item durch mehrere latente Dimensionen beeinflusst wird. Diese Konstellation wird mit Within-Item-Mehrdimensionalität im Gegensatz zu Between-Item-Mehrdimensionalität bezeichnet, bei der ein Item nur einer latenten Dimension zugeordnet werden kann. Und zweitens könnten die latenten Dimensionen kompensatorisch oder nicht-kompensatorisch verknüpft werden. Kompensatorisch bedeutet, Stärken in einer Dimension können Schwächen in anderen Dimensionen ausgleichen – die Lösungswahrscheinlichkeit einer Aufgabe hängt von der Summe der Fähigkeitsparameter ab. Nicht-kompensatorisch bedeutet, für die erfolgreiche Bearbeitung einer Aufgabe sind alle Teilfähigkeiten notwendig – die Lösungswahrscheinlichkeit hängt vom Produkt der Fähigkeitsparameter ab. Die Autoren stellen in Aussicht, weitere Vergleiche anzustellen und dazu Testaufgaben neu zu entwickeln, bei denen die Annahmen über die Dimensionalität systematisch berücksichtigt werden.

Im Projekt „Kognitive Diagnosemodelle: Theoretisches Potential und methodische Probleme“ widmen sich Kunina-Habenicht et al. (2010) den neuartigen Kognitiven Diagnosemodellen (CDM). Darunter werden IRT-Modelle verstanden, die den Personen aufgrund ihrer Testscores wahrscheinliche Kompetenzprofile zuweisen. Die Autoren verweisen auf die Neuentwicklung verschiedener CDM, über die im angelsächsischen Sprachraum berichtet wird, für die aber noch wenige erfolgreiche empirische Anwendungen bekannt sind (Kunina-Habenicht et al. 2010, S. 77). Es werden Vorteile durch die mehrdimensionale Abbildung von Kompetenzen, die Verfügbarkeit statistisch begründeter Kompetenzprofile und die Modellierung nicht-kompensatorischer Zusammenhänge erwartet. In der ersten Projektphase werden die Modelle an einem Mathematiktest in der Grundschule erprobt. Im Vergleich mit der korrespondierenden konfirmatorischen Faktorenanalyse gelingt es zwar nicht, zusätzliche Information bereitzustellen, aber die

CDM erlauben eine direkte Aufbereitung der Information in Gestalt von Kompetenzprofilen der Schüler. Für die zweite Projektphase steht aus, die Validität der so berechneten Kompetenzprofile zu überprüfen sowie geeignete Modell-Fit-Maße zu untersuchen, die Auskunft über die Modellpassung geben.

Bezug zu dieser Arbeit

Das gesuchte Anforderungsmodell für Informatikaufgaben ist nach Rost ein qualitatives Dimensionen-Modell. Die Fähigkeitsmodelle, die aus den Schülerantworten entwickelt werden, können darüber hinaus als Muster-Modelle bezeichnet werden, weil sie Fähigkeitstypen nach den Antwortmustern differenzieren. Sowohl das Aufgabenmodell als auch die Schülermodelle sind geeignet, die von Hartig und Höhler (2010) beschriebene Eigenschaft der Within-Item-Mehrdimensionalität abzubilden. Dabei wird ein Item durch mehrere Dimensionen beeinflusst, was der Auffassung entspricht, dass eine Aufgabe mehrere Schwierigkeitsmerkmale trägt. Wie die Kognitiven Diagnosemodelle, die Kunina-Habenicht et al. im Vergleich zu den korrespondierenden konfirmatorischen Faktorenanalysen untersuchen, ermöglicht auch das angestrebte Schülermodell statistisch belegte, aussagestarke Merkmalsprofile der Schüler.

3.1.4 Mathematik

Die wichtigsten Ergebnisse der internationalen und nationalen PISA-Erhebung fasst regelmäßig das PISA-Konsortium Deutschland zusammen, das die Intention der Aufgaben zur Mathematik wie folgt charakterisiert. „Die Aufgaben umfassen realitätsbezogene und innermathematische Problemstellungen, die übergreifenden mathematischen Ideen zugeordnet werden und wichtige Voraussetzungen für eine weiterführende Auseinandersetzung mit mathematischen Fragen erfassen“ (PISA 2005, S. 5). Schlüsselt man das internationale Mathematik-Rahmenkonzept auf, so erhält man eine *Inhaltsdimension* mit vier übergreifenden Ideen und eine *Prozessdimension* mit acht Teilkompetenzen. Hinzu kommt eine Dimension kognitiver Aktivitäten, die in drei *Kompetenzclustern* beschrieben werden (OECD 2006; Deutsch nach Frey et al. 2010):

Mathematical Content - The Four Overarching Ideas (Inhaltsdimension)

- Space and shape (Raum und Form)
- Change and relationships (Veränderung und Beziehungen)
- Quantity (Quantität)
- Uncertainty (Unsicherheit)

Mathematical Processes - The Competencies (Kompetenzkategorien)

- Thinking and reasoning (mathematisches Denken)
- Argumentation (mathematisches Argumentieren)

- Communication (Kommunizieren)
- Modelling (mathematisches Modellieren)
- Problem posing and solving (mathematisches Problemlösen)
- Representation (Darstellungen verwenden)
- Using symbolic, formal and technical language and operations (mit Symbolen und Formalismen umgehen)
- Use of aid and tools (Hilfsmittel verwenden)

Competency Clusters (Kompetenzcluster, Kognitive Anspruchsniveaus)

- Reproduction (Reproduktion)
- Connections (Verbindungen)
- Reflections (Reflexion)

Entgegen der mehrdimensionalen Anlage erfolgt die statistische Auswertung über ein eindimensionales Modell der Item-Response-Theorie (IRT) und führt zu sechs Kompetenzstufen. „Auf der sechsten und höchsten Kompetenzstufe müssen komplexe Modelle von unvertrauten Problemsituationen gebildet und über mehrschrittige Strategien gelöst werden. Auf der ersten Kompetenzstufe müssen zum Beispiel einfache Rechnungen vollzogen werden, die sich auf Beziehungen zwischen zwei Variablen beziehen“ (PISA 2004, S. 6).

Die Diskrepanz zwischen mehrdimensionaler Anlage und eindimensionaler Auswertung deutet daraufhin, dass zwar national wie international Einigkeit herrscht, ein Aufgabenmodell, welches den PISA-Aufgaben gerecht wird, sei durch mehrere Dimensionen strukturiert. Offenkundig besteht aber im Bereich der mehrdimensionalen Modellierung noch Forschungsbedarf, denn die eindimensionale Analyse verflochtener Dimensionen schöpft die mehrdimensionale Information nicht aus.

Neubrand et al. (2004) beschreiben das nationale PISA-Framework, hinter dem ein Aufgabenmodell mit vier zentralen Eigenschaften und vielfältigen spezifischen Itemmerkmalen steht, zum Beispiel Curriculare Herkunft oder Präsentationsformate (Neubrand et al. 2004, S. 37). Die Merkmale und ihre Wirkung auf die Itemschwierigkeit werden im Abschnitt über Schwierigkeitsmerkmale in der Mathematik vollzählig aufgelistet (S. 43). Die zentralen Eigenschaften beziehen sich auf Wissen und Kontext, Lösungsansatz und Lösungsprozess:

Zentrale Eigenschaften des Aufgabenmodells

- In der Aufgabe angesprochenes Wissen: prozedural oder konzeptuell (Neubrand et al. verweisen auf Hiebert 1986)
- Notwendigkeit eines Mathematisierungs-, Übersetzungs- oder Problemlöseprozesses
- Im Ansatz einschrittig / im Ansatz mehrschrittig (repetitiv, integrativ)
- Kontext: außer- oder innermathematischer Kontext / kontextfreie Aufgabe

Anhand dieser Eigenschaften wurden die internationalen Kompetenzcluster *Reproduction, Connections, Reflections* zu ursprünglich fünf Kompetenzklassen für das nationale Framework ausdifferenziert. In der Weiterentwicklung wurden die Kompetenzklassen zu drei *Typen mathematischen Arbeitens* zusammengefasst, (1) Technische Aufgaben, (2) Rechnerische Modellierungs- und Problemlöseaufgaben und (3) Begriffliche Modellierungs- und Problemlöseaufgaben (vgl. S. 12). Die Typen bezeichnen eine disjunkte Einteilung mathematischer kognitiver Tätigkeiten. Eine Ausnahme bildet der Kontext, der separat aufgeführt wird (Neubrand et al. 2004, S. 39).

Etliche Arbeiten üben auch Kritik an den PISA-Analysen. Stellvertretend sei Meyerhöfer (2004) genannt, der beanstandet, „dass die Lösungshäufigkeiten der mathematischen PISA-Aufgaben allein und allgemein keine Aussage über ihre Schwierigkeit zulassen“ (Meyerhöfer 2004, S. 1). Kompetenzstufen werden in PISA durch Abschnitte auf einer Rasch-Skala repräsentiert, in deren Konstruktion die Lösungshäufigkeiten als Schwierigkeitswerte eingehen. Meyerhöfer zeigt anhand einiger PISA-Aufgaben, dass sie eine Vielzahl verschiedener Lösungswege aufweisen, die unterschiedlich anspruchsvoll sind. Seine Argumentation ist, die eindeutige Zuordnung zu einer Kompetenzstufe sei nicht erlaubt, wenn mehrere Lösungsarten existieren.

Bezug zu dieser Arbeit

Die Kritik Meyerhöfers an der Schwierigkeitsstufung hängt damit zusammen, dass der individuelle Lösungsweg ein Einflussfaktor ist, der nicht allein von der Aufgabe, sondern von der Person abhängt, die sie bearbeitet. Das gleiche gilt für den Schwierigkeitsfaktor Routine, der bereits als personen-, nicht nur aufgabenabhängiger Schwierigkeitsfaktor identifiziert wurde (vgl. S. 24). Dieselbe Aufgabe, die sich dem Schüler leicht zeigt, der das Lösungsschema beherrscht, ist möglicherweise für den Schüler schwieriger, der kreativ einen eigenen Lösungsweg aufspürt. Existieren unterschiedlich komplexe Lösungswege, lässt die Aufgabenstellung kaum eine Vorhersage zu, ob ein einfacher oder komplexer Weg eingeschlagen wird. Das gilt für Informatikaufgaben wie für Mathematikaufgaben.

3.1.5 Naturwissenschaften

Auch für die naturwissenschaftlichen Fächer informiert das PISA-Konsortium Deutschland über die wichtigsten Ergebnisse des internationalen Vergleichs und kommentiert sie aus deutscher Perspektive. Interessant ist, dass mit dem Schwerpunkt Naturwissenschaften im Jahr 2006 „erstmalig motivationale Orientierungen und Einstellungen gegenüber den Naturwissenschaften (Interesse, Wertschätzung von Forschung) systematisch in der Testkonzeption als wichtige Bildungsergebnisse berücksichtigt [wurden]“ (PISA 2007, S. 5). Die internationale Rahmenkonzeption unterscheidet dazu drei Hauptkomponenten des Testgegenstands Scientific

Literacy, *Kompetenzen, Wissen, Einstellungen und Überzeugungen* (OECD 2006, S. 43; Deutsch nach Rönnebeck et al. 2010):

Competencies (Kompetenzen)

- Identifying scientific issues (Fragestellungen erkennen, die mit naturwissenschaftlichen Zugängen bearbeitet werden können)
- Explaining scientific phenomena (naturwissenschaftliche Phänomene beschreiben, vorhersagen und erklären)
- Using scientific evidence (naturwissenschaftliche Evidenz nutzen und interpretieren, um Entscheidungen zu treffen)

Knowledge (Wissen)

- Knowledge of science: Physical systems, Living systems, Earth and space systems, Technology systems (Naturwissenschaftliches Wissen: Physikalische Systeme, Lebende Systeme, Erd- und Weltraumsysteme sowie Technologische Systeme)
- Knowledge about science: Scientific enquiry, Scientific explanations (Wissen über die Naturwissenschaften: Wissenschaftliche Forschung und Erörterung)

Attitudes (Einstellungen und Überzeugungen)

- Interest in science (Interesse an Naturwissenschaften)
- Support for scientific enquiry (Wertschätzung naturwissenschaftlicher Forschung)
- Responsibility towards resources and environment (Verantwortung gegenüber der Umwelt und natürlichen Ressourcen)

Prenzel et al. (2002) untersuchen anhand der nationalen Ergänzungsaufgaben, bei deren Konstruktion bereits die kognitiven Aspekte Faktenwissen anwenden, Grafikverständnis, Mentales Modell nutzen, Schlüsse ziehen und Verbalisieren systematisch berücksichtigt wurden, ob die Aufgabenschwierigkeit anhand solcher Merkmale vorhersagbar ist, „d.h. Aufgaben mit einem ‚schwierigen‘ Merkmal sollten generell seltener gelöst werden“ (Prenzel et al. 2002, S. 124). Pauschal schließen sie, es zeige sich „ein plausibles Muster von kognitiven Hürden oder Hilfestellungen bei der Aufgabenbearbeitung“ (Prenzel et al. 2002, S. 132; siehe auch S. 47 für eine detaillierte Darstellung).

Die Reichweite der Befunde wird jedoch durch gering besetzte Kategorien, eine heterogene Schülerpopulation und die Verletzung der Annahme additiver Komponenten im Regressionsmodell eingeschränkt. „So kann zum Beispiel ein Schüler eine Aufgabe, die eine bestimmte Wissensbasis *und* eine bestimmte Fähigkeit voraussetzt, nur dann lösen, wenn er über beide Komponenten verfügt. Fehlt das Wissen, enthält seine (Nicht-)Lösung keine Information über seine Fähigkeit und umgekehrt“ (Prenzel et al. 2002, S. 132).

Diese Erfahrungen gehen in die Arbeiten von Senkbeil et al. (2005) ein, die nun ein zweidimensionales Facettendesign erproben. Sieben *Facetten kognitiver Kompetenz* werden mit zehn *Inhaltsbereichen* gekreuzt. Die Inhaltsbereiche ergeben sich aus den Basiskonzepten der Fächer Physik, Biologie und Chemie, etwa Elektrizität, Räuber-Beute-Systeme oder Teilchenkonzept (Senkbeil et al. 2005, S. 174ff). Die kognitiven Kompetenzfacetten sind folgende:

Facetten kognitiver Kompetenz

- a) Bewerten
- b) Divergentes Denken (kreative Produktion von Lösungsmöglichkeiten bei naturwissenschaftlichen Problemstellungen)
- c) Umgang mit Graphen (das Entnehmen relevanter Informationen aus einer Graphik oder einem Diagramm)
- d) Konvergentes Denken (aus gegebenen Informationen die richtigen Schlussfolgerungen ziehen)
- e) Umgang mit mentalen Modellen
- f) Sachverhalte verbalisieren
- g) Umgang mit Zahlen

Je Basiskonzept werden sieben Aufgaben entwickelt, eine für jede kognitive Komponente. Die Autoren entwerfen vier konkurrierende Modelle, das sogenannte vollständige Facettenmodell mit zehn Dimensionen, das siebendimensionale Kompetenzenmodell, komplementär dazu das dreidimensionale Fächermodell und schließlich das eindimensionale Rasch-Modell:

„Vollständiges Facettenmodell (10 Dimensionen): Hierbei wird angenommen, dass sich die empirische Datenmatrix durch sieben kognitive Kompetenzen und drei Fächerdimensionen beschreiben lässt.

Kompetenzenmodell (7 Dimensionen): Hier besteht die Vermutung, dass nur die sieben kognitiven Komponenten eigenständige Dimensionen bilden. Die Annahme distinkter Fächerdimensionen ist nicht notwendig.

Fächermodell (3 Dimensionen): Komplementär zum vorangegangenen Modell wird vermutet, dass die Differenzierung nach den naturwissenschaftlichen Fächern ausreicht, um die Daten angemessen abzubilden.

Rasch-Modell (1 Dimension): Bei diesem Modell wird erwartet, dass alle Items des nationalen Naturwissenschaftstests eine gemeinsame und eindimensionale Skala bilden“ (Senkbeil et al. 2005, S. 178).

Die beiden höherdimensionalen Modelle zeigen eine bessere Anpassung an die empirischen Daten. Dabei ist das einfachere 7D-Kompetenzenmodell dem 10D-Facettenmodell vorzuziehen. Die geringerdimensionalen Modelle 1D-Rasch-Modell und 3D-Fächermodell zeigen signifikant schlechtere Datenanpassungen.

Das zeigt „erwartungskonträr, dass zur Beschreibung der empirischen Daten eine Unterscheidung nach den naturwissenschaftlichen Fächern nicht notwendig ist, die Differenzierung in die beschriebenen sieben Komponenten jedoch die Modellalternative mit der besten Datenanpassung darstellt“ (Senkbeil et al. 2005, S. 185).

Um weiteren Aufschluss über die Dimensionalität des Naturwissenschaftstests zu erhalten, schließt sich eine Hauptkomponentenanalyse der sieben kognitiven Komponenten an. Zwei Faktoren werden gebildet. Der erste, *Analytisches Denken*, wird durch die Facetten Umgang mit Zahlen (g), Umgang mit Graphen (c), Umgang mit mentalen Modellen (e), Konvergentes Denken (d) gebildet, der zweite, *Kreativität/Produktivität*, durch die Facetten Sachverhalte verbalisieren (f), Divergentes Denken (b), Bewerten (a).

„Während der erste Faktor diejenigen Kompetenzen umfasst, für die formallogisches und analytisches Denken benötigt werden, erfordern die Komponenten des zweiten Faktors vornehmlich Kreativität und Produktivität. Weiterhin fällt auf, dass alle Komponenten des zweiten Faktors ein offenes Antwortformat aufweisen, d. h. die Lösung dieser Aufgabe ein Mindestmaß an Verbalisierungskompetenz voraussetzt“ (Senkbeil et al. 2005, S. 182f).

Ein Modellgeltungstest des zweidimensionalen Modells mit den Dimensionen Analytisches Denken und Kreativität/Produktivität ergab zwar eine signifikant bessere Anpassung als für das 1D-Rasch-Modell, aber ebenso eine signifikant schlechtere Anpassung als für das 7D-Kompetenzen-Modell.

Schecker und Parchmann (2006) schlagen im Einklang mit der Begriffspraxis für die Beschäftigung mit Kompetenzmodellen vor, deskriptive Intention ausdrücklich von normativer Intention abzuheben. Sie unterscheiden Kompetenzstrukturmodelle von Kompetenzentwicklungsmodellen, die Annahmen über die Herausbildung von Kompetenzstrukturen machen, und definieren Strukturmodelle wie folgt.

„1) Ein normatives Kompetenzstrukturmodell ist das Gefüge einer nach Dimensionen (z. B. Kompetenzbereiche, Kompetenzausprägungen) gegliederten Beschreibung der (kognitiven) Voraussetzungen, über die ein Lernender verfügen soll, um Aufgaben und Probleme in einem bestimmten Gegenstands- oder Anforderungsbereich lösen zu können. [Schecker und Parchmann verweisen auf Klieme et al. 2003, S. 74]

2) Ein deskriptives Kompetenzstrukturmodell ist das Gefüge einer nach Dimensionen gegliederten Beschreibung eines ‚typischen‘ Musters (kognitiver) Voraussetzungen, mit denen man das Verhalten von Lernenden beim Lösen von Aufgaben und Problemen in einem bestimmten Gegenstands- oder Anforderungsbereich rekonstruieren bzw. beschreiben kann“ (Schecker und Parchmann 2006, S. 47).

Die Autoren zeichnen die Perspektive, normative Modelle, die zumeist fachsystematisch aufgebaut sind, mit deskriptiven Modellen zur Deckung zu bringen, indem die postulierten Strukturen mit den empirischen Befunden abgeglichen werden. Sie gehen von den nationalen Bildungsstandards für die Naturwissenschaften aus und erarbeiten ein fünfdimensionales Rahmenmodell. Die Dimensionen lauten Inhaltsbereich/Basiskonzept, Prozess/Handlung, Kontext, Ausprägung, Kognitive Anforderung. Die letzte, die Anforderungsdimension, geht auf das 7D-Kompetenzenmodell (vgl. S. 32) zurück. So wird das Rahmenmodell als „erste Näherung eines deskriptiven Modells, das einerseits noch von normativen Modellen geprägt ist, andererseits jedoch empirische Befunde über Schülerkompetenzen berücksichtigt“ verstanden (Schecker und Parchmann 2006, S. 58).

Bezug zu dieser Arbeit

In der Begleitforschung zu PISA-Naturwissenschaften ist eine klare Fokussierung auf den Zusammenhang zwischen Aufgabenmerkmalen und Aufgabenschwierigkeit erkennbar. Modelle der Item-Response-Theorie (IRT), Regressionsanalysen sowie Faktorenanalysen werden exploratorisch angewendet.

Prenzel et al. beschreiben die Problematik gering besetzter Kategorien, die sich auch im Datenmaterial des Informatikwettbewerbs abzeichnet, das dieser Untersuchung zugrunde liegt. Wenn es schwerfällt, Merkmale zu isolieren, etwa weil die Aufgabenzahl zu gering ist, dann bleiben die Befunde bezüglich einzelner Merkmale zwangsläufig unscharf.

Senkbeil et al. nähern sich der Problematik, indem sie für die in Frage kommenden Dimensionenmodelle vergleichen, wie gut diese zu den Daten aus PISA passen. Die Modellgeltungstests, mit denen sie die Modell-Hypothesen 1D, 3D, 7D, 10D (vgl. S. 32) überprüfen, sind geeignet, Modelle auf ihre Anpassung an empirische Daten zu überprüfen und Modellvarianten zu vergleichen.

Nach der Begriffsdefinition von Schecker und Parchmann handelt es sich bei den hier entwickelten Modellvarianten um deskriptive Strukturmodelle, die *typische* kognitive Muster abbilden, im Gegensatz zu den verbreiteten normativen Kompetenzstrukturmodellen, die *erwünschte* kognitive Muster festschreiben.

3.1.6 Andere Fächer und Fächerübergreif

Abgerundet wird der Querschnitt durch die Forschungsszenarie durch einen Blick auf weitere Fächer beziehungsweise fächerübergreifende Themen, namentlich den dritten PISA-Schwerpunkt Leseverständnis, fachunabhängiges Problemlösen und Kompetenzmodellierung in der Berufsbildung.

Leseverständnis

Die PISA-Modellierung des Leseverständnisses lässt sich nur begrenzt in Einklang mit der Fragestellung dieser Untersuchung bringen, daher wird lediglich der Vollständigkeit halber die Rahmenkonzeption skizziert. Leseverständnis bezieht das *Textformat* ein, ob kontinuierlich, zum Beispiel eine Erzählung, oder nicht-kontinuierlich, zum Beispiel eine Landkarte, sowie die *Lese-Situation* und den *Lese-Prozess* (OECD 2006; Deutsch nach Naumann et al. 2010):

Situations (Situationen und Kontexte)

- Reading for private use (private Zwecke, z. B. ein persönlicher Brief)
- Reading for public use (öffentliche Zwecke, z. B. ein offizielles Dokument)
- Reading for work (berufsbezogene Zwecke, z. B. ein Bewerbungsformular)
- Reading for education (bildungsbezogene Zwecke, z. B. ein Lehrbuchtext)

Processes (Aspects) (Aspekte des Lesens)

- Retrieving Information (Informationen suchen und extrahieren)
- Interpreting (textbezogenes Kombinieren und Interpretieren)
- Reflection and evaluation (Reflektieren und Bewerten)

Problemlösen

Zusätzlich zu den fachbezogenen Kompetenzen werden in PISA fächerübergreifende Schlüsselkompetenzen wie die Problemlösekompetenz erfasst. Wird die Kontextspezifität aufgeweicht, ist die Abgrenzung gegenüber anderen Leistungskonzepten schwierig. Nach Hartig und Klieme (2006) ist Kompetenz im Unterschied zu Intelligenz kontextualisiert, das heißt, sie tritt in spezifischen Situationen zu Tage. Außerdem ist sie durch Erfahrung lernbar und ihre Binnenstruktur ergibt sich aus Situationen und Anforderungen, während die Binnenstruktur von Intelligenz sich aus grundlegenden kognitiven Prozessen ergibt. „Je breiter ein Bereich von Situationen und Aufgaben für eine Schlüsselkompetenz definiert wird, desto schwieriger wird die begriffliche Abgrenzung von dem, was typischerweise unter allgemeiner Intelligenz verstanden wird (Hartig und Klieme 2006, S. 130).“

Auch die Trennung zwischen fächerübergreifendem und fachlichem Problemlösen ist schwierig. Fleischer et al. (2010) erkunden die Struktur der Problemlösekompetenz, indem sie PISA-Aufgaben zum analytischen Problemlösen und zur Mathematik vergleichen. Problemlöseaufgaben enthalten als Aufgabenstellung häufiger eine Aufforderung, Mathematikaufgaben eine Frage. Das kognitive Anforderungsniveau der Denk- und Handlungsanweisung ist bei ersteren höher. Problemlöseaufgaben sind sprachlogisch komplexer, während Mathematikaufgaben stärker formalisiert sind. Sie weisen meist einen Kontext mit persönlichem Bezug auf, erfordern nur Grundkenntnisse curricularen Wissens und sind dennoch kognitiv komplexer als Mathematikaufgaben. Handlungswissen steht gegenüber Sachwissen im Vordergrund. Problemlösen stellt höhere Anforderungen an das Erkennen relevanter Bedingungen (Fleischer et al. 2010, S. 244f).

Berufsbildung

Laut Zlatkin-Troitschanskaia und Seidel (2011) ist in der beruflichen Bildung als Kompetenzauffassung neben der kognitiven die handlungsorientierte Typologie Methoden-, Sozial- und Selbstkompetenz verbreitet. Gelegentlich wird unter Kompetenz außer Fähigkeit noch Zuständigkeit verstanden, dann lautet die Typologie Handlungsfähigkeit, Handlungsbereitschaft und Zuständigkeit (Zlatkin-Troitschanskaia und Seidel 2011, S. 221f).

Nickolaus (2011) vermittelt einen breiten Überblick über den Forschungsstand der Kompetenzmodellierung in der beruflichen Bildung. Eine typische Frage ist, „inwieweit auf der Basis computersimulierter Arbeitsprozesse und -systeme valide Kompetenzabschätzungen möglich werden“ (Nickolaus 2011, S. 331). Unter der Überschrift Strukturmodellierung stehen teilweise kontroverse Ansätze zur Ausdifferenzierung der Kompetenzdimensionen nebeneinander. Unter der Überschrift Niveaumodellierung sind Beiträge versammelt, in denen Leistungsstufen identifiziert werden. Die systematische Variation der potentiellen Schwierigkeitsmerkmale ist im Kontext beruflicher Bildung besonders erschwert, da sich nicht unbedingt authentische Aufgabenstellungen ergeben (Nickolaus 2011, S. 337).

Bezug zu dieser Arbeit

In der Rahmenkonzeption für das Leseverständnis sind eine Prozessdimension sowie mit dem Textformat eine Repräsentationsdimension und mit der Lese-Situation eine Kontextdimension enthalten. Diese Konstellation kristallisiert sich auch im mathematisch-naturwissenschaftlichen Bereich als häufig auffindbare Grundstruktur heraus, wobei der Kontext nicht unbedingt als Dimension auftritt, sondern als fachliche Situation, in der das Modell zur Anwendung kommt. Fällt wie beim fächerübergreifenden Problemlösen die Fachspezifität weg, erschwert das die Abgrenzung von Kompetenz gegenüber anderen Leistungskonzepten. Entsprechend wichtig ist es, in die Definition der Informatikkompetenz den Kontext aufzunehmen. Schließlich ergibt der Blick auf die Kompetenzmodellierung in der beruflichen Bildung, dass die Aktivitäten dort weitgehend parallel, gleichwohl unvernetzt mit den allgemeinbildenden Fächern verlaufen.

3.2 Ansätze in der Informatik-Didaktik

Das Fach Informatik steht in Deutschland erst am Anfang der systematischen Bestimmung von Lernvoraussetzungen und Leistungsanforderungen. Die Position der informatischen Bildung inmitten allgemeiner Bildungsziele ist nicht geklärt. Eine Vereinbarung über Bildungsstandards für den Mittleren Schulabschluss, wie für andere Fächer in der Kultusministerkonferenz (KMK) beschlossen, gibt es bislang nicht (KMK 2004a; 2005). Allein die Einheitlichen Prüfungsanforderungen in der Abiturprüfung Informatik stellen ein länderweit verbindliches Regelwerk

dar (KMK 2004b). Der Unterricht profitiert mehr vom Engagement der Lehrpersonen mit Expertise im Fach Informatik, die Unterrichtskonzepte veröffentlichen und Praxisbeispiele ausarbeiten, als von einer reichen didaktischen Tradition und Aufgabekultur wie in der Mathematik. Dementsprechend ist die Situation der Forschung nicht mit dem über Jahrzehnte gewachsenen und geförderten Erkenntnisstand anderer Fächer vergleichbar, wenn es auch mittlerweile einige Forschergruppen gibt, die sich mit dem Thema Informatikkompetenz beschäftigen.

3.2.1 Lernziel-Taxonomien

Blooms Taxonomie kognitiver Lernziele prägt viele internationale Beiträge zur systematischen Kompetenzerfassung, vor allem im englischsprachigen Raum. Inhaltlich liegt der Schwerpunkt in der Computer Science Education (Informatik-Bildung) curriculumsbedingt meist auf der Programmierung. Zum Beispiel nimmt Lahtinen (2007) eine Typisierung von Programmieranfängern nach Blooms Lernzielstufen vor; Starr et al. (2008) beschreiben einen Ansatz, Blooms originale Taxonomie zur Curriculumsauslegung und Aufgabengestaltung heranzuziehen. Auch Neuentwürfe für Taxonomien werden in der Computer Science Education vorgestellt. So entwerfen Biggs und Collis (1982; vgl. S. 24) die SOLO-Taxonomie mit den fünf Stufen Prestructural, Unistructural, Multistructural, Relational und Extended Abstract, die von Sheard et al. (2008) zur Klassifizierung von Prüfungsantworten eingesetzt wird; Bower (2008) führt eine zehnstufige Taxonomie zur Aufgabekategorisierung ein, die von Declarative über Comprehension, Debugging, Prediction, Provide-an-example, Provide-a-model, Evaluate, Meet-a-design-specification bis Solve-a-problem und Self-reflect tasks reicht.

Fuller et al. (2007) nennen als intrinsische Charakteristika der Disziplin Computer Science Problemlösen, Modellierung, Wissensrepräsentation, Effizienz der Problemlösung, Abstraktion/Modularität, Neuartigkeit/Kreativität, Kategorisierung, Kommunikationsgeschick mit Experten anderer Gebiete und Übernahme bewährter Praktiken in den Software Engineering Prozess:

„Intrinsic characteristics of computer science:

- Problem solving
- Domain modelling
- Knowledge representation
- Efficiency in problem solving
- Abstraction/modularity
- Novelty/creativity
- Categorization
- Communication skills with experts in other domains
- Adoption of good practice in software engineering“ (Fuller et al. 2007, S. 163, im Original ohne Hervorhebung und Aufzählungszeichen)

Für die spezifischen Ansprüche der Computer Science, insbesondere des Programmierens, empfehlen Fuller et al. eine zweidimensionale Adaption der Bloom'schen Taxonomie, die Matrix Taxonomie (Fuller et al. 2007, S. 163f). Sie ordnen die Prozessstufen der revidierten Taxonomie in zwei Dimensionen an, Apply und Create, ursprünglich Stufen 3 und 6, entlang einer Dimension *Producing* und Remember, Understand, Analyze, Evaluate, ursprünglich Stufen 1, 2, 4, 5, entlang einer Dimension *Interpreting*. In der entstehenden Matrix lassen sich beispielsweise Lernpfade abbilden.

Bezug zu dieser Arbeit

Gemeinsam ist den eher pragmatisch als theoretisch begründeten Ansätzen, dass sie auf ihre Praktikabilität überprüft wurden, aber nicht auf ihre empirische Datenpassung. Ihre Vielzahl zeugt davon, dass in der Informatik-Bildung ein Bedarf an Modellen zur kognitiven Taxierung von Lernenden, Curriculumsinhalten, Prüfungsfragen und Prüfungsantworten besteht. Lerntypen und Lernzielhierarchien sind nicht fächerübergreifend konstant. Ein wichtiges Lernziel in der Informatik ist „the ability to develop artifacts“ (Fuller et al. 2007, S. 162f) – Software ist ein solches Artefakt. Ein wichtiges Lernprinzip ist „learning through doing“ (ebd.). Problemlösen, Effektivität und Effizienz sind Kernanliegen professioneller Informatiker, Abstrahieren eine essentielle Fähigkeit. Das bedeutet, ein Modell der Anforderungen und Fähigkeiten in der Informatik zeichnet sich durch besondere Lernziele wie Artefakte entwickeln und Kernfähigkeiten wie Abstrahieren aus.

3.2.2 Computer Literacy

Bildungsstudien wie PISA konzentrieren sich auf die Kernfächer. Begleitend wird unter anderem die Computer Literacy (Computer-Bildung) erfasst. Wirth und Klieme (2002) interpretieren die Ergebnisse aus PISA 2000 zur Computer Literacy für Deutschland so, dass die Beherrschung des Computers eine anerkannte Kulturtechnik ist, mit Computerbeherrschung allerdings die souveräne Bedienung der Geräte und nicht die Beherrschung informatischer Kalküle und Methoden gemeint ist. Werden die Computer-Nutzung, das Interesse am Computer und selbst eingeschätzte computerbezogene Kenntnisse und Fähigkeiten fünfzehnjähriger Schüler beleuchtet, fällt auf, dass bei deutschen Schülern im internationalen Vergleich ein hohes Interesse am Computer besteht, die eigenen Fähigkeiten indessen gering eingeschätzt werden. Meist wird der Computer zum Zeitvertreib benutzt, seltener als Arbeitsmittel in der Schule.

Bezug zu dieser Arbeit

Computerbeherrschung als Fertigkeit im Umgang mit dem Computer steht in dieser Arbeit nicht im Vordergrund. Vielmehr interessieren die kognitiven Vor-

aussetzungen, die zum Verständnis und zielgerichteten Einsatz informatischer Konzepte und Methoden befähigen. In dieser Hinsicht sind die Ergebnisse aus PISA 2000 zur Computer Literacy weniger bedeutsam.

3.2.3 Bildungsstandards

In Deutschland beschloss die Kultusministerkonferenz im Jahr 1989 die „Einheitlichen Prüfungsanforderungen in der Abiturprüfung Informatik“, kurz EPA Informatik (heute in der Fassung von 2004), um vergleichbare Abiturergebnisse sicherzustellen (KMK 2004b). Dort sind Qualifikationen und fachliche Inhalte sowie taxonomische Anforderungsbereiche vereinbart, erstere gliedern sich in *Kompetenzen* und *Inhalte*, letztere in drei *Anforderungsbereiche*¹:

Fachliche und methodische Kompetenzen

- Erwerb und Strukturierung informatischer Kenntnisse
- Kennen und Anwenden informatischer Methoden
- Kommunizieren und Kooperieren
- Anwenden informatischer Kenntnisse, Bewerten von Sachverhalten und Reflexion von Zusammenhängen (KMK 2004b, S. 7ff)

Fachliche Inhalte

- Grundlegende Modellierungstechniken
- Interaktion mit und von Informatiksystemen
- Möglichkeiten und Grenzen informatischer Verfahren (ebd.)

Anforderungsbereiche

- Wiedergabe (Anforderungsbereich I)
- Anwendung (Anforderungsbereich II)
- Problemlösung (Anforderungsbereich III) (KMK 2004b, S. 14ff)

Die Anforderungsbereiche Wiedergabe, Anwendung und Problemlösung sind über alle Fächer hinweg im Wesentlichen gleich und werden durch fachspezifische Beispiele konkretisiert. In der Abiturprüfung und in den Bildungsstandards der Kultusministerkonferenz, die für etliche Fächer bereits vorliegen, aber noch nicht für das Fach Informatik, fungieren sie als Kompetenzniveaus, da es noch keine fachdidaktisch fundierten und empirisch abgesicherten Kompetenzstufen gibt (KMK 2004b, S. 14ff).

„Für Aussagen über die Angemessenheit, Qualität und Komplexität der Anforderungen, die mit den Aufgabenbeispielen verbunden sind, stellen die Anforderungsbereiche eine Orientierung dar, in der sich die Leistungen von Schülerinnen und Schülern erfahrungsgemäß bewegen.

¹Die Spezifikation der Anforderungsbereiche befindet sich in Anhang B (S. 197).

Anforderungsbereiche resultieren somit nicht aus empirisch validierten Testverfahren, sondern aus der beruflichen Erfahrung von Lehrkräften und einschlägigen Aufgabenformaten aus bereits vorhandenen Testmaterialien. Vor diesem Hintergrund besitzen die Anforderungsbereiche vorläufigen Charakter“ (KMK 2005, S. 17).

Die EPA betreffen die Abiturprüfung am Ende der Sekundarstufe II. Für die Sekundarstufe I verabschiedete die Gesellschaft für Informatik e. V. (GI) als Interessensverband der Informatik im Jahr 2008 eine Empfehlung für „Grundsätze und Standards für die Informatik in der Schule“. Sie sind durch verzahnte *Inhalts- und Prozessbereiche*² strukturiert, die für die Jahrgangsstufen 5-7 und 8-10 separat weiter aufgeschlüsselt werden (GI 2008, S. 14ff):

Inhaltsbereiche

- Information und Daten
- Algorithmen
- Sprachen und Automaten
- Informatiksysteme
- Informatik, Mensch und Gesellschaft

Prozessbereiche

- Modellieren und Implementieren
- Begründen und Bewerten
- Strukturieren und Vernetzen
- Kommunizieren und Kooperieren
- Darstellen und Interpretieren

Bezug zu dieser Arbeit

Die Anforderungsbereiche der EPA Informatik, die in der Abiturprüfung als Kompetenzniveaus fungieren und die Prozessbereiche der GI-Empfehlung für Standards in der Schulinformatik werden in die Ausgangsmenge der Schwierigkeitskriterien für die Modellierung informatischer Kompetenz aufgenommen. Die Anforderungsbereiche beruhen auf Erfahrung. Sie sind zwar nicht als Abstufung empirisch validiert, haben aber große praktische Bedeutung erlangt und sind in allen Schulfächern richtungsweisend für die Ausbalancierung des Prüfungsniveaus. In der vorliegenden Untersuchung erweisen sie sich als tragende Kriterien. Die Prozessbereiche beruhen ebenfalls mehr auf der Erfahrung von Fachdidaktikern und Schulpraktikern als auf systematischer theoretischer und empirischer Forschung. Die GI-Empfehlung überbrückt das Fehlen offizieller und verbindlicher Bildungsstandards im Fach Informatik und findet so Eingang in die Kriterienauswahl dieser Untersuchung.

²Die Spezifikation der Inhalts- und Prozessbereiche befindet sich in Anhang B (S. 199f).

3.2.4 Kompetenzmodelle

Den Standards gehen theoretische Entwürfe voran, die die informatische Bildung insgesamt adressieren. Friedrich (2003) stützt sich auf die Empfehlungen der GI (2000) für ein Gesamtkonzept zur informatischen Bildung an allgemein bildenden Schulen. Die vier Leitlinien informatischer Bildung, *Interaktion* mit Informatiksystemen, *Wirkprinzipien* von Informatiksystemen, *Informatische Modellierung* und *Wechselwirkungen* zwischen Informatiksystemen, Individuum und Gesellschaft, bezieht er auf die fünf Kompetenzstufen Bedienung von Informatik-anwendungen, Benutzung von Informatiksystemen, Kenntnis fachsystematischer Grundlagen, Verständnis von Konzepten der Informatik sowie Entwicklung und Bewertung von Informatiksystemen. Puhlmann (2003) schlägt eine Definition informatischer Literalität vor, die er durch die drei Kompetenzklassen Anwendung, Gestaltung und Entscheidung konkretisiert. Magenheim (2005) entwirft ein Kompetenzmodell mit den Dimensionen Informatiksysteme als Medium (Usage of Media Functions of the Informatics System), Anwendung (Level of Application) und Informatiksystemverständnis (Level of System Comprehension).

Weitere Ansätze beziehen sich auf bestimmte Inhaltsbereiche. Brinda (2004) entwickelt für die strukturierte Aufgabengestaltung zum Objektorientierten Modellieren (OOM) eine Klassifizierung nach Fachkern, Gegenstand und Aufgabentyp mit einer Einordnung der Aufgabentypen in Blooms Taxonomie. Schulte und Brinda (2005) konsolidieren ihre Ergebnisse aus unterschiedlichen Zugängen zum OOM, lernbezogen beziehungsweise fachbezogen, zu einem Kompetenzmodell informatischen Modellierens mit den drei Bereichen Grundverständnis zum OOM – Überblickswissen, Modellierungskompetenz und Bewertungskompetenz. Freischlad (2009) nimmt für den Unterrichtsinhalt „Internetworking“ drei unabhängige Kompetenzdimensionen an und gestaltet Aufgaben zur Erfolgskontrolle, deren Merkmale er innerhalb der drei Dimensionen Inhalt, Prozess und Darstellung variiert:

- „1. Inhaltsdimension: Zuordnung des notwendigen Wissens mit Aufgabenklassen und Beschreibung des Anforderungsniveaus durch Wissenskategorie und Anzahl der vernetzten Aufgabenklassen.
2. Prozessdimension: Bezug zum Operator als Teil der Fragestellung und Beschreibung des Niveaus durch Lernzielstufe.
3. Darstellungsdimension: Bezug zum Stimulus und Beschreibung des Anforderungsniveaus durch Repräsentationsstufe“ (Freischlad 2009, S. 77).

Stechert (2009) reflektiert in seiner Dissertation Sichten auf die Kompetenzentwicklung mit Informatiksystemen im Informatikunterricht. Unter der Leitung von Magenheim, Schaper und Schubert (2009) erarbeiten Informatiker und Psychologen im interdisziplinären Projekt „Entwicklung von qualitativen und

quantitativen Messverfahren zu Lehr-Lern-Prozessen für Modellierung und Systemverständnis in der Informatik – kurz MoKoM“ ein vielschichtiges Kompetenzrahmenmodell. Nelles, Rohde und Stechert (Nelles et al. 2010) gehen dafür von vier Dimensionen aus. Die ersten beiden, Aufgabenbereiche und Nutzung informatischer Sichten, werden durch die Anforderungsniveaus der EPA zusätzlich gegliedert. Anforderungen an den Umgang mit Komplexität und Nicht-kognitive Komponenten definieren die dritte und vierte Dimension. Im finalen MoKoM-Modell wurden die anfänglich vier Dimensionen in fünf Dimensionen überführt, Systemanwendung, Systemverständnis, Systementwicklung, Umgang mit Systemkomplexität und Nicht-kognitive Fähigkeiten (Linck et al. 2013).

Aufschlussreiche Beiträge zur Modellierung von Schwierigkeit und Fähigkeit leisten Steinert (2007) und Kujath (2007), Steinert von einer theoretischen Warte, Kujath dagegen aus der qualitativen Beobachtung von Studierenden heraus. Steinert konstruiert zu Prüfungsaufgaben Lernzielgraphen, die die revidierte Taxonomie von Anderson und Krathwohl et al. (2001) und eine Vorrangrelation für Lernziele implementieren. Vorrang hat ein Lernziel, falls es die Voraussetzung für ein anderes Lernziel bildet. Die Aufgabenschwierigkeit entspricht dann der maximalen Stufe des Lernziels und seiner Vorgänger im Graphen.

Kujath kontrastiert das Problemlöseverhalten von sogenannten Hochleistern und sogenannten Niedrigleistern. Studierende, die als Bundessieger aus dem Bundeswettbewerb Informatik³ hervorgingen, vertreten die Hochleister. Informatikstudenten mit mittleren bis schlechten Studienleistungen in theoretisch-formal ausgerichteten Fächern vertreten die Niedrigleister. Die leistungsstarken Problemlöser fallen durch schnelles und sicheres Problemverständnis auf sowie durch eine klare Trennung in Teilprobleme, intensive Problemanalyse, ein hohes Abstraktionsniveau, ausgeprägte fundamentale Ideen der Informatik und die Herausbildung der Frage: Was passiert beim Übergang von einem n zum nächsten? (Kujath 2007, S. 67).

Bezug zu dieser Arbeit

Die bisherigen Entwürfe gelten noch öfter den Niveaustufen als der dimensionalen Struktur der Kompetenz. Erst seit wenigen Jahren werden Strukturen modelliert, in denen Aufgaben gleichzeitig in mehreren Dimensionen bewertet werden. Insgesamt beziehen sich die Modelle meist auf bestimmte Fachgebiete, etwa Programmieren oder Objektorientiertes Modellieren.

Für alle Ansätze gilt, dass die Modelle noch kaum auf einem entwicklungspsychologischen oder lerntheoretischen Fundament errichtet werden. Die Entwicklung verläuft entlang der Fachsystematik. Evaluieren werden die Entwürfe etwa durch probeweises Einordnen von Unterrichtszielen oder Aufgaben in das

³Der Bundeswettbewerb Informatik ist Kern der Initiative Bundesweit Informatiknachwuchsförderer (www.bundeswettbewerb-informatik.de; www.bwinf.de), zu der auch der Informatik-Biber gehört.

Modell oder durch Expertenabstimmung. Es fehlen Untersuchungen zum Erklärungsgehalt der Modelle, das heißt zu ihrer Eignung, empirische Schülerdaten in ihrer Dimensionalität und Stufung zu beschreiben.

3.3 Schwierigkeitsmerkmale

In den vergangenen Abschnitten lag der Fokus auf den Arbeiten, die sich mit dem Konstrukt Kompetenz in Gänze auseinandersetzen, einschließlich der Begriffs- und Methodendiskussion. Eng mit der Modellierung von Kompetenz in ihrer dimensional Struktur verwoben ist das Thema „Schwierigkeitsmerkmale“ – die Aufgabenmerkmale, die die Dimensionen charakterisieren. Diese Aufgabenmerkmale rücken jetzt in das Blickfeld. Das Interesse gilt ihrer Funktion, Niveaustufen zu veranschaulichen, und ihrer Relevanz als Prädiktoren der Aufgabenschwierigkeiten, „diese wiederum sind konstruktionsbedingt variabel. Man kann ‚denselben‘ Aufgabeninhalt so in Form einer Testaufgabe bringen, dass diese eher schwer oder eher leicht zu lösen ist“ (Rost 2004b, S. 663).

Ohne Unterscheidung wurde bis zu dieser Stelle überwiegend der Begriff Aufgabe, seltener Item benutzt. Laut Hartig und Jude (2007) sind Aufgabe und Item identisch, sofern nur eine einzelne Antwort verlangt wird.

„In einem Leistungstest bezeichnet eine *Aufgabe* eine inhaltlich zusammenhängende, nicht aufteilbare Einheit, in der ein Problem oder eine bestimmte Einheit von Material (z. B. ein Lesetext) vorgegeben wird. Ein *Item* hingegen stellt die kleinste Analyseeinheit eines Tests dar, die bei der Auswertung berücksichtigt wird. Typischerweise ist ein Item dabei gleichbedeutend mit einer einzelnen Frage oder Teilfrage. Eine Testfrage kann also aus mehreren Items bestehen. Dies muss aber nicht der Fall sein. Wenn zu einer Aufgabe nur eine einzelne Antwort zu geben ist, sind Aufgabe und Item identisch“ (Hartig und Jude 2007, S. 19).

Generell wird in dieser Arbeit dem Begriff Aufgabe der Vorzug gegeben, wenn auch im Kapitel zum Stand der Forschung die Wortwahl den vorgestellten Beiträgen folgt, im Kontext der Item-Response-Theorie also von Items gesprochen wird.

3.3.1 Mathematik

Ein ausführlicher Katalog von Aufgabenvariablen des mathematischen Problemlösens findet sich bei Goldin und McClintock (1979) in einem Gutachten zur mathematischen Bildungsforschung. Kulm (1979) teilt darin die Aufgabenvariablen in Syntaxvariablen, Inhalts- und Kontextvariablen, Strukturvariablen und Variablen des heuristischen Lösungsprozesses ein.

„The categories are: (a) variables which describe the problem *syntax*, (b) variables which characterize the problem’s mathematical *content* and non-mathematical *context*, (c) variables which describe the *structure* of the problem, and (d) variables which characterize the *heuristic processes* evoked by the problem“ (Kulm 1979, S. 15).

Zum Beispiel enthalten Syntaxvariablen die durchschnittliche Wortlänge (Barnett 1979, S. 32), Inhaltsvariablen das Auftreten von Schlüsselwörtern wie „größer als“ oder das Aufgabengebiet, z. B. Algebra oder Geometrie (Webb 1979, S. 78). Kontextvariablen halten fest, ob der Problemkontext vertraut oder unvertraut ist (Webb 1979, S. 96), Strukturvariablen zählen die Zustände im Zustandsraum des Problems (Goldin 1979, S. 165). Der heuristische Lösungsprozess wird durch „Analogie“ oder „Rückschluss“ beschrieben (McClintock 1979, S. 174). Die letzte Kategorie wird von den anderen abgesetzt, da sie nicht unbedingt die Schwierigkeit beschreibt, obwohl denkbar sei, dass heuristisch lösbare Probleme weniger schwierig sind als solche ohne heuristische Lösung. Die Vermutung, dass die übrigen Kategorien eine Komplexitätshierarchie bilden, wird empirisch nicht bestätigt (Kulm 1979, S. 19).

Dem heutigen Forschungsinteresse gemäß, das sich dem Kompetenzerwerb zuwendet, sind Variablen des Problemlösens nicht nur auf Syntax, Struktur, Inhalt, Kontext, Heuristiken des Problems bezogen, sondern mehr auf den Problemlöser und die Fähigkeiten und Fertigkeiten, die der Löseprozess anspricht. So werden in PISA-Mathematik 2006 die internationalen Stufen der Itemschwierigkeit und damit die Kompetenzniveaus durch die Anforderungen Interpretation und Reflexion, Repräsentationsfertigkeiten, mathematische Fertigkeiten und mathematisches Schlussfolgern gestützt (OECD 2006, S. 113):

Internationale Berichtsmerkmale

- The kind and degree of interpretation and reflection required
- The kind of representation skills required
- The kind and level of mathematical skills required
- The kind and degree of mathematical argumentation required

Verglichen mit den internationalen Berichtsmerkmalen ist das nationale Aufgabenmodell in Mathematik (vgl. S. 29) höher aufgelöst, das heißt, die Merkmale sind weiter ausdifferenziert (Neubrand et al. 2004, S. 37):

Nationales Aufgabenmodell, Merkmale

- Anzahl der bei der Lösung benötigten Zwischenergebnisse
- Curriculare Herkunft
- Präsentationsformate
- Spezifische mathematische Tätigkeiten wie Beweisen, Begründen, Argumentieren, Modellieren, Interpretieren usw.

- Anzahl und Qualität der zur Bearbeitung der Aufgabe erforderlichen Grundvorstellungen (Neubrand et al. verweisen auf vom Hofe 1995)
- Sprachliche und/oder logische Komplexität (Neubrand et al. verweisen auf Cohors-Fresenborg 1996)
- Umkehraufgabe oder direkte Aufgabe
- Existenz multipler Lösungswege

Aus den nationalen und internationalen Aufgabenmerkmalen stellen Neubrand et al. (2002) sechs Variablen zusammen und untersuchen mit einer Regressionsanalyse, welche Eigenschaften als Prädiktoren einen Einfluss auf die Schwierigkeit in den drei Typen mathematischen Arbeitens ausüben, repräsentiert durch Technische Aufgaben, Rechnerische und Begriffliche Modellierung (Neubrand et al. 2002, S. 105ff; zu den Typen mathematischen Arbeitens vgl. auch S. 30):

- a) Komplexität der Modellierung mit den Ausprägungen Reproduktion (Reproduction), Verknüpfung (Connections), Verallgemeinerung (Reflections) (Internationale Kompetenzklassen)
- b) Curriculare Wissensstufe mit den Stufen Grundkenntnisse, einfaches Wissen der Sekundarstufe I, anspruchsvolles Wissen der Sekundarstufe I
- c) Kontext, außermathematisch, innermathematisch oder ohne Kontext
- d) Offenheit der Mathematisierung, das heißt, es existieren vielfältige Lösungen
- e) Umfang der Verarbeitung
- f) Argumentieren

Die Schwierigkeitsparameter werden im Rasch-Modell aus dem beobachteten Schülerverhalten geschätzt. Generell haben die Komplexität der Modellierung (a) und die curriculare Herkunft (b) den größten Einfluss. Die technischen Aufgaben werden hauptsächlich mit ansteigenden curricularen Anforderungen (b) schwieriger. Die Modellierungsaufgaben, ob rechnerisch oder begrifflich, reflektieren die internationalen Kompetenzklassen (a). Auf die rechnerischen Modellierungsaufgaben wirken außerdem die curriculare Wissensstufe (b) und der Umfang der Verarbeitung (e), in den begrifflichen Modellierungsaufgaben erzeugen der innermathematische Kontext (c) und die multiple Lösbarkeit (d) Schwierigkeit. Die Autoren räumen ein, dass der Zusammenhang nicht eindeutig ist, sondern durch eine Vielzahl anderer Einflüsse moderiert wird (Neubrand et al. 2002).

Zur konzeptionellen Bandbreite der zahlreichen schwierigkeitsbeeinflussenden Merkmale, die aus Aufgabenanalysen im Rahmen von PISA resultieren, nehmen Cohors-Fresenborg et al. (2004) Stellung. „Die Merkmalsbezeichnungen lassen bereits erkennen, dass ihnen sehr unterschiedliche Konzeptionen zu Grunde liegen. Beispielsweise entstammen Merkmale wie ‚curriculare Wissensstufe‘, ‚Inhaltsbereich‘ oder ‚Kontext‘ offensichtlich einer traditionellen, andere wie ‚multiple Lösbarkeit‘ oder ‚Grundvorstellung‘ einer neueren stoffdidaktischen Diskussion“ (Cohors-Fresenborg et al. 2004, S. 110). Sie selber gehen demgegenüber von den

Denkprozessen aus, die stoff- und aufgabenübergreifend bei der Bearbeitung ablaufen und stellen ausschließlich kognitionsorientierte Aufgabenmerkmale vor:

- „Das Merkmal ‚Sprachlogische Komplexität‘ erfasst Anforderungen beim Identifizieren und Verstehen von relevanten Informationen eines (durch logische Struktur und sprachliche Verflechtung geprägten) Aufgabentextes [...]“ (Cohors-Fresenborg et al. 2004, S. 114).
- „Das Merkmal ‚Kognitive Komplexität‘ erfasst Anforderungen an Ausmaß, Intensität und Vielschichtigkeit von Denkvorgängen beim Lösen einer Aufgabe [...]“ (ebd., S. 115).
- „Das Merkmal ‚Formalisierung von Wissen‘ erfasst Fähigkeiten des Abstrahierens und Formalisierens einerseits sowie des Erfassens und abstrakten Vorstellens formaler Ausdrücke (Terme, Gleichungen, Funktionen) andererseits [...]“ (ebd., S. 118).
- „Das Merkmal ‚Formelhandhabung‘ erfasst kognitive und metakognitive Fähigkeiten, formale mathematische Ausdrücke zuverlässig zu handhaben [...]“ (ebd., S. 119).

Ihr Vorgehen ist, die PISA-Aufgaben nach den Merkmalen zu klassifizieren. Da jede Aufgabe eindeutig einer Kompetenzstufe zugeordnet ist, entstehen Häufigkeitsverteilungen der Merkmalsausprägungen je Kompetenzstufe. Die Auswertung ergibt getreu den Erwartungen, je höher der kognitive Anspruch ist, desto seltener befinden sich die Aufgaben auf den unteren Kompetenzstufen und desto häufiger auf den höheren Kompetenzstufen. Weiter berechnen Cohors-Fresenborg et al. Rangkorrelationen der Merkmale und Aufgabenschwierigkeit sowie Regressionsmodelle. Sie schließen unter anderem, „dass innerhalb des Merkmals ‚Kognitive Komplexität‘ das Merkmal ‚Formalisierung von Wissen‘ in besonderer Weise schwierigkeitsgenerierend wirkt“ (Cohors-Fresenborg et al. 2004, S. 139).

Bezug zu dieser Arbeit

Die Auszüge aus der Bildungsforschung zur Mathematik lassen erkennen, dass eine breite Grundmenge von Merkmalen und vielfältige statistische Ansätze, Kategorien daraus zu formen, existieren. Einige Ergebnisse bestätigen die meist erwartungsgetreue Schwierigkeitsrelevanz einzelner Merkmale. Nicht abzulesen ist eine etwaige Konvergenz der Kategoriensysteme.

3.3.2 Naturwissenschaften

Analog zur Mathematik gibt es in den Naturwissenschaften, besonders im Kontext von PISA, etliche Vorschläge zur Aufgabenkategorisierung sowie Analysen des Zusammenhangs zwischen Merkmalen und Aufgabenschwierigkeit. Auf internationaler Ebene wird in PISA 2006 mit Verfahren der IRT eine fünfstufige Skala

naturwissenschaftlicher Kompetenz konstruiert. Für eine qualitative Auslegung der Stufen werden die folgenden Merkmale verwendet (OECD 2006, S. 41):

Internationale Berichtsmerkmale

- General complexity of the context
- Level of familiarity with the scientific ideas, processes and terminology involved
- Length of the train of logic required to respond to a question - that is, the number of steps needed to arrive at an adequate response and the level of dependence of each step on the previous one

Auf diesen Merkmalen, Komplexität der Situation, naturwissenschaftliche Grundbildung und Komplexität der Schlussfolgerung, fußt die standardisierte Rückmeldung (Reporting) der internationalen Studie. Die nationale PISA-Ergänzung betreffend nennen Baumert et al. (2003) fünf zentrale kognitive Aspekte, denen die PISA-Items je nach dem dominierenden Prozess bei der erfolgreichen Aufgabenbearbeitung eindeutig zugeordnet werden (Baumert et al. 2003, S. 19):

Kognitive Prozesse

- Heranziehen von konzeptuellem und Faktenwissen bei der Aufgabenlösung
- Entnehmen relevanter Informationen aus einer Graphik oder einem Diagramm
- Nutzung eines mentalen Modells über einen naturwissenschaftlichen Sachverhalt
- Aus gegebener verbaler Information die richtigen Schlüsse ziehen
- Einen Sachverhalt angemessen verbal beschreiben

Die Verbindung von Aufgaben mit Prozessmerkmalen ermöglicht eine aussagekräftige Ergebnisdarstellung. Zum Beispiel interpretieren Baumert et al. die Ergebnisse des Ländervergleichs anhand der Merkmalsprofile dahingehend, „dass in den neuen Ländern vornehmlich die Testaufgaben, zu deren Lösung Konzept- und Faktenwissen erforderlich ist, zu den hohen Leistungen beitragen. Eine erste Schlussfolgerung könnte daher lauten, dass in der Unterrichtstradition der neuen Länder die Vermittlung von Konzept- und Faktenwissen eine größere Rolle spielt als in den meisten alten Ländern. Eine Ausnahme stellt Bayern dar, wie auch Schleswig-Holstein und Saarland im Gymnasialbereich“ (ebd., S. 19).

In vertiefenden Analysen naturwissenschaftlicher Kompetenz verwenden Prenzel et al. (2002) eine umfangreiche Aufstellung formaler Aufgabenmerkmale, kognitiver Anforderungen und wissensbasierter Merkmale (Prenzel et al. 2002, S. 130):

Formale Aufgabenmerkmale

- Langer Aufgabentext
- Aufgabe enthält Grafik
- Aufgabe enthält bildliche Informationen

- Freie kurze Antwort
- Freie lange Antwort

Kognitive Anforderungen

- Textinformationen verarbeiten
- Logisch verknüpfen
- Räumliches Modell aufbauen
- Etwas ausrechnen
- Divergent denken

Merkmale mit Wissensbasis

- Terminologisches Wissen
- Faktenwissen
- Kontraintuitives Wissen
- Funktionale Zusammenhänge

Mittels einer Regressionsanalyse wird der Zusammenhang der Aufgabenmerkmale mit der Aufgabenschwierigkeit berechnet, wobei für die Aufgabenschwierigkeit der Schätzwert des Itemparameters im Rasch-Modell eingesetzt wird. Unter den formalen Merkmalen scheint die Vorgabe einer Grafik oder Bildinformation die Lösung der Aufgabe zu erleichtern, die aktive Verbalisierung in Form einer freien Antwort dagegen erschwert sie. Unter den kognitiven Anforderungen bereiten etwas ausrechnen, ein räumliches Modell aufbauen und divergentes Denken Schwierigkeiten. Als Merkmal mit Wissensbasis ist die Erfordernis terminologischer Wissens relevant (vgl. S. 31).

Gut et al. (2010) führen in sechsten und neunten Jahrgangsstufen kombinierte Experimentier- und Papier- und Bleistift-Tests zu Mensch, Umwelt, Natur und Technik durch und untersuchen die Wirkung von äußeren Aufgabenmerkmalen wie Textlänge des Inputs oder Offenheit des Outputs, inneren Aufgabenmerkmalen wie Anzahl Lösungen, Inhalts-Merkmalen und Kodierungs-Merkmalen. Ordinale Rasch-Analysen ergeben auch in dieser Stichprobe Hinweise, dass Informationen grafischer oder bildlicher Art Experimentieraufgaben vereinfachen. Paarweise varianzanalytische Vergleiche können die vermuteten Zusammenhänge von Itemmerkmalen und Itemschwierigkeiten aber nicht bestätigen.

Kauertz (2008) analysiert die Merkmale *Komplexität*, *Physikalische Leitidee* und *Kognitive Aktivität* innerhalb eines „Inhaltsstrukturmodells“ der Schwierigkeit von Physikaufgaben für Schüler der zehnten Jahrgangsstufe an einem Gymnasium:

Komplexität

- Ein Fakt (1)
- Mehrere Fakten (2)
- Ein Zusammenhang (3)

- Mehrere unverbundene Zusammenhänge (4)
- Mehrere verbundene Zusammenhänge (5)
- Übergeordnetes Konzept (6)

Physikalische Leitidee

- Energie
- Materie
- Wechselwirkung
- System

Kognitive Aktivität

- Erinnern
- Strukturieren
- Explorieren (Kauertz 2008; Viering et al. 2010, S. 93)

Im Vergleich zu einem „Aufgabendesignmodell“ mit den formalen Merkmalen Antwortformat, Position der Aufgabe im Testheft, Abstand zwischen Lern- und Testzeitpunkt kann er zeigen, dass Komplexität und Leitidee signifikant zur Varianzaufklärung beitragen, während das Inhaltsstrukturmerkmal Kognitive Aktivität und die formalen Merkmale kaum ins Gewicht fallen. Über die hierarchischen Komplexitätsstufen lassen sich die Aufgabenschwierigkeiten vorhersagen, wenn auch mit unbefriedigender Genauigkeit (Kauertz 2008, S. 101f). Ein sechsdimensionales IRT-Modell, bei dem jeweils eine Leitidee eine Dimension bildet, passt besser zu den erhobenen Daten als ein eindimensionales Modell. Auffällig ist, dass ein dreidimensionales IRT-Modell, bei dem jeweils eine kognitive Aktivität eine Dimension bildet, keine signifikant bessere Datenpassung hat als das eindimensionale Modell (ebd., S. 104ff).

Viering et al. (2010) erweitern dieses Modell um die Dimension der Konzeptentwicklung und planen, in einer Längsschnittuntersuchung zu überprüfen, ob die Schülerfähigkeit bezüglich der *Komplexität* und der *Konzeptentwicklung* mit der Jahrgangsstufe wächst. Die Pilotierung der Aufgaben zeigt, dass die Schwierigkeit mit der Konzeptentwicklungsstufe wächst, während ein Einfluss der Komplexität noch nicht zufriedenstellend nachgewiesen werden kann.

Bezug zu dieser Arbeit

Ebenso wie in der Mathematik sind auch in den Naturwissenschaften in den letzten zehn Jahren viele Arbeiten entstanden, die sich mit der Wirkung schwierigkeiterzeugender Merkmale beschäftigen. Die unterschiedlichen statistischen Ansätze liefern überwiegend Ergebnisse, die in der erwarteten Richtung liegen, aber ungenau und daher unbefriedigend sind.

3.3.3 Andere Fächer und Fächerübergreif

Forschungsbestrebungen, in deren Mittelpunkt potentiell schwierigkeiterzeugende Merkmale stehen, gibt es selbstverständlich auch außerhalb der mathematisch-naturwissenschaftlichen Kernfächer. Relevant für das Forschungsvorhaben in der Informatik sind Projekte in der Berufsbildung sowie im Kontext fächerübergreifenden Problemlösens und im Kontext fächerübergreifender Aufgabenbewertung. Die Zahl der fächerübergreifenden Veröffentlichungen bleibt hinter der Zahl der Publikationen aus den Fachdidaktiken zurück.

Berufsbildung

Schumann und Eberle (2011) beschäftigen sich mit Schwierigkeitsindizes für die Erfassung ökonomischer und beruflicher Kompetenzen von Maturandinnen und Maturanden in der Deutschschweiz. Sie schlagen vier Merkmale vor (Schumann und Eberle 2011, S. 83f):

- Modellierung (als geringe, einfache oder fortgeschrittene Leistung je nach der Komplexität des Entscheidungsmodells)
- Kognitive Verarbeitungsprozesse (als Reduzierung der Bloomschen Taxonomie auf die Stufen Informationen wiedergeben/erinnern, Informationen verstehen und verarbeiten, Informationen nutzen)
- Offenheit der Lösungswege und Lösungen (eindeutige Lösung oder mehrere)
- Definitionen oder Gesetzmäßigkeiten (keine, wenige oder mehrere Definitionen oder Gesetzmäßigkeiten aus dem Gedächtnis abrufen)

Ähnlich zum Vorgehen von Cohors-Fresenborg et al. (2004), das im Abschnitt über Schwierigkeitsmerkmale in der Mathematik (S. 46) beschrieben wurde, klassifizieren sie Leistungstestaufgaben, die allerdings gezielt nach den Merkmalen konstruiert wurden, und berechnen in einer Pilotierungsstudie die Lösungsquoten. Sie vergleichen die Mittelwerte der Lösungsquoten je Merkmalsausprägung und stellen ebenfalls für alle Merkmale ein Gefälle fest: „Items mit a priori höher eingeschätzter Schwierigkeit werden weniger häufig gelöst, *et vice versa*“ (Schumann und Eberle 2011, S. 86, Hervorhebung im Original).

Winther und Achtenhagen (2009) analysieren kaufmännische Kompetenz nach drei Kriterien, *Funktionale Modellierung*, damit ist die eigenständige Übersetzung der Situation in die notwendigen Schritte des Lösungsprozess gemeint, *Inhaltliche Komplexität* und *Art der kognitiven Taxonomierung*, vom Informationsabruf bis zur sinnvollen Verwendung kognitiver Ressourcen. In einer Regressionsanalyse mit Dummy-Variablen, das heißt, für jede Ausprägung der Kriteriumsvariablen wird eine binäre Scheinvariable eingeführt, identifizieren sie vier bedeutsame Merkmale für die Vorhersage der Aufgabenschwierigkeit (Winther und Achtenhagen 2009, S. 545):

- Umfangreiche Modellierungsleistung, eine Ausprägung des Kriteriums Funktionale Modellierung
- Korrekte Modellierungsleistung, ebenfalls eine Ausprägung des Kriteriums Funktionale Modellierung
- Systemkohärenz (gemeint ist der Zusammenhang zwischen verschiedenen betriebswirtschaftlichen Konzepten), eine Ausprägung des Kriteriums Inhaltliche Komplexität
- Analysieren und Validieren als Stufe der Kognitiven Taxonomierung

Problemlösen

Mit schwierigkeitsgenerierenden Merkmalen des Problemlösens sind Greiff und Funke (2010) im Projekt „Dynamisches Problemlösen“ des DFG-Schwerpunktprogramms Kompetenzmodellierung befasst. Sie gehen von fünf Anforderungen an eine problemlösende Person aus, ob bei der Bedienung eines Fahrkartenautomaten oder beim richtigen Verhalten in einem medizinischen Notfall: „(a) die Reduktion überbordender Information auf einen handhabbaren Umfang (Informationsreduktion); (b) die Bildung adäquater Situationsmodelle zum Verständnis der gegebenen Situation (Modellbildung); (c) die Prognose weiterer Entwicklungen aufgrund der gegebenen Situation und im Lichte getroffener Maßnahmen (Prognose); (d) die Beschaffung fehlender, aber für die Problemlösung notwendige [sic] Information (Informationssuche und -generierung) und (e) das Treffen von Wertentscheidungen und Prioritätensetzungen, mit denen Ziele gesetzt und Zielkonflikte gelöst werden können (Bewertung)“ (Greiff und Funke 2010, S. 217). Mittels Kovarianzanalysen wird der Einfluss der Merkmale Effektqualität, Effektzahl und Variablenzahl auf die Systemschwierigkeit untersucht. Die Autoren finden ihre Hypothese bestätigt, dass die Schwierigkeit mit der Effektqualität variiert und mit der Effektzahl oder Variablenzahl ansteigt.

Fächerübergreifende Aufgabenbewertung

Kleinknecht et al. (2011) beabsichtigen, ein allgemeindidaktisches Kategoriensystem zu schaffen, um zu Zwecken der Lehrerbildung und der Unterrichtsplanung das kognitive Potential von Aufgaben zu beurteilen. Das Raster besteht aus dem losen Nebeneinander von sieben Facetten (Kleinknecht et al. 2011, S. 333):

- Wissensarten (Fakten, Prozeduren, Konzepte, Metakognition)
- Kognitiver Prozess (Reproduktion, naher -, weiter Transfer, Problemlösen)
- Anzahl der Wissenseinheiten (eine, bis zu 4, mehr als 4 Wissenseinheiten)
- Offenheit (definiert/konvergent, definiert/divergent, ungenau/divergent)
- Lebensweltbezug (kein Bezug, konstruiert, authentisch, real)
- Sprachlogische Komplexität (niedrig, mittel, hoch)
- Repräsentationsformen (eine Form, Integration, Transformation)

Das Kategoriensystem wird an schulpraktischen Aufgaben der Fächer Deutsch, Mathematik, Natur- und Gesellschaftswissenschaften erprobt. Als Gütekriterium wird die gute bis sehr gute Übereinstimmung des Ratings herangezogen.

Bezug zu dieser Arbeit

Der Blick auf die Forschungstätigkeit bezüglich Berufsbildung, Problemlösen und fächerübergreifender Kategorien offenbart beim Thema Schwierigkeitsmerkmale wie schon beim Thema Kompetenzmodelle Ansätze und Ergebnisse, die parallel zu den mathematisch-naturwissenschaftlichen Fächern, speziell Informatik, verlaufen. Als Kriterien der Schwierigkeit wiederholen sich vier Merkmale, der Modellierungsanspruch des Problems beziehungsweise der Situation, die kognitiven Prozessanforderungen bei der Problembearbeitung, die Komplexität des Inhalts oder erforderlichen Wissens und die Offenheit der Aufgabe.

3.4 Zusammenschau

Die dargestellten Forschungsbeiträge zur Kompetenzmodellierung bauen fast ausnahmslos auf der von Weinert (2001) geprägten Kompetenzdefinition auf, die durch Klieme und Leutner (2006) weiter eingeschränkt wurde. Ausgesprochen oder unausgesprochen ist dieser Kompetenzbegriff vielfach wiedererkennbar. Auch die vorliegende Untersuchung folgt der Begriffsbestimmung durch Klieme und Leutner und konzentriert sich auf die kognitiven Aspekte kontextspezifischer Leistungsdispositionen.

Grundsätzlich werden Kompetenzstrukturmodelle von Kompetenzstufenmodellen unterschieden. Im ersten Fall wird mit klassischen faktorenanalytischen oder probabilistischen Methoden ergründet, in welche Dimensionen sich ein Kompetenzkonstrukt auffächert. Im zweiten Fall werden Abschnitte einer Leistungsskala als hierarchische Kompetenzniveaus ausgewiesen. Die Skala wird in der Regel durch die Anpassung eines eindimensionalen Rasch-Modells gebildet und erfasst Aufgabenschwierigkeiten und Personenfähigkeiten gemeinsam.

Insgesamt überwiegt die eindimensionale Modellierung mit Methoden der Item-Response-Theorie. Selbst wenn eine mehrdimensionale Struktur mit mehreren latenten Variablen vermutet wird, erfolgt oft eine pragmatische Festlegung von Dimensionen und anschließend die separate Modellierung der Dimensionen mit jeweils einer latenten Variablen. Gegebenenfalls werden die Dimensionen iterativ verfeinert beziehungsweise korrigiert, wie beispielsweise in der nationalen Ergänzungsstudie zu PISA-Naturwissenschaften. Im Gegensatz dazu wird in der vorliegenden Arbeit ein mehrdimensionales Kompetenzstrukturmodell entworfen, dessen Dimensionen nicht pragmatisch festgelegt, sondern unter Zuhilfenahme explorativer faktorenanalytischer Methoden aus empirischen Daten herausgearbeitet werden.

In den Abschnitten 3.1 und 3.2 wurde eine repräsentative Auswahl bestehender Kompetenzmodelle der Domänen Informatik, Mathematik, Naturwissenschaften, Berufsbildung und Problemlösen konturiert. Auf einen Nenner gebracht weisen die Modelle folgendes Grundmuster fachdidaktischer Dimensionen auf:

- Eine *Prozessdimension*, die etwa Denkprozesse bei der Problembearbeitung berücksichtigt, zum Beispiel Kognitive Prozesse bei Anderson und Krathwohl et al. (S. 22), Verarbeitungsstufen bei Marzano und Kendall (S. 23), Mathematical Processes in PISA-Mathematik (S. 28), Mathematisierungs-, Übersetzungs- oder Problemlöseprozess bei Neubrand et al. (S. 29), Competencies in PISA-Naturwissenschaften (S. 31), Kognitive Anforderungen bei Prenzel et al. (S. 47), die Facetten ‚Bewerten‘, ‚Divergentes Denken‘, ‚Konvergentes Denken‘, ‚Sachverhalte verbalisieren‘ bei Senkbeil et al. (S. 32), die Prozessbereiche in den GI-Standards (S. 40, 200) oder die Prozessdimension bei Freischlad (S. 41). Eine solche Prozessdimension ist allen Modellen gemeinsam.
- Eine *Wissensdimension*, curricular oder auch taxonomisch, zum Beispiel Wissensarten bei Anderson und Krathwohl et al. (S. 22), Wissensgebiete bei Marzano und Kendall (S. 23), in der Aufgabe angesprochenes Wissen bei Neubrand et al. (S. 29), Knowledge in PISA-Naturwissenschaften (S. 31) oder Merkmale mit Wissensbasis bei Prenzel et al. (S. 47). Nicht in allen Modellen wird eine Wissensdimension extra ausgewiesen.
- Eine *Repräsentationsdimension*, die formale Kriterien wie das Antwortformat einbezieht, etwa Formale Aufgabenmerkmale bei Prenzel et al. (S. 47), die Facetten ‚Umgang mit Graphen‘, ‚Umgang mit mentalen Modellen‘, ‚Umgang mit Zahlen‘ bei Senkbeil et al. (S. 32) oder die Darstellungsdimension bei Freischlad (S. 41). Eine Repräsentationsdimension wird nicht in allen Modellen ausgewiesen.
- Eine Inhalts- oder *Kontextdimension*, meistens die Anwendungssituation. Beispiele sind der Mathematical Content in PISA-Mathematik (S. 28), der außer- oder innermathematische Kontext bei Neubrand et al. (S. 29), die Inhaltsbereiche bei Senkbeil et al. (S. 32), Fachliche Inhalte in den EPA-Informatik (S. 39), die Inhaltsbereiche in den GI-Standards (S. 40, 199) und die Inhaltsdimension bei Freischlad (S. 41). Die Anwendungssituation wird oft nicht als eigene Dimension, sondern als Kontext aufgeführt, auf den das Modell bezogen wird.

Dieses Modellmuster mit je einer Prozess-, Wissens- und Repräsentationsdimension wird in Abbildung 3.1 verbildlicht. Die Anwendungssituation wird nicht als vierte Dimension des Modells aufgenommen, sondern als Kontext repräsentiert. Ein solches Dimensionenmuster erwächst aus fachdidaktischen Überlegungen und bildet das Gerüst für eine detaillierte Beschreibung und Erklärung

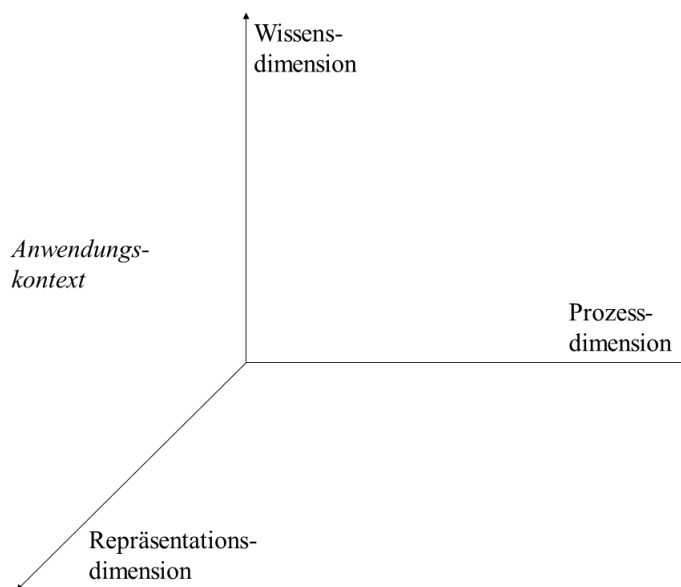


Abbildung 3.1: Fachdidaktisches Strukturmodell mit den Dimensionen Prozess, Wissen und Repräsentation

von Leistungserwartungen und Leistungsdispositionen. In den seltensten Fällen sind derartige fachdidaktische Dimensionen orthogonal, das heißt unabhängig. Sie hängen vielfach voneinander ab, das heißt, eine Veränderung innerhalb einer Dimension, zum Beispiel ein anderes Antwortformat in der Repräsentationsdimension, zieht möglicherweise unbeabsichtigte Veränderungen in anderen Dimensionen nach sich, wie die Adressierung eines anderen Denkprozesses in der Prozessdimension.

In dieser Untersuchung wird eine Dimensionenstruktur ermittelt, die nicht fachdidaktischen Kategorien entspringt, sondern aus empirischen Beobachtungen abgeleitet wird und idealerweise unabhängige Dimensionen aufweist.

Im Abschnitt 3.3 wurden Arbeiten über Schwierigkeitsmerkmale herausgegriffen, die sich auf der Ebene der Aufgaben mit deren Anforderungskriterien auseinandersetzen. Diese sind in der Kompetenzmodellierung in zweierlei Hinsicht wichtig, denn sowohl die Charakterisierung der Dimensionen in Strukturmodellen als auch die Interpretation der Kompetenzniveaus in Stufenmodellen beruhen auf Anforderungsmerkmalen. Traditionell wird untersucht, welche Aufgabeneigenschaften einen messbaren Einfluss auf die Schwierigkeit ausüben. Erschwerend wirkt hierbei das Problem der Merkmalskonfundierung: treten Merkmale gehäuft gemeinsam auf, lassen sich ihre Einzeleffekte nicht mehr isolieren.

Dieses Problem erfordert hier besondere Aufmerksamkeit. Die Datenbasis der Untersuchung besteht aus den Aufgaben und Teilnehmerantworten des Online-Schülerwettbewerbs Informatik-Biber (siehe Kapitel 6). Die Wettbewerbsaufga-

ben werden nach ihren Schwierigkeitsmerkmalen klassifiziert, wobei sich ergibt, dass die charakterisierenden Merkmale in den Aufgaben hochgradig konfundiert sind. Diese unerwünschte Korrelation überlagert in den Analysen die Wechselbeziehung der kognitiven Fähigkeiten und Fertigkeiten bei der Aufgabenbearbeitung. Trägt eine Aufgabe zwei Merkmale und löst ein Schüler sie nicht, bleibt unklar, ob er wirklich keine der beiden Fähigkeiten besitzt, die mit den Merkmalen verbunden sind. Anders betrachtet deutet das häufige gemeinsame Auftreten zweier Merkmale daraufhin, dass die damit verbundenen Fähigkeiten häufig gleichzeitig benötigt werden. Aus diesem Blickwinkel liegen sie in einer gemeinsamen Dimension. Aus dieser Überlegung heraus wird hier zuerst, nur basierend auf den Aufgaben, ein Anforderungsmodell abgeleitet, bevor, nun basierend auf den Schülerantworten, die Fähigkeitsmodelle entwickelt werden.

Bemerkenswert ist die Uneinheitlichkeit der Befunde zur Schwierigkeitsrelevanz der Aufgabenmerkmale. Zum Beispiel schließen Senkbeil et al. (2005) aus dem empirischen Vergleich des 7D-Modells *kognitiver Kompetenzen* mit dem 3D-Fächermodell und dem 1D-Rasch-Modell, dass kognitive Anforderungen für die Struktur naturwissenschaftlicher Kompetenz bedeutsam sind (vgl. S. 32). Kauertz (2008) hingegen zeigt für das Fach Physik, dass die *kognitive Aktivität* für die Bestimmung der Aufgabenschwierigkeit vergleichsweise geringes Gewicht hat (vgl. S. 48). Das scheint widersprüchlich. Der Vergleich der Spezifikation fördert zutage, dass die Facetten kognitiver Kompetenz bei Senkbeil et al. in konkreten, heterogenen Anforderungssituationen bestehen, nämlich mit Zahlen umgehen, Sachverhalte verbalisieren, mit mentalen Modellen umgehen, konvergent denken, mit Graphen umgehen, divergent denken, bewerten. Die kognitiven Aktivitäten bei Kauertz benennen hingegen abstrakte, disjunkte Kategorien von Denkprozessen, nämlich Erinnern, Strukturieren oder Explorieren. Auf der Modellebene lassen sich die beiden Auffassungen kaum zur Deckung bringen. So spiegeln die abweichenden Befunde die bunte Modelllandschaft.

Insgesamt trifft die Feststellung von Klieme und Leutner (2006) für den Entwicklungsstand theoretisch begründeter und empirisch geprüfter Kompetenzmodelle noch zu, dass ihr Abstraktionsgrad zu hoch ist. „Insofern bleibt den Testentwicklern in aller Regel keine andere Wahl, als sehr große Mengen von Aufgaben zu entwickeln, diese Aufgaben empirisch zu erproben und dann im Nachhinein diejenigen Aufgaben beizubehalten, die dem jeweils verwendeten Testmodell entsprechen“ (Klieme und Leutner 2006, S. 7).

In einem kritischen Artikel zur empirischen Bildungsforschung gehen Zlatkin-Troitschanskaia und Seidel (2011) der Frage nach, ob diese an einem „Theorie-Empirie-Problem“ krankt, wie es Beck (1987) bereits in den 80er-Jahren ange mahnt habe. Das Problem bestehe in der fehlenden theoretischen Fundierung vieler empirischer Ansätze, die mit unangemessenen Methoden einhergehe, was wiederum zu widersprüchlichen und inkonsistenten Ergebnissen führe. Die Autorinnen vergleichen die vorherrschenden Kompetenzbegriffe, die sich im Wesentlichen darin unterscheiden, ob sie kognitionsbezogen oder handlungsbezogen

sind. Kognitionsbezogen heißt, Kompetenz wird eher eng als individuelle Disposition zur Bewältigung spezifischer Anforderungen ausgelegt, für das DFG-Schwerpunktprogramm zum Beispiel als kontextspezifische kognitive Leistungsdisposition. Durch affektiv-motivationale Faktoren wird sie höchstens beeinflusst. Handlungsbezogen heißt, Kompetenz konstituiert sich durch Selbst-, Sach- und Methodenkompetenz. In diesem Fall sind affektive und motivationale Bedingungen mehr als Einflussfaktoren, nämlich integrale Bestandteile des Konstrukts.

Zlatkin-Troitschanskaia und Seidel bemängeln, dass Kompetenz als Untersuchungsgegenstand häufig unzureichend definiert wird. Sie monieren beispielsweise, dass in PISA neben den schwerpunktmäßig untersuchten fachbezogenen auch bereichsübergreifende Kompetenzen wie Problemlösen und motivationale Aspekte wie die Einstellung zum Lernen einbezogen werden, ohne Einordnung in das Kompetenzkonzept. Weiter kritisieren sie, dass in Studien, die von mehreren Merkmalen als Einflussfaktoren ausgehen, diese ungenügend voneinander abgegrenzt werden. „Selten thematisiert wird das Problem, dass aufgrund des (erwartbaren) sich gegenseitigen Bedingens die (Interaktions-)Effekte nur sehr schwer erfasst und kontrolliert werden können; das Problem der Variablenkonfundierung wird kaum berücksichtigt“ (Zlatkin-Troitschanskaia und Seidel 2011, S. 226). Ihr Resümee ist, dass die Struktur von Kompetenz, das Verhältnis von kognitiven, motivationalen und volitionalen Komponenten zentral ist. Der Zusammenhang zwischen diesem Theoriemodell, dem Analyse- und dem Messmodell müsse in Zukunft stärker und konsequenter in den Blick genommen werden (ebd., S. 229).

Für dieses Vorhaben wird der maximal enge kognitionsbezogene Kompetenzbegriff gewählt, Kompetenz als individuelle Disposition zur Bewältigung spezifischer Anforderungen. Das Problem der Variablenkonfundierung wird berücksichtigt, wenn es auch nicht gelöst werden kann, so dass ein wohldurchdachter Beitrag zur Exploration der dimensional Struktur von Kompetenz geleistet wird.

Gerade in der Informatik wurde die Kompetenzstruktur, als Grundlage für die Modellierung von Kompetenzniveaus oder ihrer Entwicklung, erst wenig erforscht. Am Curriculum orientierte Modelle liegen vor, die einen Rahmen bieten, um Lerninhalte zu strukturieren, Unterricht zu konzipieren und Aufgaben zu entwickeln. Die Modelle sind meist fachsystematisch, selten lerntheoretisch begründet; kognitive Modelle informatischer Schwierigkeit und Fähigkeit werden bislang nicht beschrieben. Eine empirische Untersuchung ihrer Eignung, Schülerdaten zu erklären, steht noch aus.

Hier wird ein empirisch-statistischer Ansatz verfolgt, die Anforderungsstruktur von Aufgaben mit Informatikbezug und die Fähigkeitsstruktur von Schülern, die die Aufgaben bearbeiten, zu modellieren. Der Schlüssel dazu sind die Aufgabenmerkmale, die die kognitiven Anforderungen charakterisieren. Mittels statistischer Analysen wird aus den Merkmalen anhand der empirischen Daten einer Expertenbefragung und eines Schülerwettbewerbs bottom-up eine Modellvorstellung der Anforderungs- und Fähigkeitsstruktur aufgebaut.

Kapitel 4

Zielstellung und Methodik

Die Recherche zum Forschungsstand lässt zwei Auswertungsebenen erkennen. Auf der *Merkmalsebene* werden Aufgaben analysiert, relevante Merkmale identifiziert und gegebenenfalls wird empirisch untersucht, ob die Merkmale Schwierigkeit erzeugen. Auf der *Modellebene* werden aus fachsystematischen oder kognitionspsychologischen Gliederungen Kompetenzdimensionen abgeleitet. Nur in wenigen Fällen werden sie empirisch auf ihre Erklärungsmächtigkeit untersucht. Die beschriebenen Ansätze sind *konfirmatorisch*, das heißt, mit statistischen Methoden werden vorliegende Zusammenhangs- oder Strukturhypothesen überprüft. Alternativ zum konfirmatorischen Ansatz bedeutet *exploratorisch*, mit statistischen Methoden erstmalig Verdachtsmomente und Hypothesen aufzudecken. Über exploratorische Ansätze, mit strukturentdeckenden Verfahren insbesondere die Dimensionen informatischer Kompetenz zu erkunden, wird nicht berichtet. Hier besteht eine Forschungslücke.

In der Informatik gibt es einzelne normative Modelle, die sich meist auf die Kompetenzdefinition von Weinert (2001) stützen und bestimmte Fachinhalte behandeln. Als Beispiel sei das Projekt MoKoM genannt, das sich auf das Thema Modellierung und Systemverständnis in der Informatik bezieht (Nelles et al. 2010; vgl. auch S. 42). Grundlegende Fragen sind für die Informatik noch nicht geklärt, etwa welche Merkmale geeignet sind, den kognitiven Anspruch typischer Informatikprobleme zu beschreiben, durch welche kognitiven Charakteristika Informatik sich von anderen Fächern abhebt und welche dimensionale Struktur informatische Kompetenz aufweist.

In dieser Untersuchung wird exemplarisch der Ausschnitt informatischer Anforderungen und Fähigkeiten betrachtet, der durch die Aufgaben des Schülerwettbewerbs Informatik-Biber (siehe S. 82) operationalisiert wird. Das Ziel ist, entlang der Aufgaben und Schülerdaten eine Modellvorstellung der Kompetenzstruktur zu entwickeln, innerhalb derer die kognitiven Aufgabenanforderungen wie auch die Leistungsfähigkeit von Personen abgebildet werden können. Die Zielformulierung wird im Folgenden präzisiert.

4.1 Forschungsfragen

Für den Begriff der Informatikkompetenz wird die Definition von Weinert hier auf „die bei Individuen verfügbaren oder durch sie erlernbaren kognitiven Fähigkeiten und Fertigkeiten, um bestimmte Probleme zu lösen“ (Weinert 2001, S. 27f) begrenzt, das heißt, motivationale, volitionale und soziale Bestandteile werden soweit möglich ausgenommen. Nach Klieme und Leutner wird der Begriff explizit als „kontextspezifische kognitive Leistungsdispositionen, die sich funktional auf Situationen und Anforderungen in bestimmten Domänen beziehen“ (Klieme und Leutner 2006, S. 879) ausgelegt. Informatikkompetenz wird demnach als kognitive Leistungsdisposition in spezifisch informatischen Kontexten verstanden, in Abgrenzung zu generalisierter, kontextunabhängiger Intelligenz. Das entspricht nach Schaper dem klassischen Verständnis von Kompetenz als Wissens- und Fähigkeitsvoraussetzung (Schaper 2008, S. 92; vgl. auch S. 20). Dabei wird nicht nach einzelnen Inhalten differenziert. Ergründet wird der Kompetenzanteil, der wesentlich für die Informatik und unabhängig vom speziellen Thema ist.

Wie Schecker und Parchmann ausführen, sind die bestehenden normativen Kompetenzstrukturmodelle in der Regel fachsystematisch geprägt. „Es bleibt offen, ob eine solche Struktur geeignet ist, um das tatsächlich vorfindbare Fähigkeitsmuster von Schülern zu erfassen“ (Schecker und Parchmann 2006, S. 50). Normative Modelle dienen der Erschließung und Handhabung der inhaltlichen Fülle und Vielfalt didaktischer Zugänge, zum Beispiel in Curricula. Geht es aber um Anforderungen und Fähigkeiten, aus Gründen der Unterrichtsgestaltung oder Leistungsbeurteilung, werden Modelle benötigt, die den Ist-Zustand beschreiben. Das hier angestrebte Strukturmodell wird deskriptiv angelegt – im Gegensatz zum normativen Ansatz wird keine Aussage über die erwünschte Leistungsdisposition bei Schülern getroffen, vielmehr werden die tatsächlich vorgefundenen Gegebenheiten beschrieben.

Zur Beschreibung der gegebenen Anforderungen und Fähigkeiten sind verschiedene Modellvarianten geplant. Auf der Merkmalebene werden Informatikaufgaben auf die kognitiven Kriterien hin untersucht, die ihre Lösungswahrscheinlichkeit beeinflussen. Um herauszuarbeiten, welche Aufgaben in gemeinsamen Dimensionen liegen, werden sie nach den Kriterien gruppiert. Die resultierenden Aufgabencluster werden anhand der vorherrschenden Merkmale als Anforderungsdimensionen interpretiert. So entsteht das *Anforderungsmodell* mit mehreren Dimensionen, charakterisiert durch kognitive Merkmale.

Auf der Modellebene werden die Ergebnisse der Aufgabenbearbeitung ausgewertet. Als erste Variante werden die Antwortmuster der Schüler (Welcher Schüler löste welche Aufgabe?) analysiert. Die Aufgaben, die dieselben Schüler lösten, werden in Dimensionen zusammengefasst. Anhand der Merkmalsprofile werden die Dimensionen interpretiert. Als zweite Variante werden die Merkmalsquoten der Schüler (Wie viele Aufgaben eines Merkmals löste ein Schüler?) berechnet. Die Merkmale, deren Quote bei denselben Schülern ähnlich ausgeprägt

ist, werden wiederum in Dimensionen zusammengefasst. So entstehen zwei alternative *Fähigkeitsmodelle*. Es handelt sich bei beiden Varianten um Dimensionen-Modelle, im Unterschied zu Stufen- oder Muster-Modellen (Rost 2004b; vgl. auch S. 26).

Konkret kommen informatische Anforderungen hier als kognitiver Anspruch der Aufgaben des Schülerwettbewerbs Informatik-Biber zum Ausdruck. Daraus ergibt sich die Einschränkung auf das Format Multiple-Choice und auf überschaubare Probleme, deren Lösung in kurzer Zeit zu schaffen ist. Nötiges Vorwissen wird kompakt in der Angabe dargeboten, denn bedingt durch die heterogene Informatikausbildung in den Ländern können weder Kalkülwissen noch Programmierfertigkeiten als bekannt und eingeübt vorausgesetzt werden. Weiter kommen nur Probleme in Betracht, die Sekundarstufenschüler nach ihren entwicklungspsychologisch zu erwartenden Voraussetzungen lösen können, da die Adressaten des Informatik-Bibers Schülerinnen und Schüler der fünften bis dreizehnten Jahrgangsstufe, also im Alter von zehn bis ungefähr neunzehn Jahren sind. In der Summe heißt das, nur derjenige Anteil von Informatikkompetenz, der durch das Fenster des Informatikwettbewerbs zugänglich ist, wird untersucht. Dabei kann zum Beispiel nicht differenziert werden, ob Erfolge auf Personeneigenschaften wie Intelligenz oder erworbene Kenntnisse wie Strategien zurückgehen.

Aus der pragmatischen Entscheidung, mit den Wettbewerbsdaten zu arbeiten, ergeben sich auch methodische Einschränkungen. Deren gravierendste ist die fehlende Systematik in der Variation der Merkmale (siehe S. 104, zur Konfundierung der Effekte).

Die Exploration der dimensional Struktur informatischer Anforderungen und Fähigkeiten, die sich in den Wettbewerbsaufgaben und Schülerantworten abzeichnet, wird von der Frage nach der Anzahl und dem Merkmalsgepräge der Dimensionen geleitet. Angestrebt werden Strukturhypothesen der Art, informatische Anforderungen und Fähigkeiten bilden sich entlang mehrerer Dimensionen aus, die durch kognitive Merkmale wie das erforderliche Abstraktionsvermögen, den Prozessbereich, die Lernzielstufe oder die Wissensart gekennzeichnet sind. Zur Kontrolle wird die Modellgeltung der Strukturhypothesen einer vergleichenden Prüfung unterzogen, bei der die empirische Anpassungsgüte beurteilt wird: Wie gut passen die Modelle zu den echten Daten? Aus den vorangehenden Überlegungen ergeben sich die Forschungsfragen.

1. Inwieweit lässt sich aus kognitiven Anforderungsmerkmalen von Aufgaben mittels klassifizierender statistischer Methoden (Clusteranalysen, Faktorenanalysen) ein Strukturmodell ableiten, welches die Dimensionen informatikspezifischer *Aufgabeanforderungen* abbildet?
2. Inwieweit lässt sich darüber hinaus ein Strukturmodell ableiten, das die Dimensionen informatikspezifischer *Schülerfähigkeiten* abbildet, wenn zusätzlich die Bearbeitungsergebnisse der Wettbewerbsteilnehmer einbezogen werden?

Die elementare Idee ist, dass kognitive Merkmale den Anspruch von Aufgaben charakterisieren und die Lösung der Aufgaben entsprechende Fähigkeiten erfordert, so dass Schüler, die die Aufgaben erfolgreich bearbeiten, dem Aufgabenanspruch gerecht werden und erwiesenermaßen über die notwendigen Fähigkeiten verfügen. Zum Zweck der vereinfachten Darstellung des Kompetenzgefüges, als Fundament systematischer Aufgabenkonstruktion sowie als Ausgangspunkt der künftigen Entwicklung von Messverfahren, wird ein deskriptives Dimensionen-Modell kognitiver Leistungsdispositionen erstellt. Der konstruktive Prozess, der auf der Basis empirischer Daten bottom-up von Schwierigkeitsmerkmalen zu Strukturhypothesen führt, ist übertragbar, beispielsweise auf andere Personengruppen oder durch Anpassung der Schwierigkeitskriterien auf engere Fachinhalte. Die geplante Vorgehensweise wird im Weiteren detailliert.

4.2 Forschungsmethodik

Der Argumentation folgend, dass ein Schüler, der beispielsweise abstrakte Aufgaben löst, Abstraktionsvermögen besitzt, wird ein Bogen von den Aufgaben und ihren Schwierigkeitsmerkmalen über die Fähigkeitsmerkmale der Schüler zu ihrer Problemlösekompetenz geschlagen. Im Grunde beruht die geplante Vorgehensweise auf der statistischen Indexbildung, wie sie Bortz und Döring (2006) beschreiben. Demnach ist ein Index „ein Messwert für ein komplexes Merkmal, der aus den Messwerten mehrerer Indikatorvariablen zusammengesetzt wird“ (Bortz und Döring 2006, S. 143). Im Falle einer normativen Indexbildung werden die Gewichte durch ein Expertenrating bestimmt, sonst empirisch-analytisch, wofür die Faktorenanalyse das geeignete Analyseverfahren ist.

„Resultieren in der Faktorenanalyse mehrere substantielle Faktoren, ist dies ein Beleg dafür, dass die Indikatoren kein eindimensionales Merkmal, sondern mehrere Dimensionen erfassen, was für die Theoriebildung über das interessierende Merkmal sehr aufschlussreich sein kann. Gegebenenfalls wird man dann das komplexe Merkmal nicht nur mit einem, sondern mit mehreren gewichteten Indizes erfassen. Die Gewichte für diese Indizes entsprechen den Ladungen der Indikatorvariablen auf den jeweiligen Faktoren“ (ebd.).

Die „komplexen Merkmale“ sind auf der einen Seite der Aufgabenanspruch und auf der anderen Seite die korrespondierende Schülerfähigkeit, die sich tatsächlich als mehrdimensional erweist. Die Indikatorvariablen sind die Schwierigkeitskriterien, die auch als Schwierigkeitsmerkmale, Aufgabenmerkmale oder einfach als Merkmale bezeichnet werden. Der Forschungsprozess umfasst sechs Phasen.

1. Theoretische Begründung der Schwierigkeitsmerkmale
2. Erhebung der Aufgabeneinschätzung und der Wettbewerbsdaten

3. Aufbereitung der Aufgabenklassifizierung und der Teilnahmedaten
4. Analyse der Aufgaben, Antwortmuster und Merkmalsquoten
5. Interpretation der Modelldimensionen
6. Prüfung der Modellgeltung

Schritt für Schritt werden so aus einfachen, beobachtbaren Merkmalen von Informatikaufgaben, zum Beispiel der Art des benötigten Wissens, die komplexen Dimensionen der Anforderungen und Fähigkeiten modelliert. Abbildung 4.1 gibt einen Überblick über die ersten fünf Schritte, die exploratorisch von Aufgabenmerkmalen zu mehrdimensionalen Modellen führen. Der sechste Schritt besteht in der Überprüfung der Modellgeltung für die Wettbewerbsdaten.

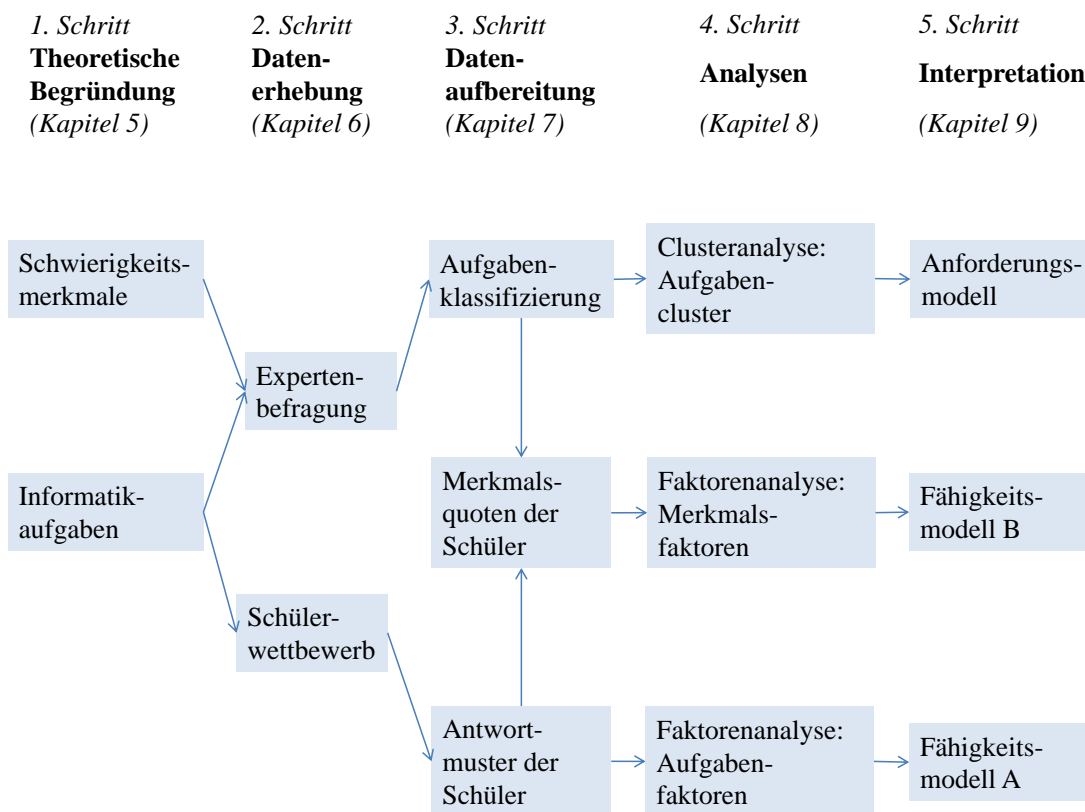


Abbildung 4.1: Methodik

Die Schwierigkeitsmerkmale und die Aufgaben gehen in die Expertenbefragung ein, deren Ergebnis die Aufgabenklassifizierung ist. Die Aufgaben werden im Informatikwettbewerb von Schülern bearbeitet und es entstehen die Antwortmuster. Aus der Klassifizierung (eine bestimmte Aufgabe trägt ein bestimmtes Merkmal) und den Antwortmustern (ein bestimmter Schüler löst die Aufgabe,

die das Merkmal trägt) werden die Merkmalsquoten der Schüler berechnet. Die Klassifizierung geht in eine Clusteranalyse ein, die Aufgabencluster werden als Dimensionen eines Anforderungsmodells interpretiert. Die Antwortmuster gehen in eine Faktorenanalyse ein, die Aufgabenfaktoren werden als Dimensionen eines Fähigkeitsmodells interpretiert, Modell A genannt. Die Merkmalsquoten gehen ebenfalls in eine Faktorenanalyse ein, die Merkmalsfaktoren werden wieder als Dimensionen eines Fähigkeitsmodells interpretiert, des Modells B. Der letzte Schritt, die Überprüfung der Modellgeltung, wird in der Grafik nicht dargestellt.

4.2.1 Theoretische Begründung: Aufgabenmerkmale

Ganz am Anfang dieser Arbeit stehen die Aufgabenmerkmale. Klassische Taxonomien, Bildungsstandards, Bildungsstudien sowie neuere, grundlegende Arbeiten auf dem Gebiet der Kompetenzmodellierung werden gesichtet, um Schwierigkeitsmerkmale zu identifizieren. Auch Veröffentlichungen benachbarter Fachdidaktiken, insbesondere aus der Mathematik und den Naturwissenschaften werden berücksichtigt (Kapitel 3). In die Startmenge für die Untersuchung werden kognitive Merkmale aufgenommen, die für Informatikaufgaben relevant sind und die in den Wettbewerbsaufgaben, das heißt im Multiple-Choice-Format, überhaupt variieren. Für jedes der ausgewählten Merkmale, Erfahrungsweltnähe, Abstraktionsgrad, Formalisierungsgrad, Komplexitätsgrad, Redundanz, Anforderungsbereich, Prozessbereich, Kognitive Lernzielstufe und Art des Wissens, wird der entwicklungspsychologische Hintergrund und der Bezug zu den Aufgaben beleuchtet (Kapitel 5). Der nächste Schritt ist, die Merkmale empirisch mit Aufgaben und Personen zu verknüpfen.

4.2.2 Datenerhebung: Expertenbefragung, Wettbewerb

Die Erhebungen schließen eine Expertenbefragung und den Informatik-Schülerwettbewerb ein. In der Expertenbefragung werden die Wettbewerbsaufgaben und die Merkmale direkt verknüpft. 45 Personen mit Expertise in der Informatik-Didaktik wird je ein Satz von fünf Aufgaben zur Einschätzung bezüglich der Merkmale vorgelegt. Nur wenn sich ein Konsens oder zumindest eine Mehrheit für eine Einschätzung ergibt, wird sie als Klassifizierung übernommen, sonst bleibt die betreffende Aufgabe ohne Zuordnung zu diesem Merkmal (Abschnitt 6.2). Die solcherart klassifizierten Aufgaben werden im Online-Wettbewerb Informatik-Biber von 21.802 Schülern der Sekundarstufe bearbeitet. Je Schüler werden unter anderem das Geschlecht, das Alter, die Jahrgangsstufe und das Bundesland erhoben. Je Aufgabe wird erfasst, welche Antwortmöglichkeit der Schüler wählt. Über die Aufgaben und die Antworten werden die Schüler indirekt mit den Merkmalen verknüpft (Abschnitt 6.1). Die Erhebungen sind damit abgeschlossen. Der eigentlichen Analyse geht jedoch noch die Aufbereitung der Daten voran.

4.2.3 Datenaufbereitung: Klassifizierung, Antwortmuster, Merkmalsquoten

Die Rohdaten umfassen die Aufgabenklassifizierung und die Teilnahmedaten aus dem Informatikwettbewerb. Die Klassifizierung wird durch Maßnahmen wie die Zusammenlegung von Kategorien und eine Dummy-Kodierung für die Analysen aufbereitet. Für jede Kategorie wird eine dichotome Variable eingeführt, die die Werte „trifft zu“ oder „trifft nicht zu“ annehmen kann. So entstehen aus den Merkmalskategorien eigene Merkmalsvariablen. Die Variablenzahl erhöht sich von neun auf sechzehn. Tabelle 4.1 zeigt den Aufbau der Datentabelle, die für 29 Aufgaben die Klassifizierung nach 16 Merkmalen enthält (Abschnitt 7.1).

Tabelle 4.1: Aufbau der Datentabelle Aufgabenmerkmale (vgl. Tab. 7.1, S. 99)

	16 Merkmale
29 Aufgaben	Klassifizierung

Von den Wettbewerbsdaten interessieren hier weniger die genauen Antworten als die Ergebnisse „Aufgabe korrekt beantwortet“ oder „nicht korrekt beantwortet“, die in den Tabellen der Antwortmuster bereitgestellt werden. Darin gibt es für jede Aufgabe eine Spalte und für jeden Schüler eine Zeile, sein persönliches Antwortmuster. In den drei Altersgruppen fünfte bis siebte (5-7), achte bis zehnte (8-10), elfte und höhere Klassen (11+) sind das 8118, 8225 und 5403 Antwortmuster, den Teilnehmerzahlen entsprechend. Tabelle 4.2 zeigt den Aufbau der Datentabellen, die die kodierten Resultate enthalten (Abschnitt 7.2).

Tabelle 4.2: Aufbau der Datentabellen Antwortmuster (vgl. Tab. 7.2–7.4, S. 101)

	Aufgaben
Schüler	Antwortmuster

Über alle Aufgaben wird in Prozent berechnet, wie viele Aufgaben eines Merkmals ein Schüler löste. Dieser Wert wird Merkmalsquote genannt. Die Tabellen der Merkmalsquoten haben eine Spalte je Merkmal und eine Zeile für jeden der 8118, 8225 und 5403 Schüler in den Altersgruppen 5-7, 8-10 und 11+. Tabelle 4.3 zeigt den Aufbau der Datentabellen, die nach der Kodierung die Merkmalsquoten enthalten (Abschnitt 7.3). Nach dem Abschluss der Kodierungen stehen einige Tausend Datensätze für die Analysen bereit.

Tabelle 4.3: Aufbau der Datentabellen Merkmalsquoten (vgl. Tab. 7.5–7.7, S. 103)

	Merkmale
Schüler	Merkmalsquoten

4.2.4 Statistische Analysen: Dimensionale Struktur

Die strukturentdeckenden Analysen zielen auf drei Modellalternativen ab, die sich hinsichtlich der Analyseebene unterscheiden. Diese sind ein Modell der kognitiven Aufgabenanforderungen und die Modelle A und B der realisierten kognitiven Schülerfähigkeiten.

- Das Anforderungsmodell entsteht durch eine Typisierung der Aufgaben. Um herauszufinden, welche Anforderungstypen unter den Aufgaben vorkommen, werden sie einer Clusteranalyse nach den kognitiven Merkmalen unterzogen (Abschnitt 8.1). Zusätzlich wird die empirische Schwierigkeit der Aufgaben jeweils als Lösungsquote berechnet, das heißt, als Prozentzahl der korrekten Lösungen von den Inangriffnahmen (Abschnitt 8.2).
- Das Fähigkeitsmodell A bündelt die Aufgaben nach den Antwortmustern. Mittels einer Faktorenanalyse werden die Aufgaben, die dieselben Schüler lösten, in Dimensionen zusammengefasst (Abschnitt 8.3).
- Um das Fähigkeitsmodell B zu gewinnen, werden die Merkmale auf Personen bezogen. Eine Faktorenanalyse bündelt die kognitiven Merkmale, deren Quote bei denselben Schülern hoch ausfiel, zu Dimensionen (Abschnitt 8.4).

Die Analysen liefern zunächst tabellarische Ergebnisse, aus denen im Fall der Clusteranalyse die Zuordnung der Aufgaben zu den Typen bereits ablesbar ist. Im Fall der Faktorenanalysen gibt es einen gewissen Spielraum für die Zuordnung. In beiden Fällen bedarf die Zuordnung zu den vermuteten Dimensionen der Interpretation anhand der charakterisierenden Merkmale.

4.2.5 Interpretation: Anforderungen und Fähigkeiten

Die Dimensionen des Anforderungsmodells werden durch Aufgabencluster repräsentiert. Die Merkmalsausprägungen der Aufgaben erlauben eine treffende Charakterisierung der Cluster. Sie werden als Anforderungstypen Wiedergabe, Verständnis, Anwendung und Problemlösung interpretiert (Abschnitt 9.2).

Als Ergebnis der Faktorenanalysen der Antwortmuster liegen Faktoren vor, die wiederum durch Aufgaben repräsentiert werden. Die Fähigkeitsdimensionen des Modells A werden anhand der Merkmalsprofile der Aufgaben interpretiert (Abschnitt 9.3).

Die Faktorenanalyse der Personen-Merkmalsquoten erbringt Faktoren, die unmittelbar durch Merkmale repräsentiert werden, so dass der Zwischenschritt entfällt, Merkmalsprofile der Dimensionen abzuleiten. Die Fähigkeitsdimensionen des Modells B werden anhand der Merkmale interpretiert (Abschnitt 9.4).

4.2.6 Modellprüfung: Anpassungsgüte

Der Interpretation der Modelldimensionen schließt sich als letzter Schritt eine strukturprüfende mehrdimensionale Rasch-Analyse an – strukturprüfend im Gegensatz zu strukturentdeckend, um die Anpassungsgüte an die Stichprobendaten als Indiz für die Geltung der Modellhypothesen zu überprüfen (Abschnitt 10.1). Die Ergebnisse werden in der Modellkritik reflektiert, einbezüglich der methodischen Probleme, die im Verlauf der Arbeit auftreten (Abschnitt 10.2). Das Fazit ist, dass die generelle Vorgehensweise sich bewährt. Am Ende steht eine Empfehlung, welche Modellhypothesen beizubehalten sind, und ein Ausblick auf die nächsten Schritte zur Konkretisierung der Modellvorstellung informatischer Anforderungen und Fähigkeiten (Kapitel 11).

Kapitel 5

Aufgabenmerkmale

Im vergangenen Kapitel wurde die Zielstellung formuliert, zu erproben, inwieweit aus Aufgabenmerkmalen und Schülerantworten mit statistischen Verfahren Strukturmodelle informatischer Anforderungen und Fähigkeiten abgeleitet werden können. In diesem Kapitel wird als erster Schritt des Forschungsplans die theoretische Begründung der Aufgabenmerkmale realisiert.

Die Kriterien, die den kognitiven Anspruch festlegen, verkörpern über die Veranschaulichung von Kompetenzniveaus hinaus die Bausteine, aus denen ein Modell informatischer Anforderungen und Fähigkeiten errichtet wird. Quellen aus der traditionellen Didaktik und Entwicklungspsychologie, aktuelle Prüfungsanforderungen und Zielvereinbarungen der Fächer sowie neue Ansätze der Bildungsforschung wurden ausgewertet, um geeignete Merkmale zu identifizieren. Das Augenmerk lag auf Taxonomien, Kompetenzmodellen, auf denen Bildungsstandards fußen, und Schwierigkeitskriterien zur Charakterisierung kognitiver Aufgabenanforderungen, ansatzweise auch zur Vorhersage der Lösungswahrscheinlichkeit. Auszüge aus der Fülle möglicher kognitiver Merkmale wurden in Kapitel 3 zum Forschungsstand dargestellt. Nicht alle diese Kriterien, die theoretisch zur Definition des Kompetenzkonstrukts beitragen, können in der Untersuchung berücksichtigt werden. Folgende Bedingungen limitieren die Auswahl.

1. Es handelt sich um *Aufgaben-Merkmale*, die unabhängig von der Person des Bearbeiters sind. Variablen des heuristischen Lösungsprozesses etwa, über die ohne Kenntnis des individuellen Lösungsweges keine Aussage getroffen werden kann, scheiden aus. Auch Faktoren wie die Routine, die als Vertrautheit mit dem Löseprozess die Aufgabe erleichtert (Marzano und Kendall 2007), kommen nicht in Betracht, weil sie Merkmale der Person und nicht der Aufgabe sind.
2. Es handelt sich um *informatikrelevante Merkmale*. Variablen wie „Umkehr-aufgabe oder direkte Aufgabe“ des nationalen PISA-Aufgabenmodells (Neubrand et al. 2004) sind in der Informatik nicht relevant.

3. Die *Merkmale variieren* in den Aufgaben, die als Untersuchungsmaterial zur Verfügung stehen, nämlich den Aufgaben eines Online-Informatikwettbewerbs. Diese sind auf wenige Minuten Lese- und Lösezeit und auf das Multiple-Choice-Format beschränkt. Die Adressaten sind Schüler der Sekundarstufe; das inhaltliche und sprachliche Niveau ist auf den Ausbildungs- und Entwicklungsstand der jeweiligen Altersstufe abgestimmt.

Die ersten zwei Bedingungen ergeben sich unmittelbar aus der Zielsetzung, aus kognitiven Merkmalen von Aufgaben ein informatikspezifisches Anforderungs- respektive Fähigkeitsmodell zu entwickeln. Die dritte Bedingung geht auf das Datenmaterial zurück, das ursprünglich für die Durchführung und Auswertung des Informatik-Bibers erhoben wurde. Die Wettbewerbsaufgaben sind nicht auf die Erfordernisse einer empirischen Studie ausgerichtet (vgl. S. 59). Etliche Merkmale, zum Beispiel syntaktische Variablen wie Textmenge, Grammatik, Wortschatz, Fachvokabular oder „Sprachlogische Komplexität“ (Cohors-Fresenborg et al. 2004), variieren innerhalb der Wettbewerbsaufgaben nicht. Der Aspekt „Sprachniveau“ kann daher nicht berücksichtigt werden. Das Gleiche gilt für den Aspekt „Offenheit der Lösungswege und Lösungen“, unter anderem von Draxler (2006) als Schwierigkeitsmerkmal vorgeschlagen, denn offene Aufgaben sind im Multiple-Choice-Format nicht abbildbar.

Wird nun für jedes der identifizierten Merkmale geprüft, ob es die obigen Bedingungen erfüllt, reduziert sich die Menge der in Frage kommenden Merkmale auf neun, die erstens *aufgabeninhärent*, zweitens *informatikrelevant* und drittens *variabel* in den vorliegenden Aufgaben sind.

Weiterhin decken die neun Merkmale die fachdidaktischen Dimensionen ab: Die Auswertung der existierenden normativen Modelle für die Bereiche Informatik, Mathematik und Naturwissenschaften führte zu einem übereinstimmenden Grundmuster fachdidaktischer Dimensionen. Stets ist eine Prozessdimension (a) enthalten, fast immer eine curriculare oder taxonomische Wissensdimension (b), eine Repräsentationsdimension (c) mit formalen Kriterien und eine Kontextdimension (d), meistens die Anwendungssituation (vgl. S. 53). Als Prozessmerkmale, den Denkprozess (a) charakterisierend, werden die Anforderungsbereiche der EPA Informatik (KMK 2004b; Abschnitt 5.6) und die informatikspezifischen Prozessbereiche aus der Empfehlung für Bildungsstandards (GI 2008; Abschnitt 5.7) gewählt. Die Lernzieltaxonomie (Anderson und Krathwohl et al. 2001) liefert Lernzielstufen (Abschnitt 5.8), ebenfalls den Problemlöseprozess (a) betreffend, sowie taxonomische Wissensarten (Abschnitt 5.9), die die Wissensdimension (b) repräsentieren. Die Merkmale Erfahrungsweltnähe (z. B. Klafki 1985; Abschnitt 5.1), Abstraktionsgrad (z. B. Piaget, nach Lefrancois 1986; Abschnitt 5.2), Formalisierungsgrad (z. B. Bruner 1966; Abschnitt 5.3), Komplexitätsgrad (z. B. Goldin 1979; Abschnitt 5.4) und Redundanz (z. B. Prenzel et al. 2002; Abschnitt 5.5) entstammen der Allgemeindidaktik und der Kognitionspsychologie. Als Repräsentationskriterien (c) sind sie auf Informatikaufgaben anwendbar und wirken

nachweislich auf Verständnis und Lösungserfolg. Schließlich werden die Inhaltsbereiche der GI-Empfehlung für Bildungsstandards (GI 2008; Abschnitt 5.10) herangezogen, um den fachlichen Kontext (d) der Aufgaben zu beschreiben. Der Aufgabeninhalt wird als informatives Aufgabenmerkmal mitgeführt, aber nicht als Einflussvariable untersucht. Zwar wird mit Sternberg (1998) angenommen, dass für die Kompetenz wie für die Intelligenz gilt, dass der Inhalt eine Wirkung hat: „that content effects can be as large as or larger than process effects on individual differences“ (Sternberg 1998, S. 131). Getreu der Zielstellung wird hier aber ausdrücklich derjenige Kompetenzanteil untersucht, der allen Informatikinhalten gemeinsam ist (vgl. S. 58).

Auf den kommenden Seiten werden die Kriterien und ihre Kategorien beschrieben. Die Auswahl wird begründet. Im Rahmen einer Expertenbefragung, ausführlich dargestellt in Kapitel 6, werden die Merkmale mit den Aufgaben des Schülerwettbewerbs Informatik-Biber¹ verknüpft (S. 82f, S. 173ff). Im Vorgriff auf die Ergebnisse der Merkmalsklassifizierung werden die Aufgaben bereits in diesem Kapitel zur Verdeutlichung der Kriterien herangezogen.

5.1 Erfahrungsweltnähe

Die Nähe zur Erfahrungswelt, synonym zu Lebensweltnähe oder Realitätsbezug, ist ein Qualitätsmerkmal in der Didaktik. Als Gegenwartsbedeutung wurde sie zuerst 1985 von Klafki als erste Frage der Didaktischen Analyse eingefordert (Klafki 1985, S. 273). Lebensweltnähe ist nicht mit Authentizität zu verwechseln, die Echtheit bedeutet, während ein realitätsbezogener Lernimpuls durchaus erfunden sein kann, lediglich als Schülererfahrung denkbar sein muss.

Unter der Erfahrungsweltnähe einer Aufgabe wird hier verstanden, wie nah sie dem persönlichen Erleben der Schüler ist. Die Anwendungssituation wird betrachtet, nicht der Fachinhalt selbst. So könnte das Thema „Scheduling“ in unterschiedlich lebensweltnahe Situationen Stundenplan, Fahrplan oder Produktionsplanung eingebettet werden. Die Anwendung Stundenplan ist dem Schüleralltag zuzuordnen, das heißt *nah* zur Erfahrungswelt. Die Anwendung Fahrplan ist nicht unbedingt täglich erlebt, aber jederzeit erlebbar durch die Schüler, das heißt, sie weist *mittlere Entfernung zur Erfahrungswelt* auf. Die Anwendung Produktionsplanung ist im Allgemeinen nicht erlebbar durch die Schüler, jedoch vorstellbar, das heißt *entfernt* von der Erfahrungswelt.

Unbezweifelt ist Erfahrungsweltnähe förderlich, wenn es um die Motivation geht. In neueren Arbeiten zur Kompetenzmodellierung wird untersucht, inwieweit sie sich auf die Lösequote auswirkt. Fleischer et al. (2010) identifizieren den persönlichen Bezug als charakteristisches Merkmal von Problemlöseaufgaben im Unterschied zu Mathematikaufgaben (Fleischer et al. 2010, S. 245).

¹www.informatik-biber.de

Während der Expertenklassifizierung der Wettbewerbsaufgaben des Informatik-Bibers ergaben sich Unsicherheiten bei der Einschätzung der Erfahrungsweltnähe, da die Erfahrungswelt sich im Verlauf der Sekundarstufe verändert. Um dem gerecht zu werden, hätte es zumindest einer altersgruppenbezogenen Einschätzung bedurft. Mehr noch, die Lebenswelt ist ein Merkmal des individuellen Schülers, was schon Klafki zu bedenken gibt.

„Insofern kann die Gegenwartsbedeutung eines ins Auge gefassten Ziel- und Themenzusammenhanges – abkürzend angedeutet: ‚Wohnen‘, ‚Freunde‘, ‚Was bedeutet mir meine Familie?‘, ‚Spielzeug‘, ‚Jungen und Mädchen‘, ‚Fahrrad und Radfahren‘, ‚Brauchen wir eigentlich Gewerkschaften?‘ – für Kinder und Jugendliche – kognitiv, emotional und handlungsbezogen – sehr unterschiedlich sein“ (Klafki 1985, S. 273).

Dennoch schätzten die Experten die Biber-Aufgaben bezüglich ihrer Lebensweltnähe weitgehend konform ein, erklärbar dadurch, dass im Zweifel eine mittlere Jahrgangsstufe des Bearbeiters unterstellt wurde. In Kenntnis der starken Personenabhängigkeit, die gegen die erste Bedingung (S. 67) verstößt, geht das Merkmal aber nicht als Variable in die explorativen Analysen ein.

5.2 Abstraktionsgrad

Abstraktes Denken und logisches Schlussfolgern kennzeichnen das vierte Entwicklungsstadium nach Jean Piaget, die Stufe der formalen Operationen (Lefrancois 1986). Das formal-operative Stadium wird mit elf oder zwölf Jahren, also in den ersten Jahren der Sekundarstufe erreicht. Lefrancois illustriert das kognitive Anspruchsniveau mit einer einfach erscheinenden Frage.

„John ist dünner als Bill; John ist dicker als Sam; wer ist der Dickste von den dreien? Kinder, die jünger als 11 oder 12 Jahre sind, haben große Schwierigkeiten mit solchen Aufgaben, außer wenn es sich um Objekte handelt, die sie sehen können. Der Grund ist, daß die Lösung der Aufgabe propositionales Denken erfordert, d. h. Nachdenken über hypothetische Aussagen“ (Lefrancois 1986, S. 138).

Abstrakte Aufgaben verlangen dem Bearbeiter nicht allein bei der Informationsentnahme höher entwickelte Fähigkeiten ab, auch der Lösungsprozess findet auf der Ebene theoretischer Analyse statt. Abstrahieren, mit Abstraktion umgehen und geläufige Wechsel der Abstraktionsebene sind typische Ansprüche in der Informatik, was auch im Duden Informatik (Claus und Schwill 2003) zum Ausdruck kommt.

„Jede Umsetzung in Algorithmen, Programme oder Informatiksysteme erfordert, sich von der Realität zu lösen, ein geeignetes (nur gedachtes, also mit den natürlichen Sinnen nicht erfassbares) Modell zu bilden und dieses einem Computer zur Bearbeitung zu übergeben“ (Claus und Schwill 2003, Duden Informatik, Stichwort ‚abstrakt‘).

Hartmann et al. (2006) sehen darin ein informatikspezifisches Unterrichtsziel. „Für den Unterricht stellt das abstrakte Wesen der Informatik eine besondere Herausforderung dar: Die Fähigkeit zur Abstraktion müssen sich Schülerinnen und Schüler erarbeiten, insbesondere auch die Fähigkeit, Dinge auf verschiedenen Abstraktionsebenen gleichzeitig zu betrachten“ (Hartmann et al. 2006, S. 91). So nennt Kujath (2007) beim Bearbeiten typischer Informatikprobleme ein hohes Abstraktionsniveau als charakteristische Strategie leistungsstarker Problemlöser im Vergleich zu schwächeren Problemlösern, die enaktiv vorgehen und ausschließlich konkret operieren (vgl. S. 42).

Für die Aufgabenstudie wird unter dem Abstraktionsgrad einer Aufgabe der Grad der Loslösung vom konkreten Beispiel verstanden. Lautet die Aufgabe, einen Weg in einem gegebenen Netzwerk zu suchen, handelt es sich um einen konkreten Fall. Die Aufgabe ist *konkret*. Lautet die Aufgabe, mögliche Knotenausfälle in einem Netzwerk durchzuspielen, so handelt es sich um die Betrachtung gedachter Fälle. Die Aufgabe ist von *mittlerem Abstraktionsgrad*. Gilt es, einen Algorithmus zur Wegsuche in Netzwerken zu finden, so handelt es sich um ein allgemeines Verfahren, losgelöst vom Einzelfall. Die Aufgabe ist *abstrakt*.

In den Abbildungen A.13 und A.9 im Anhang sind die Aufgaben Link (S. 184) und Endlosschleife (S. 180) zu sehen. Link enthält eine Wissensfrage über Hyperlinks und ist eine konkrete Aufgabe, Endlosschleife dagegen ist ein typischer Vertreter einer abstrakten Aufgabe, in der ein Flussdiagramm gegeben ist, das einen Algorithmus beschreibt. Um die Frage zu beantworten, muss ein Schüler den Algorithmus nachvollziehen und verstehen, wie sich die Variablenwerte bei jedem Durchlauf verändern.

5.3 Formalisierungsgrad

Wie Abstrahieren sind auch Formalisieren und der Umgang mit Formalismen typisch informatische Anforderungen. Für die Mathematik zeigen Cohors-Fresenborg et al. (2004), dass das Merkmal Formalisierung von Wissen Schwierigkeit generiert (vgl. S. 46). Für die Aufgabenklassifizierung werden als Ausprägungen, die in den Aufgaben zu beobachten sind, folgende Grade der Formalisierung betrachtet. Die Darstellung der Information „Von drei Uhren ist A die größte, B die kleinste und C die mittlere.“ als Text oder als Bild (Abb. 5.1, Bsp. 1) ist *informell*. Die symbolische Repräsentation (Abb. 5.1, Bsp. 2) weist einen *mittleren Formalisierungsgrad* auf. Die modellhafte Darstellung, etwa als Baum oder Struktogramm (Abb. 5.1, Bsp. 3) wird als *formal* angenommen.

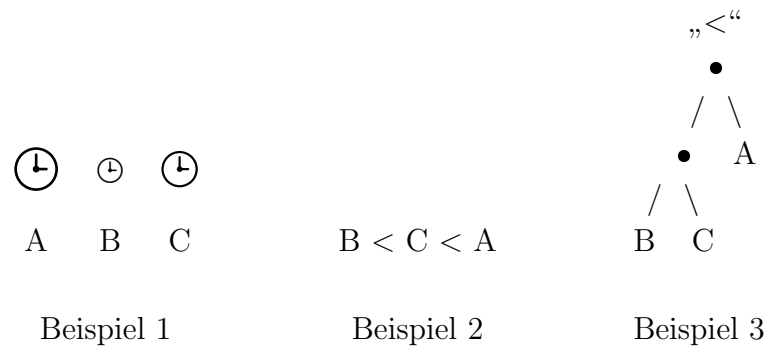


Abbildung 5.1: Informationsdarstellung. Bildliche (Beispiel 1), symbolische (Beispiel 2) und modellhafte (Beispiel 3) Darstellung der Information „Von drei Uhren ist A die größte, B die kleinste und C die Uhr mittlerer Größe.“

Das Merkmal Formalisierungsgrad ist nicht deckungsgleich zu Jerome Bruners Repräsentationsstufen enaktiv, ikonisch und symbolisch (Bruner 1966), aber es bestehen Parallelen. Obwohl Bruners Repräsentationsformen sich beim Kind nacheinander entwickeln, ersetzen sie einander nicht (Lefrancois 1986, S. 110), auch Erwachsene stellen Dinge enaktiv, ikonisch oder symbolisch dar. Der Formalisierungsgrad *informell*, Text oder Bilder, gleicht der ikonischen Stufe Bruners, auf der Information durch geistige Bilder repräsentiert wird. Der *mittlere Formalisierungsgrad*, symbolische Repräsentation, gleicht Bruners Stufe symbolischer Darstellung. Lefrancois grenzt die ikonische und die symbolische Stufe folgender Art gegeneinander ab. „Der fundamentale Unterschied zwischen einem Symbol und einem Ikon ist, dass das Ikon, im Gegensatz zum Symbol, eine tatsächliche Ähnlichkeit zu dem von ihm bezeichneten Objekt besitzt. Ein Symbol ist dagegen völlig willkürlich [...]“ (Lefrancois 1986, S. 110). Der Grad *formal* geht über die symbolische Darstellung des Sachverhaltes hinaus, indem Formalismen beziehungsweise Modelle herangezogen werden, die einen zusätzlichen kognitiven Anspruch bergen.

Das soll die Charakterisierung von Aufgaben wie **Umparken** (S. 192) ermöglichen. Die Aufgabe konfrontiert die Schüler über die Problemlösung hinaus mit einem Formalismus zu deren Beschreibung. In der *formal* eingeschätzten Aufgabe wird durch Garagen und parkende Autos die Datenstruktur Kellerspeicher verbildlicht. Die Aufgabe, eine bestimmte Parkreihenfolge herzustellen, wird mit einer formalen Operation bewerkstelligt, die das Ausparken des vordersten Autos aus einer Garage (Kelleroperation pop) und das Einparken in eine andere Garage (Kelleroperation push) beschreibt.

Die Aufgabeneinschätzung nach dem Formalisierungsgrad führte zu einem Diskurs der befragten Experten, die in Frage stellten, ob die symbolische oder die modellhafte Repräsentation einen höheren Formalisierungsgrad aufweise. Als Konsequenz werden in den Analysen „symbolisch“ und „modellhaft“, die strittig

sind, zu „formal“ vereinigt. Damit wird nur noch unterschieden, ob Formalismen bei der Aufgabenbewältigung von Belang sind oder nicht. Die Zusammenlegung von Kategorien ist nicht grundsätzlich nachteilig für das Vorhaben, schwierigkeitenrelevante Merkmale zu einem Strukturmodell anzuordnen, da es nur vergrößert, nicht verfälscht wird.

5.4 Komplexitätsgrad

Der Komplexitätsbegriff ist gemeinhin sehr weit gespannt. Im Kontext der Kompetenzmodellierung wird beim PISA-Reporting der Naturwissenschaftsergebnisse die Komplexität der Situation und der Schlussfolgerung unterschieden (vgl. S. 47). In der Physik definiert Kauertz (2008) Komplexität über die Inhaltsstruktur: ein Fakt (1), mehrere Fakten (2), ein Zusammenhang (3), mehrere unverbundene (4) bis zu mehreren verbundenen Zusammenhängen (5) oder übergeordnete Konzepte (6) (vgl. S. 48). Im Fach Mathematik bewerten Neubrand et al. (2002) die Komplexität der Modellierung nach den Kompetenzklassen Reproduktion, Verknüpfung, Verallgemeinerung (vgl. S. 45). Cohors-Fresenborg et al. (2004) trennen die sprachlogische Komplexität der Aufgabenstellung von der kognitiven Komplexität des Lösevorgangs (vgl. S. 46).

Biggs und Collis (1982) führen die SOLO-Taxonomie ein, um Lernqualität nach der Komplexität des Erlernten zu beurteilen (SOLO steht für Structure of the Learning Outcome), ob Prestructural, Unistructural, Multistructural, Relational oder Extended Abstract (vgl. S. 24, S. 37). Goldin (1979) stellt eine umfangreiche Liste von Variablen auf, denen ein Effekt auf die mathematische Komplexität zugesprochen wird, und hebt dabei die aufgabeninhärenten Variablen (Problem Complexity Variables, zum Beispiel Total number of states, Length of shortest solution path, Number of blind alleys, Number of possible first moves, Number of goal states, Total number of solution paths) von den Algorithmus- oder Strategiebezogenen Variablen (Variables Defined with Respect to Algorithms or Strategies) ab (Goldin 1979, S. 165f).

Auf diese Auffassung von Problemschwierigkeit als Lösungsraumkomplexität stützt sich die Spezifikation des Aufgabenmerkmals Komplexitätsgrad. Danach wächst der Grad der Komplexität mit der Anzahl der Fälle, Regeln, Vergleiche, Umformungen, die die Größe des Lösungsraums bestimmt. Die unterschiedlich komplexe Gestaltung einer Aufgabe im Kontext Schachspiel könnte wie folgt aussehen. Ist nur ein einzelner Spielzug einer Figur zu betrachten, gilt die Aufgabe als *einfach*. Sind die Spielzüge mehrerer Figuren zu betrachten, ist die Aufgabe von *mittlerer Komplexität*. Ist eine Folge von Spielzügen mehrerer Figuren zu betrachten, gilt die Aufgabe als *komplex*. Die Aufgabe **Bibers Geheimcode** (S. 175), in der ein Vierwortsatz durch die alphabetische Verschiebung der Konsonanten, nicht aber der Vokale, zu verschlüsseln ist, wird als Aufgabe mittlerer Komplexität klassifiziert.

5.5 Redundanz

Die Redundanz ist ein Merkmal der Informationsdarstellung. Wird der gleiche Sachverhalt mehrfach dargestellt, etwa als Text: „Von drei Uhren ist A die größte, B die kleinste und C die mittlere.“ und zusätzlich als Bild, wie im Beispiel 1 in der Abbildung 5.1 (S. 72), so ist die Informationsrepräsentation *redundant*, sonst nicht. Prenzel et al. (2002) sowie Gut et al. (2010) stellen fest, dass in den Naturwissenschaften die Vorgabe einer Grafik oder Bildinformation die Lösung einer Aufgabe zu erleichtern scheint (vgl. S. 48). Das Merkmal Redundanz wurde nur in der Aufgabenklassifizierung berücksichtigt. In die Analysen wird es nicht einbezogen. Von sämtlichen Wettbewerbsaufgaben enthält einzig **Labyrinth** (S. 183) redundante Information: der schrittweisen Erklärung der Methode, durch das Entfernen von Trennwänden aus einem Gitter ein Labyrinth zu erstellen, wird ein Beispiel helfend zur Seite gestellt. Keine weitere Aufgabe enthält redundante Information. Somit variiert das Merkmal Redundanz in den Untersuchungsaufgaben nicht ausreichend. Die dritte Auswahlbedingung (S. 67f) ist nicht erfüllt.

5.6 Anforderungsbereich

Das Merkmal Anforderungsbereich bezieht sich auf die geforderte Aktivität. In den EPA Informatik (KMK 2004b) werden die Anforderungsbereiche Wiedergabe, Anwendung und Problemlösung unterschieden.²

Wiedergabe kann unter anderem eine Reproduktionsleistung wie „Übersetzen in eine andere Darstellungsform in einem wiederholenden Zusammenhang“ oder „Einfaches Erweitern einer vorgegebenen Problemlösung in geübtem Zusammenhang“ (KMK 2004b, S. 10f) sein. In der Aufgabe **Link** (S. 184) ist die Wiedergabe-Anforderung, diejenige Begriffserklärung herauszufinden, die mit dem impliziten Wissen, was ein Link ist, übereinstimmt.

Zu *Anwendung* gehören Reorganisation und Transfer, beispielsweise „Implementieren von Prozeduren, Funktionen und Methoden im vorgegebenen Kontext“ oder „Analysieren eines gegebenen Algorithmus oder Fallbeispiels“ (ebd.). In der Aufgabe **Endlosschleife** (S. 180) besteht die Anwendungs-Anforderung darin, einen Algorithmus mit verschiedenen Abbruchbedingungen durchzuspielen.

Problemlösung kann unter anderem im „Zerlegen eines gegebenen anspruchsvollen Problems in geeignete Teilprobleme“ oder „Entwickeln eines Verfahrens bzw. Algorithmus zur Lösung eines neuen Problems“ (ebd.) bestehen. So kann die Aufgabe **Labyrinth** (S. 183) nicht durch bloße Anwendung der Labyrinth-Methode gelöst werden. Sie erfordert den Einsatz einer Heuristik, zum Beispiel die Rückwärtsanwendung der Methode, und wird daher dem Anforderungsbereich Problemlösung zugeordnet.

²Die Spezifikation der Anforderungsbereiche ist in Anhang B abgedruckt (S. 197).

5.7 Prozessbereich

Der Prozessbereich bezieht sich wie das Merkmal Anforderungsbereich auf die geforderte Aktivität. Die Prozessbereiche sind vielschichtig, doch werden sie hier bewusst nicht in ihr kognitives Spektrum zerlegt. Sie stehen ganzheitlich für spezifisch informatische Tätigkeitsausprägungen. Im Entwurf der Bildungsstandards Informatik (GI 2008) werden fünf Prozessbereiche unterschieden, Modellieren und Implementieren, Begründen und Bewerten, Strukturieren und Vernetzen, Kommunizieren und Kooperieren, Darstellen und Interpretieren.³

Modellieren und Implementieren beinhaltet das Erstellen, Implementieren und Reflektieren informatischer Modelle zu gegebenen Sachverhalten (GI 2008, S. 19). Für Schülerinnen und Schüler der Jahrgangsstufen 8 bis 10 heißt das zum Beispiel, sie modellieren reale Automaten mithilfe von Zustandsdiagrammen oder setzen einfache Datenmodelle in relationale Modelle um und realisieren diese mit einem Datenbanksystem. Die Aufgabe POP und PUSH (S. 187) wird dem Prozessbereich Modellieren und Implementieren zugeordnet. Sie verlangt, ganz ähnlich zur Aufgabe Umparken (S. 192), die Implementierung der Belegungsänderung eines Kellers mit der zusammengesetzten Operation „popush“.

Begründen und Bewerten beinhaltet, Nutzungsentscheidungen zu begründen, informatische Sachverhalte zu hinterfragen und zu bewerten (GI 2008, S. 19f). Für Schülerinnen und Schüler der Stufe 8 bis 10 heißt das, sie formulieren angemessene Bewertungskriterien und wenden diese an. Die Aufgabe Primärschlüssel (S. 188) wird in die Kategorie Begründen und Bewerten eingeordnet. Um den Fehler in einer Datentabelle zu entdecken, müssen die Schüler erkennen, dass der Primärschlüssel nicht eindeutig ist, wenn ein Wert doppelt vorkommt.

Strukturieren und Vernetzen beinhaltet, Sachverhalte durch Zerlegen und Anordnen zu strukturieren und Verbindungen außerhalb und innerhalb der Informatik zu erkennen und zu nutzen (GI 2008, S. 20). Für Schülerinnen und Schüler der Stufe 8 bis 10 heißt das, sie planen Arbeitsabläufe und Handlungsfolgen, ordnen Sachverhalte hierarchisch an oder erstellen netzartige Strukturen. Die Herausforderung in der Aufgabe Dateisuche (S. 178) ist zum Beispiel, anhand dreier Suchbeispiele mit ihren Ergebnissen die Funktionsweise einer Textsuche mit regulären Ausdrücken zu verstehen.

Kommunizieren und Kooperieren beinhaltet, fachgerecht über informatische Sachverhalte zu kommunizieren, bei der Lösung informatischer Probleme zu kooperieren und dazu Werkzeuge zu nutzen (GI 2008, S. 21). Für Schülerinnen und Schüler der Jahrgangsstufen 8 bis 10 heißt das, sie kooperieren in Projekten und nutzen synchrone und asynchrone Kommunikationsmöglichkeiten zum Informationsaustausch. In diese Kategorie fällt die Aufgabe Private E-Mail (S. 189), die die sachgerechte Nutzung des E-Mail-Dienstes betrifft und Kenntnisse einfacher E-Mail-Optionen voraussetzt.

³Die Spezifikation der Prozessbereiche ist in Anhang B abgedruckt (S. 200).

Darstellen und Interpretieren beinhaltet, unterschiedliche Darstellungen von Sachverhalten zu interpretieren, informatische Sachverhalte zu veranschaulichen und geeignete Darstellungsformen auszuwählen (GI 2008, S. 22). Für Schülerinnen und Schüler der Jahrgangsstufen 8 bis 10 heißt das zum Beispiel, sie interpretieren Diagramme, Grafiken, Ergebnisdaten oder wenden informatische Werkzeuge zum Erstellen von Diagrammen und Grafiken an. In *Binärbaum* (S. 177) wird die Baumdarstellung eines Ausdrucks mit Klammern der Darstellung als Zeichenkette gegenübergestellt. Die Aufgabe ist, einer gegebenen Zeichenkette den passenden Binärbaum zuzuordnen. Das erfordert die Interpretation der Graphen als Ausdrücke, somit trifft die Spezifikation des Prozessbereichs *Darstellen und Interpretieren* zu.

5.8 Lernzielstufe

„Kognitiver Prozess“ und „Wissen“, die Dimensionen der Lernzieltaxonomie von Anderson und Krathwohl et al. (2001) gehen als potentielle Schwierigkeitsmerkmale in die Untersuchungsmenge ein. Allerdings wird die zweidimensionale Anordnung nicht übernommen. Vielmehr gehen die Lernzielstufen der Prozessdimension und die Wissensarten der Wissensdimension getrennt als Merkmale ein. Auch die unsichere These hierarchischer Lernziele im Sinne wachsender Komplexität und Abstraktheit wird nicht übernommen. Die Lernzielstufe wird als Merkmal mit den sechs Ausprägungen *Erinnern*, *Verstehen*, *Anwenden*, *Analyisieren*, *Bewerten* und *Erschaffen* berücksichtigt, und die Art des Wissen mit den vier Ausprägungen *Fakten-*, *Konzept-* und *Prozedurwissen* sowie *Metakognition*.

Die Lernzielstufe *Erinnern* (Remember – Retrieve relevant knowledge from long-term memory, Anderson und Krathwohl 2001, S. 31) ist erreicht, wenn relevantes Wissen aus dem Gedächtnis abgerufen wird. Die Aufgabe *Fenster schließen* (S. 182), die rudimentäre Bedienerfahrung mit graphischen Benutzeroberflächen abfragt, spricht die Lernzielstufe *Erinnern* an.

Die Lernzielstufe *Verstehen* (Understand – Construct meaning from instructional messages, including oral, written, and graphic communication, ebd.) ist erreicht, wenn den Angaben und Anweisungen einer Aufgabe ihre Bedeutung entnommen wird. Auf *Primärschlüssel* (S. 188) trifft das zu. In der Angabe befinden sich lösungskritische Informationen, nämlich dass ein Primärschlüssel Datensätze eindeutig identifiziert und „Schüler ID“ ein Primärschlüssel sein soll.

Die Lernzielstufe *Anwenden* (Apply – Carry out or use a procedure in a given situation, ebd.) ist erreicht, wenn ein Verfahren in einem gegebenen Kontext korrekt angewendet wird, wie in der Aufgabe *Bibers Geheimcode* (S. 175). Darin wird eine gegebene, einfache Verschlüsselung auf eine kurze Nachricht angewendet.

Die Lernzielstufe *Analyisieren* (Analyze – Break material into constituent parts and determine how parts relate to one another and to an overall structure or purpose, ebd.) ist erreicht, wenn ein Ganzes in seine elementaren Bestandteile zerlegt

und deren Zusammenhalt, Struktur und Zweck erkannt werden. Das gilt für die Aufgabe **Platzwechsel** (S. 186), in der die Schrittzahl eines einfachen Sortierverfahrens für ein gegebenes Sortierproblem zu bestimmen ist.

Die Lernzielstufe *Bewerten* (Evaluate – Make judgments based on criteria and standards, ebd.) ist erreicht, wenn auf der Basis von Argumenten und Normen geurteilt wird, so wie in der Aufgabe **Sicheres Passwort** (S. 191), in der es darum geht, das kombinatorisch möglichst schwierige Passwort auszuwählen.

Die Lernzielstufe *Erschaffen* (Create – Put elements together to form a coherent or functional whole; reorganize elements into a pattern or structure, ebd.) ist erreicht, wenn Elemente zu einem kohärenten Ganzen zusammengeführt oder zu einem neuen Muster umgeordnet werden, zum Beispiel algorithmische Bausteine zu einem neuen Verfahren.

5.9 Art des Wissens

Unabhängig von der Prozessdimension wird die Wissensdimension der revidierten Taxonomie nach Anderson und Krathwohl et al. (2001) als Aufgabenmerkmal untersucht. Sie umfasst vier Arten der Information, auf die Lernende aktiv zugreifen, um neues Wissen zu konstruieren. „This move away from passive views of learning toward more cognitive and constructivist perspectives emphasizes what learners **know** (knowledge) and how they **think** (cognitive processes) about what they know as they actively engage in meaningful learning“ (Anderson und Krathwohl 2001, S. 38, Hervorhebungen im Original).

Faktenwissen (Factual Knowledge – The basic elements students must know to be acquainted with a discipline or solve any of the problems in it, Anderson und Krathwohl 2001, S. 46) besteht im Grundwissensschatz einer Domäne. Die Aufgabe **Private E-Mail** (S. 189) setzt das Wissen voraus, dass in E-Mails eine Adresse für eine Blindkopie angegeben werden kann. Deren Empfänger wird weder dem Hauptadressaten noch den Adressaten für Kopien übermittelt. **Private E-Mail** ist eine typische Aufgabe, deren Lösung ohne Faktenwissen kaum möglich ist.

Konzeptwissen (Conceptual Knowledge – The interrelationships among the basic elements within a larger structure that enable them to function together, ebd.), die Kenntnis von Gesetzmäßigkeiten, Schemata, mentalen Modellen und Theorien, wird zur Ordnung von Wissen benötigt. Zur Lösung der Aufgabe **POP und PUSH** (S. 187) ist es notwendig, das Kellerspeicherkonzept „nur die jeweils vorderste Tonne ist zugänglich“ zu durchdringen, somit entstammt die Aufgabe dem Bereich Konzeptwissen.

Prozedurwissen (Procedural Knowledge – How to do something, methods of inquiry, and criteria for using skills, algorithms, techniques, and methods, ebd.) umfasst Routinen, Fertigkeiten und Algorithmen zur Problemlösung. Ein einfaches Beispiel einer Aufgabe aus dem Bereich Prozedurwissen ist **Platzwechsel** (S. 186), die das Bubblesortverfahren thematisiert.

Metakognition (Metacognitive Knowledge – Knowledge of cognition in general as well as awareness and knowledge of one’s own cognition, ebd.) ist Wissen *über* Kognition. Nach Sternberg (1998), der als Intelligenzforscher die Bedeutung von Metakognition für den Erwerb und die Entwicklung von Expertise untersucht, ist Metakognition der Schlüssel zur Verwirklichung aller kognitiven Fähigkeiten und steht in einer Wechselbeziehung mit anderen, etwa motivationalen Kompetenzaspekten. So spricht einerseits jede einzelne der Wettbewerbsaufgaben implizit metakognitive Fähigkeiten an, andererseits ist keine darunter, die in der Expertenbefragung mehrheitlich dem Wissensbereich Metakognition zugeordnet wird.

5.10 Inhaltsbereich

Im Entwurf der Bildungsstandards für die Informatik in der Schule (GI 2008) werden fünf Inhaltsbereiche vorgeschlagen, Information und Daten, Algorithmen, Sprachen und Automaten, Informatiksysteme sowie Informatik, Mensch und Gesellschaft, die den Fachkontext der Aufgaben benennen.⁴ Die Inhaltsbereiche sind sehr breit. In der Kategorie Information und Daten zum Beispiel befinden sich Aufgaben wie **Biberszahlen** (S. 176) zum Dualsystem und auch Aufgaben wie **Binärbaum** (S. 177) zu Baumstrukturen. Da das allzu stark vereinfachend erscheint, wird zusätzlich zum Inhaltsbereich das Aufgabenthema erfasst: Algorithmen, Anwenderwissen, Automaten, Datenbanken, Datenstrukturen, Graphen, Kodierung, Logisches Schließen, Benutzerverhalten oder Sprachen.

5.11 Zusammenfassung

Das Ergebnis dieses Kapitels ist eine Ausgangsmenge kognitiver Merkmale für die anschließenden Analysen der Merkmalsstruktur von Anforderungs- und Fähigkeitsmustern in der Informatik. Die Auswahl der Merkmale wurde anhand der Bedingungen getroffen, dass sie von der Aufgabe, nicht von der Person des Bearbeiters abhängen (1) und dass sie auf Informatikprobleme zutreffen (2). Sie müssen innerhalb der zur Verfügung stehenden Aufgaben variabel sein (3) und die Aspekte des Grundmusters fachdidaktischer Dimensionen ansprechen, Prozess, Wissen, Repräsentation und Kontext (vgl. S. 53).

- Abstraktionsgrad: konkret bis abstrakt
- Formalisierungsgrad: informell oder formal
- Komplexitätsgrad: einfach bis komplex
- Anforderungsbereich: Wiedergabe; Anwendung; Problemlösung

⁴Die Spezifikation der Inhaltsbereiche ist in Anhang B abgedruckt (S. 199).

- Prozessbereich: Modellieren und Implementieren; Begründen und Bewerten; Strukturieren und Vernetzen; Kommunizieren und Kooperieren; Darstellen und Interpretieren
- Lernzielstufe: Erinnern; Verstehen; Anwenden; Analysieren; Bewerten; Erschaffen
- Art des Wissens: Faktenwissen; Konzeptwissen; Prozedurwissen; Metakognition
- Inhaltsbereich: Information und Daten; Algorithmen; Sprachen und Automaten; Informatiksysteme; Informatik, Mensch und Gesellschaft

Die Merkmale Erfahrungsweltnähe und Redundanz wurden zunächst in die Ausgangsmenge aufgenommen, aufgrund der Erkenntnisse aus der Aufgabenklassifizierung durch die Experten jedoch wieder entfernt. Die Nähe zur Erfahrungswelt ist nicht allein von der Aufgabe, sondern auch von der Person des Bearbeiters abhängig. Das Merkmal Redundanz variiert innerhalb der Aufgabengrundmenge nicht genügend. Andere Merkmale, wie das Sprachniveau oder die Offenheit der Aufgabenstellung, können nicht betrachtet werden, da sie in der Aufgabengrundmenge konstant sind. Das ist insofern kritisch zu sehen, als die Modellhypothesen, die aus den Analysen der Schülerdaten resultieren, nur diejenigen Kompetenzkomponenten verkörpern können, die durch die Aufgabenmerkmale angesprochen werden.

Weiter muss gesagt werden, dass die Merkmale teilweise überlappen oder gar Abhängigkeiten aufweisen. Zum Beispiel tendieren formale Aufgaben dazu, in gleichem Maß abstrakt zu sein. Abstraktion und Komplexität hängen zusammen, denn Abstraktion ist ein probater Weg zur Bewältigung von Komplexität. Die Anforderungsbereiche sind mit den Lernzielstufen verwandt, da die Stufen der Bloomschen Taxonomie in der Definition der Anforderungsbereiche verwendet wurden. Andererseits kann keines der Merkmale ein anderes vollständig ersetzen. So ergibt sich als nächstes Ziel, das Merkmalsgeflecht zu durchdringen und die Abhängigkeiten sichtbar zu machen.

Während der Darstellung der Merkmalsauswahl wurde einige Male auf die Ergebnisse der Aufgabenklassifizierung vorgegriffen. Das folgende Kapitel beschreibt die Datengrundlage, die aus den Aufgaben und Durchführungsdaten des Schülerwettbewerbs Informatik-Biber gebildet wird. Der nächste Schritt ist die Klassifizierung der Untersuchungsaufgaben nach den Merkmalen Abstraktionsgrad, Formalisierungsgrad, Komplexitätsgrad, Anforderungsbereich, Prozessbereich, Lernzielstufe, Wissensart und Inhaltsbereich.

Kapitel 6

Erhebungen

Thema des vorangehenden Kapitels war die Auswahl der Merkmale, die die Bedingungen erfüllen, aufgabeninhärent und informatikrelevant zu sein und in den Untersuchungsdaten zu variieren (vgl. S. 67f). Der nächste Schritt ist die Datenerhebung. In diesem Kapitel wird der Informatikwettbewerb vorgestellt, der die Aufgaben und die Schülerdaten liefert, und die Expertenbefragung, die zur Klassifizierung der Wettbewerbsaufgaben durchgeführt wurde.

Das Vorhaben, Strukturhypothesen informatischer Kompetenz empirisch zu untermauern, ist nur über Informatikaufgaben und die Beobachtung der Problembearbeitung realisierbar. Hierfür stehen die Daten des Schülerwettbewerbs Informatik-Biber zur Verfügung, dessen inhaltliche Ausrichtung ideal zur Zielsetzung passt. Aus den kognitiven Merkmalen der Wettbewerbsaufgaben wird ein informatisches *Anforderungsmodell* abgeleitet; auf Basis der Antwortmuster werden *Fähigkeitsmodelle* entwickelt. In diesem Kapitel wird die Datenerhebung beschrieben. Danach liegen die Wettbewerbsdaten, aufgeschlüsselt nach Teilnehmer und Aufgabe, und die Aufgabenklassifizierung nach kognitiven Kriterien vor.

Die Arbeit versteht sich als Erprobung des Vorgehens, Informatikaufgaben nach Schwierigkeitsmerkmalen zu klassifizieren, um aus der Klassifizierung und der Aufgabebearbeitung durch Versuchspersonen Erkenntnisse über die Struktur informatischer Anforderungen und Fähigkeiten zu gewinnen. In der Informatik-Didaktik ist dieser Weg noch nicht erforscht. Die Maßgabe ist, unter den gegebenen Bedingungen das Verfahren exemplarisch von der Einführung der Schwierigkeitskriterien bis zur Interpretation der Strukturmodelle zu durchlaufen. Das Fach Informatik kann nicht auf eine empirische Tradition zurückgreifen, wie andere Fächer sie besitzen. Empirisches Material aus Bildungsstudien, wie sie in der Mathematik seit Jahrzehnten durchgeführt werden, existiert in Deutschland für die Informatik noch kaum. Aus Machbarkeitsgründen werden in diesem ersten Versuch keine eigenen Aufgaben konstruiert und keine eigenen Schülerdaten erhoben. Stattdessen wird das vorhandene, umfangreiche Datenmaterial des Schülerwettbewerbs Informatik-Biber genutzt. Dazu werden die Wettbewerbsaufgaben durch Personen mit Expertise in der Informatik-Didaktik, meist Lehrende, nach den Un-

tersuchungsmerkmalen eingeordnet. Unter den sich bietenden Möglichkeiten werden die Modalitäten der Expertenbefragung und der Aufgabenbewertung stets so gewählt, dass für jede Aufgabe eine Klassifizierung ausgewiesen wird, damit alle Aufgaben bei der Modellfindung verwendet werden können.

Der Schülerwettbewerb Informatik-Biber wird in Abschnitt 6.1 vorgestellt. In Abschnitt 6.2 werden die Modalitäten und Ergebnisse der Expertenbefragung erläutert, in deren Rahmen die Aufgaben mit den Merkmalen verknüpft wurden (Schlüter 2008; 2009).

6.1 Schülerwettbewerb Informatik-Biber

Der Informatik-Biber¹ ist Teil der internationalen Initiative des Bebras Contest² (International Contest on Informatics and Computer Fluency), die mit einem leicht zugreifbaren und motivierenden Online-Quiz das Interesse von Lehrern und Schülern aller Altersstufen an der Informatik fördern will. Im Jahr 2013 haben 30 Länder an dem jährlichen Wettbewerb teilgenommen. International existiert ein Aufgabenpool, der für jede Altersstufe optionale und Pflichtaufgaben enthält. Ausführlicher schildern Dagiené und Futschek (2008) die Wurzeln, die Intention und die Geschichte der Bebras Initiative. In Deutschland bildet der Informatik-Biber das Einstiegsformat zum Bundeswettbewerb Informatik³, der 1980 ins Leben gerufen wurde. Die Teilnahme am Bundeswettbewerb, der auf die intensive Beschäftigung mit der Informatik abzielt, erfordert einen wesentlich höheren Zeitaufwand als die Teilnahme am Informatik-Biber. Dieser bietet schnell zu erfassende, möglichst spannende Aufgaben, deren Lösung im Normalfall kein spezifisches Vorwissen über Basis-Bedienfertigkeiten des Computers und Mathematik-Grundkenntnisse hinaus erfordert, wie sie bereits in der Primarstufe angelegt werden. Vielmehr sollen informatische Konzepte angesprochen und informatische Denkmuster angeregt werden (Pohl et al. 2009). Der Wettbewerb wird seit dem Jahr 2006 für Sekundarstufenschüler aller Schularten ausgerichtet, also Mädchen und Jungen der Klassenstufen fünf bis dreizehn. Er findet in einer festgelegten Woche statt, in der die Schüler der angemeldeten Schulklassen oder Kurse in einer Schulstunde online die Aufgaben bearbeiten. Träger des Informatik-Bibers sind die Gesellschaft für Informatik (GI)⁴, der Fraunhofer-Verbund IuK-Technologie⁵ und das Max-Planck-Institut für Informatik⁶. Der Wettbewerb wird vom Bundesministerium für Bildung und Forschung⁷ gefördert.

¹www.informatik-biber.de

²www.bebas.org

³www.bundeswettbewerb-informatik.de

⁴www.gi.de

⁵www.iuk.fraunhofer.de

⁶www.mpi-inf.mpg.de

⁷www.bmbf.de

Tabelle 6.1: Teilnahmedaten. Klassenstufe (5-7, 8-10, 11+); Resultate (richtig beantwortet: „korrekt“, falsch beantwortet: „inkorrekt“, nicht beantwortet: „unbeantwortet“, Aufgabe wurde nicht gestellt: „lag nicht vor“).

Teilnehmer	Stufe	Aufgabe 1	Aufgabe 2	...	Aufgabe 29
1	8-10	lag nicht vor	korrekt		inkorrekt
2	8-10	lag nicht vor	unbeantwortet		korrekt
⋮					
21.746	8-10	lag nicht vor	korrekt		korrekt

Die Studie verwendet die Aufgaben und anonymisierten Teilnahmedaten des Informatikwettbewerbs 2007, an dem 21.802 Schülerinnen und Schüler teilnahmen. Das Datenmaterial umfasst Geschlecht, Bundesland und Jahrgangsstufe sowie das Bearbeitungsergebnis jedes Teilnehmers für jede der fünfzehn Aufgaben, die in einer Klassenstufe gestellt wurden, wobei die Klassenstufen fünf bis sieben (5-7), acht bis zehn (8-10) und elf oder höher (11+) unterschieden werden. Die Resultate lauten „korrekt“ im Falle der richtigen Antwort, „inkorrekt“ im Falle einer falschen Antwort, „unbeantwortet“, falls die Antwort ausblieb, oder „lag nicht vor“, falls die Aufgabe in der betreffenden Klassenstufe nicht gestellt wurde. Tabelle 6.1 enthält einen Auszug der Teilnahmedaten. Von den 21.802 Einsendungen waren 21.746 Datensätze gültig.

Im Informatik-Biber 2007 wurden insgesamt neunundzwanzig Aufgaben gestellt.⁸ Tabelle 6.2 gibt einen Überblick über die Themen und Inhaltsbereiche.⁹ Stärkster Inhaltsbereich ist Information und Daten mit zehn Aufgaben, davon zwei zum Thema Datenstrukturen, vier zu Graphen, vier zu Codierung. Sieben Aufgaben entfallen auf den Bereich Algorithmen. Fünf Aufgaben werden dem Bereich Sprachen und Automaten zugeordnet, eine davon Sprachen, eine Automaten, drei dem Logischen Schließen. Die zugrunde liegende GI-Empfehlung für Standards in der Schulinformatik (GI 2008) bezieht Logisches Schließen als Thema nicht explizit ein. Als formales Kalkül wird es hier den Sprachen zugeordnet. Von den fünf Aufgaben des Inhaltsbereichs Informatiksysteme betreffen vier das Thema Anwenderwissen und eine das Thema Datenbanken. Informatik, Mensch und Gesellschaft ist nur mit zwei Aufgaben vertreten, beide zum Benutzerverhalten. Abbildung 6.1 (S. 85) stellt die Verteilung der Inhaltsbereiche in einem Kuchendiagramm dar.

Für jede Aufgabe wird in Tabelle 6.2 festgehalten, in welchen Jahrgangsstufen sie vorlag. Teilweise wurde die gleiche Aufgabe in zwei Stufen gestellt, Computervirus (S. 178) und Schnitzeljagd (S. 190) sogar in allen drei Stufen. Das bein-

⁸Die Aufgaben des Informatik-Bibers 2007 sind in Anhang A abgedruckt (S. 173).

⁹Die Inhaltsbereiche laut GI-Empfehlung werden in Anhang B dargestellt (S. 199).

Tabelle 6.2: Wettbewerbsaufgaben. Inhaltsbereich; Thema; Jahrgangsstufe, in der die Aufgabe gestellt wurde.

Aufgabe	Inhaltsbereich	Thema	Stufe
1 Biber am Fluss	Information/Daten	Graphen	5-7
2 Biber und Bisons	Sprachen/Automaten	Schließen	8-10, 11+
3 Bibers Geheimcode	Information/Daten	Kodierung	5-7, 8-10
4 Biberzahlen	Information/Daten	Kodierung	8-10, 11+
5 Binärbaum	Information/Daten	Graphen	8-10
6 Computervirus	Mensch/Gesellschaft	Verhalten	5-7, 8-10, 11+
7 Dateisuche	Sprachen/Automaten	Sprachen	5-7
8 Dino-Ordnung	Sprachen/Automaten	Schließen	5-7
9 Endlosschleife	Algorithmen	Algorithmen	8-10, 11+
10 Falschgeld	Algorithmen	Algorithmen	8-10, 11+
11 Fenster schließen	Informatiksysteme	Anwenderwis.	5-7
12 Labyrinth	Algorithmen	Algorithmen	11+
13 Link	Informatiksysteme	Anwenderwis.	5-7
14 Links um!	Algorithmen	Algorithmen	5-7
15 Morse-Code	Information/Daten	Kodierung	11+
16 Netzwerkkabel	Information/Daten	Graphen	8-10, 11+
17 Platzwechsel	Algorithmen	Algorithmen	5-7
18 POP und PUSH	Information/Daten	Datenstrukt.	11+
19 Primärschlüssel	Informatiksysteme	Datenbanken	8-10, 11+
20 Private E-Mail	Informatiksysteme	Anwenderwis.	8-10
21 Schnitzeljagd	Sprachen/Automaten	Automaten	5-7, 8-10, 11+
22 Sicheres Passwort	Mensch/Gesellschaft	Verhalten	5-7, 8-10
23 Umparken	Information/Daten	Datenstrukt.	5-7, 8-10
24 Ungeschützter PC	Informatiksysteme	Anwenderwis.	5-7, 11+
25 Verschlüsselung	Information/Daten	Kodierung	8-10, 11+
26 Verwandlung	Information/Daten	Graphen	11+
27 Wertetausch	Algorithmen	Algorithmen	11+
28 Wetter	Sprachen/Automaten	Schließen	5-7
29 Zahlenreihe	Algorithmen	Algorithmen	5-7, 8-10

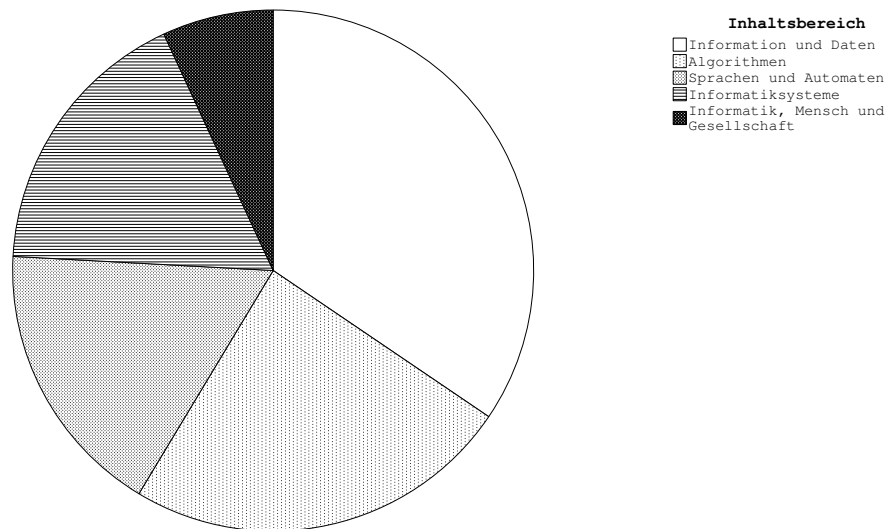


Abbildung 6.1: Verteilung der Inhaltsbereiche der Wettbewerbsaufgaben

haltet keine Aussage über die Eignung oder besser Nichteignung für bestimmte Altersstufen. Es kam zum Beispiel aus Gründen der inhaltlichen Ausgewogenheit vor, dass eine Aufgabe den jüngeren (Klassenstufe 5-7) und älteren (Klassenstufe 11+), nicht aber den übrigen Schülern (Klassenstufe 8-10) gestellt wurde.

Die Teilnahmedaten (Tabelle 6.1) und der Aufgabenüberblick (Tabelle 6.2) können den Wettbewerbsdaten ohne Weiteres entnommen werden. Zur Vervollständigung der Eingangsdaten für die Strukturanalyse informatischer Anforderungen und Fähigkeiten ist weitere empirische Arbeit erforderlich. Um die Merkmalsstruktur der Aufgaben zu analysieren, muss Aufgabe für Aufgabe geklärt werden, welche kognitiven Merkmale sie trägt. Das geschieht in einer Expertenbefragung.

6.2 Expertenbefragung

Die Aufgabenklassifizierung stützt sich auf eine Expertenbefragung. Um eine möglichst objektive Einordnung in das Kriterienschema zu erreichen, wurde jede Aufgabe durch mehrere Personen mit Expertise in der Informatik-Didaktik eingeschätzt. Die Expertengruppe setzt sich aus Personen zusammen, die an Hochschulen in der Informatik-Didaktik forschen und lehren, an Schulen Informatik unterrichten oder in Initiativen zur Informatik-Förderung, wie dem Bundeswettbewerb Informatik, aktiv sind. Von den fünfundvierzig Befragten sind neun weiblich. Einundzwanzig unterrichten in der Sekundarstufe, davon vierzehn auch an der Universität. Zweiundzwanzig sind nur an der Universität, zwei sind außerhalb der Lehre tätig. Nur jeweils einer der Befragten ordnete sich der Altersgruppe

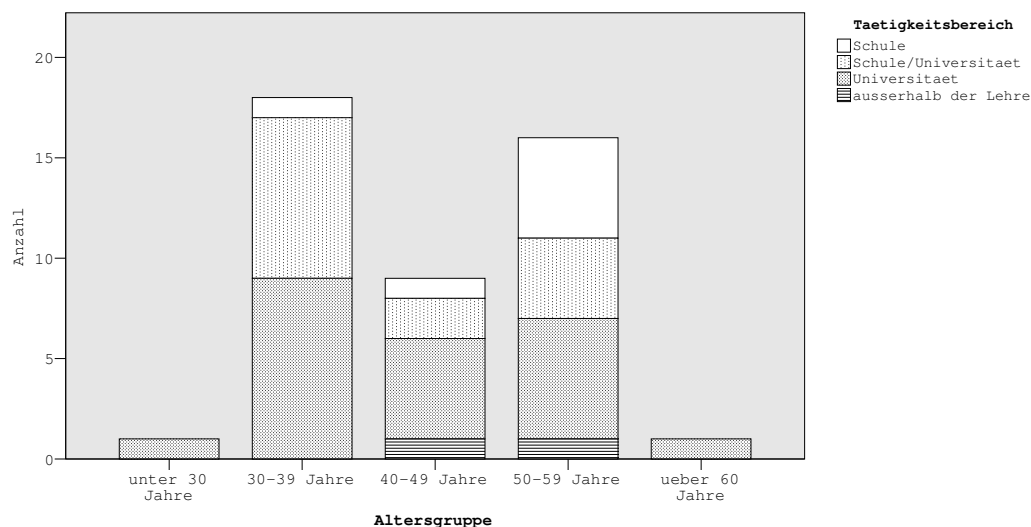


Abbildung 6.2: Altersgruppe und Tätigkeitsbereich der 45 Experten

unter 30 oder über 59 Jahre zu. Achtzehn Personen waren zum Zeitpunkt der Umfrage 30 bis 39 Jahre alt, gefolgt von sechzehn Personen mit 50 bis 59 Jahren und neun aus der mittleren Altersgruppe, 40 bis 49 Jahre. Die Balkengrafik in Abbildung 6.2 zeigt die Verteilung von Altersgruppe und Tätigkeitsbereich der Experten.

Die Kommunikation mit den Experten wurde ausschließlich schriftlich geführt. Die erste Kontaktaufnahme fand per Mail statt. Die Fragebögen lagen auf Papier vor und sowohl Versand als auch Rücklauf erfolgte per Post.¹⁰ Im vorigen Kapitel 5 wurden die Merkmale Erfahrungsweltnähe, Abstraktionsgrad, Komplexitätsgrad, Formalisierungsgrad, Redundanz, Anforderungsbereich, Prozessbereich, Lernzielstufe und Wissensart als Kriterien spezifiziert. Die Teilnehmer an der Umfrage erhielten ein Hinweisblatt zur Einordnung der Aufgaben mit einer Erläuterung der Kriterien und Aufgabenprofile in Form von Fragebögen mit je einer Aufgabe und den möglichen Merkmalswerten zum Ankreuzen. Zur Orientierung bekamen die Befragten außerdem drei Beispielprofile, Aufgaben aus einem vergangenen Wettbewerb, deren Einordnung in einer Vorstudie innerhalb der Arbeitsgruppe Didaktik um Torsten Brinda an der Universität Erlangen-Nürnberg abgestimmt wurde. Weil die Einarbeitung in die Kriterien und die sorgfältige Einschätzung je Aufgabe einen beträchtlichen Zeitaufwand erfordern, wurden die neunundzwanzig Aufgaben in Arbeitseinheiten zu je fünf Aufgaben aufgeteilt. Eine Arbeitseinheit enthielt Aufgaben nur einer der Jahrgangsstufen 5-7, 8-10 oder 11+. Jeder Experte übernahm eine solche Arbeitseinheit, so dass 225 (45 x 5) Aufgabenprofile entstanden.

¹⁰Ein Satz Unterlagen wie versendet ist in Anhang C abgedruckt (S. 201).

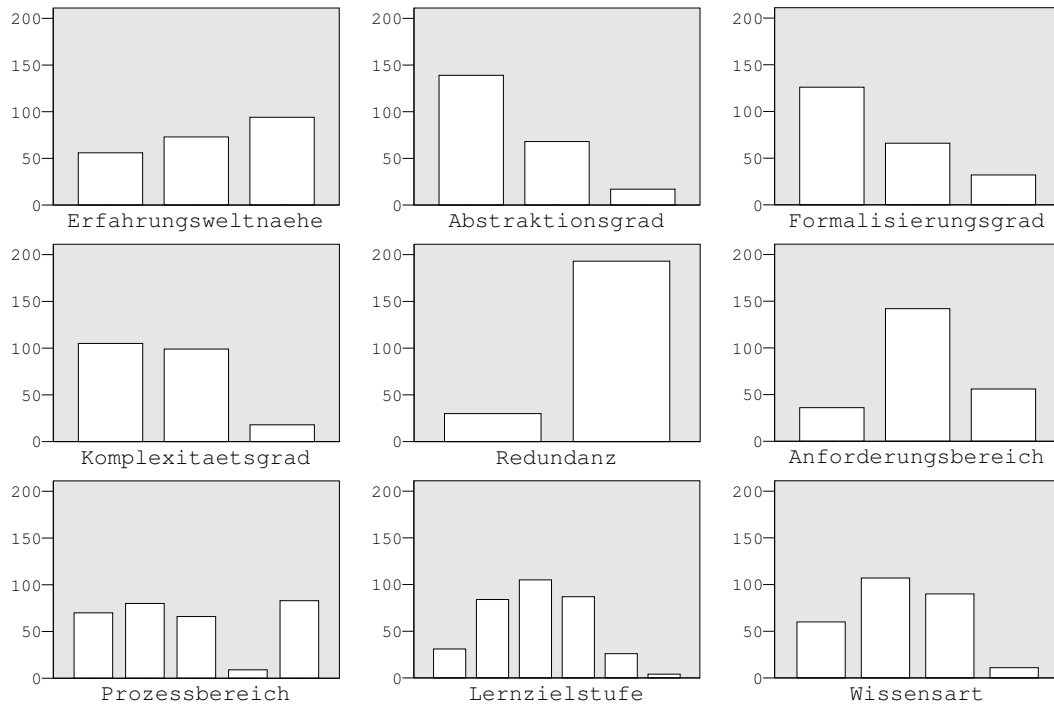


Abbildung 6.3: Häufigkeitsverteilungen der Merkmale auf der Basis von 225 Aufgabeneinschätzungen. Bei den Merkmalen Anforderungsbereich, Prozessbereich, Lernzielstufe und Wissensart waren Mehrfachantworten möglich.

6.2.1 Auswertungsmodalitäten

Die aufgabenweise Sichtung der 225 (45 x 5) Aufgabenprofile ergibt, dass je Aufgabe zwischen drei und siebzehn Einschätzungsbögen ausgefüllt wurden.¹¹ So breit ist die Spanne aus zwei Gründen. Erstens lagen Aufgaben, die im Wettbewerb in mehr als einer Jahrgangsstufe gestellt wurden, entsprechend auch mehr Experten zur Einschätzung vor und zweitens war der Rücklauf nicht vollständig. Von den 54 versendeten Fragebögen wurden nur 45 beantwortet.

Zunächst werden die Bewertungsverteilungen je Merkmal betrachtet, also die Spaltensummen der Häufigkeitstabellen (S. 210ff), in Abbildung 6.3 als Balkendiagramme dargestellt. Die Verteilung der Erfahrungsweltnähe ist steigend, die meisten Einschätzungen lauten entfernt von der Erfahrungswelt. Die Verteilungen der Merkmale Abstraktionsgrad, Komplexitätsgrad und Formalisierungsgrad sind fallend, besonders abstrakt und komplex werden selten genannt. Auch redundant wird selten genannt. Beim Anforderungsbereich fällt auf, dass der größte Anteil auf Anwendung entfällt. Die Prozessbereiche werden annähernd gleich häufig genannt, nur Kommunizieren und Kooperieren wird im Quiz-Format der Biber-

¹¹Die Häufigkeitstabellen der Merkmale sind in Anhang C abgedruckt (S. 210ff).

Aufgaben kaum angesprochen. Die Einschätzungen der Lernzielstufe sind glockenförmig um den Schwerpunkt Anwenden verteilt. Besonders Erschaffen kommt im Multiple-Choice-Format selten vor, wie auch die Wissensart Metakognition.

Anzumerken ist, dass die Verteilungen verzerrt sind. Das kommt dadurch zustande, dass die Anzahl der Expertenwertungen je Aufgabe schwankt. Beispielsweise wurden die Aufgaben **Verwandlung** mit vier von fünf Stimmen und **Zahlenreihe** mit acht von zehn Stimmen beide mehrheitlich der Wissensart Prozeduren zugeordnet. Mit acht Nennungen trägt die Aufgabe **Zahlenreihe** doppelt so stark zur Häufigkeit des Merkmals Prozeduren bei wie die Aufgabe **Verwandlung** mit vier Nennungen. Geht eine Aufgabe stärker in die Verteilung ein, weil sie in mehreren Jahrgangsstufen gestellt wurde, ist das eher zu rechtfertigen, als wenn der Grund für die Unregelmäßigkeit im fehlenden Rücklauf einiger Fragebögen liegt. In jedem Fall kann aus den Häufigkeitsverteilungen eine Tendenz abgelesen werden, die zum Verständnis der Merkmale beiträgt.

Um aus der Summe der einzelnen Bewertungen (Experte/Merkmal/Aufgabe) eine schlüssige Klassifizierung (Merkmal/Aufgabe) zu gewinnen, werden die Häufigkeitsverteilungen der Bewertungen je Aufgabe betrachtet, also die Zeilen der Häufigkeitstabellen (S. 210ff). Für die Merkmale Erfahrungsweltnähe, Abstraktionsgrad, Formalisierungsgrad, Komplexitätsgrad und Redundanz (kurz: Einfachmerkmale) gelten andere Auswertungsmodalitäten als für die Merkmale Anforderungsbereich, Prozessbereich, Lernzielstufe und Wissensart (kurz: Mehrfachmerkmale), bei denen Mehrfachantworten möglich waren, etwa zwei Prozessbereiche angekreuzt werden konnten.

Bezüglich der Einfachmerkmale, Erfahrungsweltnähe bis Redundanz, wird der Modalwert der Bewertungsverteilung als Aufgabenklassifizierung übernommen, darunter versteht man den meistgenannten Wert. Gemäß dem Bestreben, als Eingabe für die vorgesehenen Analysen möglichst immer eine Klassifizierung zu erhalten, wird bei benachbarten Modi die stärkere Ausprägung gewählt, das heißt abstrakt, falls von zehn Befragten keiner den Abstraktionsgrad konkret, fünf den Wert mittel und fünf abstrakt nannten (a). In allen anderen Fällen wird keine Klassifizierung ausgewiesen. Bezüglich der Mehrfachmerkmale, Anforderungsbereich bis Wissensart, ergibt sich die Klassifizierung einer Aufgabe aus den übereinstimmenden Nennungen von mindestens 50 % der Befragten, das heißt Konzepte und Prozeduren, falls von zehn Befragten zwei die Wissensart Fakten, acht Konzepte, fünf Prozeduren und einer Metakognition nannten (b).

Gleichzeitig erlaubt die prozentuale Nennungshäufigkeit der Modi eine Beurteilung, inwieweit die Aufgaben sich der Einordnung nach Kriterien erschließen und eine übereinstimmende Aufgabenbewertung möglich ist. Das obige Beispiel (a) weist 50 % Übereinstimmung auf, denn fünf von zehn Experten nannten den Abstraktionsgrad abstrakt. Beispiel (b) zeigt sogar 80 % Übereinstimmung, denn die häufigste Bewertung, die Wissensart Konzepte, wurde von acht der zehn Experten genannt. Die Güte der Klassifizierung eines Merkmals wird an der durchschnittlichen Übereinstimmung über allen Aufgaben gemessen. Tabelle 6.3 listet

Tabelle 6.3: Güte der Merkmalsklassifizierung. Arithmetischer Mittelwert und Standardabweichung der Übereinstimmung im Expertenurteil.

Merkmal	Übereinstimmung (\emptyset)	Standardabweichung
Erfahrungsweltnähe	73,7%	13,4%
Abstraktionsgrad	68,4%	20,1%
Formalisierungsgrad	69,1%	16,9%
Komplexitätsgrad	72,3%	19,3%
Redundanz	85,8%	18,4%
Anforderungsbereich	81,8%	14,2%
Prozessbereich	68,0%	15,8%
Lernzielstufe	77,6%	17,4%
Art des Wissens	69,6%	16,7%

die arithmetischen Mittelwerte und die zugehörigen Standardabweichungen auf, die bemessen, wie stark die Beobachter-Übereinstimmung über den Aufgaben schwankt. Ein verbreitetes statistisches Zuverlässigkeitsmaß für die Beobachter-Übereinstimmung (Intercoder-Reliability) ist beispielsweise Krippendorffs Alpha (Krippendorff 2011), ein Koeffizient, der sich durch eine Zufallskorrektur auszeichnet. Krippendorffs Alpha ist jedoch ein rechenaufwändiger Koeffizient, da er je Merkmal auf der Koinzidenzmatrix der paarweisen Übereinstimmungen und Abweichungen aller Einzelbewertungen (Experte/Aufgabe) beruht. Hier geht es darum, die wahrscheinlichste Aufgabenbewertung zu bestimmen und Indizien für die Anwendbarkeit der Merkmale zu finden. Ein Nachweis der signifikanten Übereinstimmung wird nicht geführt. Diesem gemäßigten Anspruch genügt die oben beschriebene einfache Kennzahl der mittleren Übereinstimmung, die die prozentuale Mehrheit bemisst, auf der die Klassifizierung beruht.

Ein möglicher Grund für Beobachter-Abweichungen liegt in der Definition der Merkmale, die überwiegend nicht abstrakt, sondern anhand von Beispielen erfolgt. Das zeigen die teilweise ausführlichen Bemerkungen der Experten zur Einschätzbarkeit der Aufgaben, die zum Beispiel die Kategorien des Formalisierungsgrades und des Komplexitätsgrades hinterfragten. Solche Schwächen in der Merkmalspezifikation führten möglicherweise zu unscharfen Einordnungen und beeinträchtigten die Handhabung der Kriterien.

Im ersten Durchgang wird die Auswertung der Experten-Einschätzungen mit dem Ziel vorgenommen, möglichst jede Aufgabe bezüglich jedes Merkmals einer Klassifizierung zuzuführen. Falls sich so in der anschließenden Analyse keine aussagekräftigen Ergebnisse abzeichnen, kann die Mindestübereinstimmung angepasst werden, um eine strengere Klassifizierung zu erzielen. Die Tabellen 6.4 und 6.5 enthalten die Ergebnisse der Aufgabenklassifizierung.

Tabelle 6.4: Aufgabenmerkmale, Teil 1. Einschätzung der Aufgaben nach den Merkmalen Erfahrungsweltnähe (EG), Abstraktionsgrad (AG), Formalisierungsgrad (FG), Komplexitätsgrad (KG) sowie Redundanz (RG).

Aufgabe	EG	AG	FG	KG	RG
1 Biber am Fluss	mittel	mittel	inform.	einfach	nicht red.
2 Biber und Bisons	entfernt	mittel	inform.	mittel	nicht red.
3 Bibers Geheimcode	mittel	konkret	inform.	mittel	nicht red.
4 Biberzahlen	entfernt	mittel	mittel	mittel	-
5 Binärbaum	entfernt	konkret	formal	mittel	nicht red.
6 Computervirus	nah	konkret	inform.	einfach	nicht red.
7 Dateisuche	mittel	-	-	mittel	nicht red.
8 Dino-Ordnung	mittel	konkret	inform.	mittel	nicht red.
9 Endlosschleife	entfernt	mittel	mittel	mittel	nicht red.
10 Falschgeld	mittel	konkret	mittel	mittel	nicht red.
11 Fenster schließen	nah	konkret	inform.	einfach	nicht red.
12 Labyrinth	mittel	-	inform.	mittel	redundant
13 Link	nah	konkret	inform.	einfach	nicht red.
14 Links um!	mittel	konkret	inform.	einfach	nicht red.
15 Morse-Code	entfernt	konkret	mittel	mittel	nicht red.
16 Netzwerkkabel	entfernt	mittel	mittel	mittel	nicht red.
17 Platzwechsel	entfernt	mittel	mittel	mittel	-
18 POP und PUSH	mittel	konkret	inform.	mittel	nicht red.
19 Primärschlüssel	entfernt	konkret	mittel	einfach	nicht red.
20 Private E-Mail	nah	konkret	inform.	einfach	-
21 Schnitzeljagd	entfernt	konkret	formal	mittel	nicht red.
22 Sicheres Passwort	nah	konkret	inform.	einfach	nicht red.
23 Umparken	mittel	konkret	mittel	einfach	nicht red.
24 Ungeschützter PC	nah	konkret	inform.	einfach	nicht red.
25 Verschlüsselung	entfernt	mittel	mittel	mittel	nicht red.
26 Verwandlung	entfernt	mittel	inform.	mittel	nicht red.
27 Wertetausch	mittel	abstrakt	formal	mittel	nicht red.
28 Wetter	nah	konkret	inform.	einfach	nicht red.
29 Zahlenreihe	entfernt	konkret	inform.	einfach	nicht red.

Tabelle 6.5: Aufgabenmerkmale, Teil 2. Einschätzung der Aufgaben nach den Merkmalen Anforderungsbereich, Prozessbereich, Kognitive Lernzielstufe sowie Art des Wissens.

Aufgabe	Anforderg.	Prozess	Lernziel	Wissen
1 Biber am Fluss	Anwendung	Darst./Int.	Analysieren	K., P.
2 Biber und Bisons	Problemlsg.	-	Verst., Ana.	Konz.
3 Bibers Geheimcode	Anwendung	Strukt./Vern.	Anwenden	-
4 Biberzahlen	Anwendung	Darst./Int.	Anwenden	Proz.
5 Binärbaum	Anwendung	Darst./Int.	Anw., Ana.	Konz.
6 Computervirus	Wiedergabe	Begr./Bewert.	Verstehen	Fakt.
7 Dateisuche	Anwendung	Strukt./Vern.	Verst., Anw.	Proz.
8 Dino-Ordnung	Anwendung	B./B., S./V.	Verst., Ana.	Konz.
9 Endlosschleife	Anwendung	M./I., D./I.	Anw., Ana.	K., P.
10 Falschgeld	Anwendung	Darst./Int.	Anwenden	Proz.
11 Fenster schließen	Wiedergabe	Darst./Int.	Erinnern	Fakt.
12 Labyrinth	Anw., Probl.	S./V., D./I.	Anw., Ana.	K., P.
13 Link	Wiedergabe	Komm./Koop.	Erinnern	Fakt.
14 Links um!	Anwendung	Modell./Impl.	Anwenden	-
15 Morse-Code	Anw., Probl.	Strukt./Vern.	Analysieren	Proz.
16 Netzwerkkabel	Problemlsg.	S./V., D./I.	Analysieren	Konz.
17 Platzwechsel	Problemlsg.	Modell./Impl.	Anw., Ana.	Proz.
18 POP und PUSH	Anwendung	M./I., D./I.	Anwenden	K., P.
19 Primärschlüssel	Anwendung	Begr./Bewert.	Verstehen	F., K.
20 Private E-Mail	Wiedergabe	B./B., K./K.	Erinnern	Fakt.
21 Schnitzeljagd	Anwendung	-	Anw., Ana.	Proz.
22 Sicheres Passwort	Anwendung	Begr./Bewert.	Bewerten	Fakt.
23 Umparken	Anwendung	M./I., D./I.	Anwenden	K., P.
24 Ungeschützter PC	Wiedergabe	Begr./Bewert.	Verstehen	Fakt.
25 Verschlüsselung	Anwendung	M/I, B/B, D/I	Anw., Ana.	Konz.
26 Verwandlung	Problemlsg.	B./B., D./I.	Analysieren	Proz.
27 Wertetausch	Anwendung	-	Anwenden	K., P.
28 Wetter	Anwendung	B./B., S./V.	Verst., Anw.	Konz.
29 Zahlenreihe	Anwendung	Modell./Impl.	Anwenden	Proz.

6.2.2 Aufgabenklassifizierung

Tabelle 6.4 enthält die Klassifizierung nach Einfachmerkmalen, Tabelle 6.5 nach Mehrfachmerkmalen. Solcherart bewertet, dienen die Wettbewerbsaufgaben in den Dimensionsanalysen als Testaufgaben für die Merkmale. **Wetter** (S. 195) etwa repräsentiert die Attribute nah zur Erfahrungswelt, konkret, informell, einfach, was die Lösungsraumkomplexität betrifft, und nicht redundant. Die Aufgabe beinhaltet die Anforderung Anwendung, involviert die Prozesse Begründen und Bewerten, Strukturieren und Vernetzen, spricht die Lernzielstufen Verstehen und Anwenden an und beruht auf Konzeptwissen. Folgend wird anhand der Experten-Übereinstimmung und anhand der Rückmeldungen, die die Experten im Verlauf der Befragung abgaben, für jedes Kriterium die Anwendbarkeit reflektiert.

Erfahrungsweltnähe

Allein nach dem Grad der Übereinstimmung beurteilt, den die Befragten erreichten, scheint das Kriterium Erfahrungsweltnähe mit einer durchschnittlichen Übereinstimmung von 73,7% in der meistgenannten Kategorie und einer Standardabweichung von 13,4% gut handhabbar zu sein. Im Mittel gelangten drei von vier Experten zur gleichen Bewertung. Allerdings erfüllt das Merkmal nur bedingt den Anspruch, ausschließlich aufgabenbezogen einschätzbar zu sein, zumindest in der Sekundarstufe, denn die Erfahrungswelt von Zehnjährigen und Sechzehnjährigen oder von Gymnasiasten und Berufsschülern deckt sich nicht (vgl. S. 69). Bemerkenswert ist auch, dass die Erfahrungsweltnähe von den männlichen Befragten anders beurteilt wurde als von den weiblichen. Die Häufigkeitsverteilungen der Bewertungen nach Geschlecht in Abbildung 6.4 lassen erkennen, dass von den drei Ausprägungen nah, mittel, entfernt die dritte, entfernt, durch die männlichen Experten am häufigsten und durch die weiblichen Experten am seltensten genannt wurde. In der Summe erscheint das Merkmal Erfahrungswelt weder aufgabeninhärent noch genügend objektiv. Es wird aus der Ausgangsmenge der Merkmale für die Strukturanalysen wieder entfernt.

Abstraktionsgrad

Der Abstraktionsgrad fand im Merkmalsvergleich mit 68,4% eine geringere, ungefähr eine Zweidrittel-Übereinstimmung, bei einer relativ großen Streuung von 20,1%. Zwar wählten die Experten bei den Aufgaben **Zahlenreihe** (S. 196), **Private E-Mail** (S. 189), **Link** (S. 184) und **Sicheres Passwort** (S. 191) zu hundert Prozent übereinstimmend den Abstraktionsgrad konkret, aber bei **Biberszahlen** (S. 176), **Wertetausch** (S. 195) und **Dateisuche** (S. 178) verteilten sich die Einschätzungen 4:4:2, 1:2:2 oder gar 2:1:2 auf die Ausprägungen konkret, mittel und abstrakt. Ein Erklärungsansatz ist, dass bei diesem Kriterium der Kontext und nicht wie beabsichtigt die Arbeitsanweisung beurteilt wurde. Besonders die Bewertung von **Dateisuche** scheint mit 2:1:2 kontrovers oder aber zufällig. Die Aufgabe wird

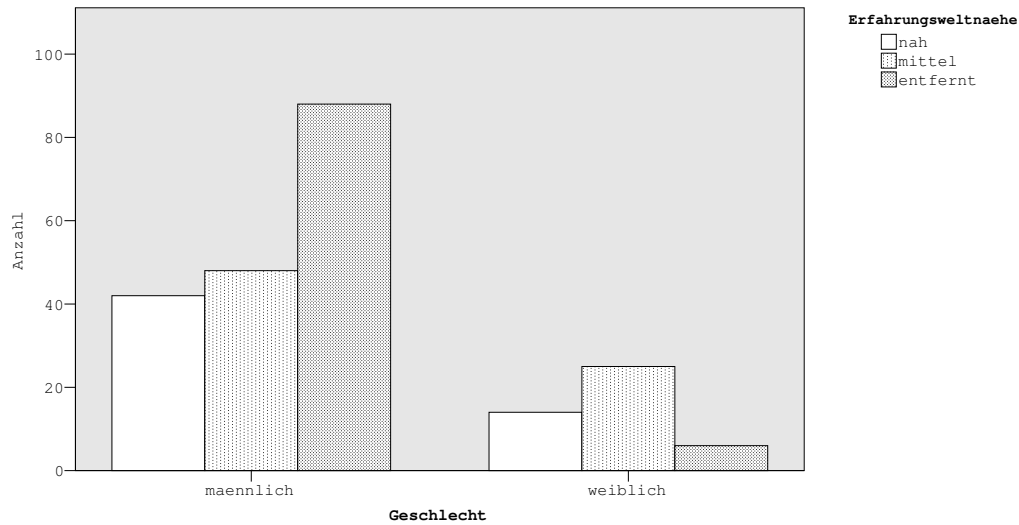


Abbildung 6.4: Häufigkeitsverteilung des Merkmals Erfahrungsweltnähe nach Geschlecht der Experten (auf der Basis von 225 Aufgabeneinschätzungen)

bezüglich ihres Abstraktionsgrades nicht klassifiziert. **Biberzahlen** wird mit 4 : 4 : 2 als konkret und **Wertetausch** mit 1 : 2 : 2 als abstrakt klassifiziert, getreu der Linie, auch bei schwacher Konformität ein Urteil zu treffen.

Formalisierungsgrad

Die Einschätzung des Formalisierungsgrades erfolgte mit einer mittleren Übereinstimmung von 72,3%. Nur eine Aufgabe, **Dateisuche** (S. 178), findet keine eindeutige Bewertung. Die relativ starke Streuung von 19,3% weist aber auf eventuelle Schwächen in der Spezifikation der Ausprägungen informell, mittel und formal hin. Die Anmerkungen der Experten bestätigen, dass die Kategorien umfassend und trennscharf sind, jedoch wurde ihre Anordnung hinterfragt. Die symbolische Darstellung, beispielsweise eines Automaten als Tupel, weist gegebenenfalls einen höheren Formalisierungsgrad auf als die modellhafte Repräsentation, etwa als Graph (vgl. S. 72). Als Reaktion auf diese Bedenken werden mittel und formal, gekennzeichnet durch symbolische und modellhafte Repräsentation, zum Formalisierungsgrad formal vereinigt.

Komplexitätsgrad

Das Merkmal Komplexitätsgrad weist mit 69,1% eine ähnlich geringe Übereinstimmung auf wie der Abstraktionsgrad, es können aber alle Aufgaben eindeutig klassifiziert werden. Während der Einschätzung erwies sich als schwierig, dass in der Definition als Lösungsraumkomplexität keine expliziten Grenzen für die „An-

zahl der Fälle, Regeln, Vergleiche, Umformungen, die die Größe des Lösungsraums bestimmt“ (S. 201, Hinweise zur Einordnung der Aufgaben) vorgegeben waren. Einige Experten berichteten, dass sie bei ihrer Einschätzung unsicher waren, ob die Komplexität mittel oder hoch anzusetzen sei. Im Zuge der Datenaufbereitung werden die Kategorien zusammengezogen zu einfach und komplex, letztere vereint mittlere und komplexe Aufgaben. Auch gaben einige Beobachter zu bedenken, obwohl die Problemkomplexität bereits eng als Lösungsraumkomplexität gefasst sei, überlappe sie mit anderen Kriterien. Möglicherweise beanspruche ein größerer Lösungsraum ein höheres Konzentrations- oder Durchhaltevermögen oder mache strategisches Denken erst nötig. Einmal mehr weist das auf den zentralen Gegenstand dieser Untersuchung hin, die Merkmalsverflechtung.

Redundanz

Die Bewertungsverteilung für das Merkmal Redundanz zeigt, dass die Einschätzung „nicht redundant“ deutlich überwiegt (Abb. 6.3, S. 87). Unter den gegebenen Auswertungsmodalitäten würden mit 85,8% durchschnittlicher Übereinstimmung alle Aufgaben als nicht redundant klassifiziert. Für Biberzahlen (S. 176), Private E-Mail (S. 189) und Platzwechsel (S. 186) bleibt die Frage der Redundanz unentschieden, da es bei gegenteiligen Werten, redundant und nicht redundant, unangemessen ist, im Fall benachbarter Modi die stärkere Ausprägung zu wählen. Eindeutig als redundant wird nur die Aufgabe Labyrinth (S. 183) klassifiziert. Der Aufgabensatz ist kaum geeignet, das Kriterium Redundanz zu untersuchen. Es wird aus der Ausgangsmenge der potentiellen Schwierigkeitsmerkmale gestrichen.

Anforderungsbereich

Der Anforderungsbereich wurde von den Befragten mit großer Übereinstimmung, 81,8%, bei geringer Standardabweichung, 14,2%, zugeordnet. Da die Wettbewerbsteilnahme keine Vorkenntnisse voraussetzt, kann das Kriterium aufgabenbezogen, das heißt unabhängig vom Adressaten, beurteilt werden. Für jede Aufgabe werden in der Vorgabe alle benötigten Informationen verfügbar gemacht, die mit Sachverstand „wiederzugeben“ oder „anzuwenden“ sind oder auf deren Basis das „Problem zu lösen“ ist. In Abbildung 6.5 ist zu sehen, wie sich die Anforderungsbereiche auf die Lernzielstufen verteilen. Zwischen den Merkmalen besteht ein natürlicher Zusammenhang, der darauf zurückzuführen ist, dass beiden die Bloomsche Taxonomie zugrunde liegt (siehe unten, Lernzielstufe).

Prozessbereich

Eine geringere Übereinstimmung zeigt mit 68% der Prozessbereich. Elf der Befragten vermerkten, vergleichsweise sei die Einschätzung des Merkmals mit der größten Unsicherheit erfolgt. Drei Experten haben regelmäßig rein nach dem Ausschlussprinzip Modellieren und Implementieren gewählt. Nichtsdestoweniger zeigt

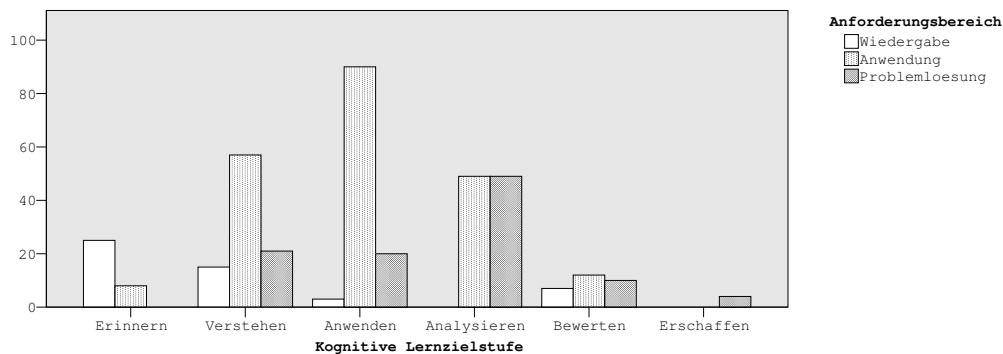


Abbildung 6.5: Häufigkeitsverteilung der Anforderungsbereiche auf die Lernzielstufen (auf der Basis von 225 Aufgabeneinschätzungen)

die Werteverteilung, dass bis auf Kommunizieren und Kooperieren die Prozessbereiche gleichermaßen in Anspruch genommen wurden. Ob die resultierende Klassifizierung geeignet ist, einen Einfluss der Prozessbereiche auf die Aufgabenschwierigkeit zu belegen, wird in einem Exkurs zur eindimensionalen Modellierung der Informatikkompetenz untersucht (siehe Abschnitt 8.2.2, S. 118).

Lernzielstufe

Die Einordnung der Aufgaben in die Lernzieltaxonomie erfolgte mit 77,6 % Übereinstimmung. Da den kognitiven Lernzielstufen wie den Anforderungsbereichen die Taxonomie von Bloom zugrunde liegt, gibt es Überschneidungen. Die gruppierte Häufigkeitsverteilung (Abb. 6.5) lässt erkennen, dass die Lernzielstufe Anwenden häufig mit dem Anforderungsbereich Anwendung zusammenfiel, was nicht überrascht, denn im Kern gleichen sich die Spezifikationen. Die Kategorien können jedoch nicht zur Deckung gebracht werden, da z. B. die Lernzielstufe Analysieren mit der Anforderung Anwendung genau so häufig zusammenfiel wie mit der Anforderung Problemlösung. Das Zusammenspiel der beiden Merkmale wird im Weiteren erforscht.

Art des Wissens

Die Wissensart wurde mit mäßiger Übereinstimmung von 69,6 % bewertet. Die Aufgaben Bibers Geheimcode (S. 175) und Links um! (S. 184) können nicht eingeordnet werden. Im Gesamtergebnis wird der Wissensart Metakognition keine Aufgabe zugesprochen, die Kategorie bleibt leer. Die höchste Bewertung erreichte die Aufgabe Biber und Bisons (S. 175). Sie wurde von neun Befragten mit 2 : 6 : 4 : 3 Nennungen den Wissensarten Fakten, Konzepte, Prozeduren und Metakognition zugeordnet.

Zusammenfassung

Aus den Erhebungen stehen mehrere Tabellen als Ausgangsdaten für die Erforschung der Anforderungsstruktur der Aufgaben und der Fähigkeitsstruktur der Personen bereit: die Teilnahmedaten im Format Teilnehmer x Aufgabe (Tabelle 6.1) und die Wettbewerbsaufgaben im Format Aufgabe x Merkmal (Tabellen 6.2, 6.4, 6.5). Das Kriterienschema erwies sich als gut handhabbar und wurde von den Befragten akzeptiert. Die Aufgaben erschlossen sich der Bewertung und konnten bis auf wenige Ausnahmen bezüglich der Kriterien eingeordnet werden. Um das zu erreichen, sind die Auswertungsmodalitäten nicht allzu streng gefasst.

Die Lernzielstufe Erschaffen und die Wissensart Metakognition sind überhaupt nicht mit Aufgaben belegt. Das Kriterium Erfahrungsweltnähe wurde ausgeschlossen, weil die Erfahrungswelt kein Aufgabenmerkmal, sondern ein Merkmal des individuellen Schülers ist (vgl. S. 67f, 1. Bedingung). Auch das Kriterium Redundanz wird nicht weiter untersucht, da es in diesem Aufgabensatz nicht variiert (vgl. S. 67f, 3. Bedingung).

Besonders bei der Bewertung des Anforderungsbereichs und der Lernzielstufe fielen Verflechtungen der Merkmale auf. Die folgenden Analysen haben zum Ziel, die Verflechtungen sichtbar zu machen und die enthaltene Struktur herauszuarbeiten. Der Entstehungsprozess des gesuchten Modells ist iterativ. Merkmale werden identifiziert, empirisch analysiert, Modellhypothesen entwickelt. In weiteren Iterationen wird die Merkmalsmenge revidiert, gegebenenfalls ergänzt, die Analyseverfahren werden angepasst, die Modellhypothesen korrigiert oder verfeinert, und so weiter. Diese Arbeit beschreibt eine erste Iteration der Modellentwicklung, mit dem Ziel, eine solide Ausgangsbasis für weitere Iterationen zu schaffen. Vermutlich können die Überschneidungen der Kriterien in einer neuen Iteration, nach einer gründlichen Untersuchung der Merkmalszusammenhänge, reduziert werden.

Kapitel 7

Daten

Im letzten Kapitel wurde der Schülerwettbewerb Informatik-Biber vorgestellt, dessen Ausrichtung im Jahr 2007 die Daten für die Studie lieferte. Aus dem Wettbewerb sind die Aufgaben und Teilnehmerantworten als Rohdaten verfügbar. Aus der Expertenbefragung liegt zudem die Merkmalsklassifizierung der Wettbewerbsaufgaben vor. Für die geplanten Analysen müssen die Daten geeignet kodiert werden. Die Aufbereitung der Daten aus den Erhebungen wird als dritter Schritt im Forschungsablauf in diesem Kapitel beschrieben.

In Abschnitt 7.1 wird die Kodierung der klassifizierten Aufgaben beschrieben, als Eingabe einer Clusteranalyse der Aufgaben nach den Merkmalen. Die Aufgabenklassifizierung wird in Tabellen der Form Aufgabe x Merkmal verfügbar gemacht. In Abschnitt 7.2 wird die Kodierung der Teilnahmedaten als Antwortmuster dargestellt, für eine Faktorenanalyse der Aufgaben nach den Antwortmustern. Die Antwortmuster werden in Tabellen der Form Teilnehmer x Aufgabe verfügbar gemacht. Zuletzt werden die Antwortmuster und die Aufgabenklassifizierung zu Personen-Merkmalsquoten aggregiert (Abschnitt 7.3). Die Merkmalsquoten, in Tabellen der Form Teilnehmer x Merkmal, werden später ebenfalls faktorenanalytisch untersucht. In dem Zusammenhang wird die Problematik der konfundierten Variablen thematisiert, die dem Datensatz innewohnt.

7.1 Aufgabenklassifizierung

Die Wettbewerbsaufgaben wurden von den Experten nach den Merkmalen Erfahrungsweltnähe, Abstraktionsgrad, Formalisierungsgrad, Komplexitätsgrad, Redundanz klassifiziert. Bis auf das letztgenannte Kriterium Redundanz mit zwei Ausprägungen sind die Merkmale dreistufig und jede Aufgabe wurde höchstens einer Stufe zugeordnet. Die Aufgabe Morse-Code (S. 185), in der die effizienteste Codierung bei gegebener Buchstabenwahrscheinlichkeit gesucht ist, wird so dem Abstraktionsgrad konkret zugeordnet. Weiter wurden die Aufgaben nach den Merkmalen Anforderungsbereich, Prozessbereich, Kognitive Lernzielstufe,

Art des Wissens klassifiziert. Bei diesen Merkmalen waren Mehrfachantworten möglich, **Morse-Code** etwa spricht zwei Anforderungsbereiche an, nämlich Anwendung und Problemlösung (siehe Tabellen 6.4 und 6.5, S. 90f).

Die Kriterien Erfahrungsweltnähe und Redundanz wurden von den Analysen ausgeschlossen, das erste hauptsächlich, weil die Erfahrungswelt vom Alter des Schülers abhängt (vgl. S. 92), das zweite, weil im Aufgabensatz konstant keine Redundanz erkennbar ist (vgl. S. 94). Einige Merkmalskategorien sind nur dünn mit Aufgaben besetzt: Der Prozessbereich Kommunizieren und Kooperieren enthält nur die Aufgaben **Link** (S. 184) und **Private E-Mail** (S. 189). Diese und **Fenster schließen** (S. 182) sind auch die einzigen Aufgaben, die der Lernzielstufe **Erinnern** zugeordnet werden. Die Lernzielstufe **Bewerten** wird lediglich durch die Aufgabe **Sicheres Passwort** (S. 191) vertreten. Die Lernzielstufe **Erschaffen** kommt in diesem Aufgabensatz überhaupt nicht vor, ebenso die Wissensart **Metakognition**. So können der Prozessbereich Kommunizieren und Kooperieren, die Lernzielstufen **Erinnern**, **Bewerten**, **Erschaffen** und die Wissensart **Metakognition** innerhalb des gegebenen Datenmaterials nicht analysiert werden.

Zur Kodierung bleiben die Merkmale **Abstraktionsgrad**, **Formalisierungsgrad**, **Komplexitätsgrad** übrig, die die Gestaltung der Aufgabe charakterisieren (kurz: **Gestaltungsmerkmale**), dazu die Anforderungsbereiche **Wiedergabe**, **Anwendung**, **Problemlösung**, die Prozessbereiche **Modellieren** und **Implementieren**, **Begründen** und **Bewerten**, **Strukturieren** und **Vernetzen**, **Darstellen** und **Interpretieren**, die Lernzielstufen **Verstehen**, **Anwenden**, **Analysieren** sowie die Wissensarten **Fakten**, **Konzepte**, **Prozeduren**. Das Ziel ist nun, eine einheitliche Kodierung zu finden.

Bei den Gestaltungsmerkmalen wurde für den **Formalisierungsgrad** bereits vorgeschlagen, *informell* beizubehalten und **formal** und **mittelformal** zu *formal* zu vereinigen, mit der Bedeutung „erfordert den Umgang mit Formalismen“ (vgl. S. 93). Der **Abstraktionsgrad** wird nur für die Aufgabe **Wertetausch** (S. 195) mehrheitlich als **abstrakt** bewertet, sonst als **konkret** oder **mittel**. Auch wird keine der Aufgaben mehrheitlich als **komplex** klassifiziert (Tabelle 6.4, S. 90). Analog zum **Formalisierungsgrad** werden die **Grade** **abstrakt** und **mittelabstrakt** vereinfachend zu *abstrakt* mit der Bedeutung „erfordert den Umgang mit Abstraktion“ zusammengefasst. Die **Grade** **komplex** und **mittelkomplex** werden zu *komplex* mit der Bedeutung „erfordert den Umgang mit Komplexität“ vereinfacht. Damit sind die **Gestaltungsmerkmale** **dichotom**, das heißt **zweigeteilt**. Sie werden jeweils mit den Werten 1 für „trifft zu“ und 0 für „trifft nicht zu“ kodiert (Tabelle 7.1).

Die Kriterien mit **Mehrfachnennungen** der Kategorien werden bislang mit **multiplen Werten** aufgeführt (Tabelle 6.5, S. 91). Bei einer **multiplen Kodierung** werden neue Kategorien hinzugenommen, wenn einer Aufgabe zwei oder mehr Ausprägungen eines Merkmals zugeschrieben werden. Hier würde den bestehenden Wissenskategorien **Fakten**, **Konzepte** und **Prozeduren** eine vierte Kategorie **Fakten + Konzepte** für die Aufgabe **Wertetausch** und eine fünfte Kategorie **Konzepte + Prozeduren** für **Biber am Fluss** und fünf weitere Aufgaben hinzugefügt. Dergestalt kodiert kämen vier Anforderungsbereiche, elf Prozessbereiche, acht

Tabelle 7.1: Aufgabenmerkmale. Gestaltungsmerkmale, Anforderungsbereiche, Prozessbereiche, Lernzielstufen, Wissensarten, dichotom kodiert (1 steht für „trifft zu“, 0 steht für „trifft nicht zu“).

Auf- gabe	Gestaltung			Anforderg.			Prozess				Lernziel			Wissen		
	AG	FG	KG	W	A	P	MI	BB	SV	DI	V	Aw	An	F	K	P
1	1	0	0	0	1	0	0	0	0	1	0	0	1	0	1	1
2	1	0	1	0	0	1	-	-	-	-	1	0	1	0	1	0
3	0	0	1	0	1	0	0	0	1	0	0	1	0	-	-	-
4	1	1	1	0	1	0	0	0	0	1	0	1	0	0	0	1
5	0	1	1	0	1	0	0	0	0	1	0	1	1	0	1	0
6	0	0	0	1	0	0	0	1	0	0	1	0	0	1	0	0
7	-	-	1	0	1	0	0	0	1	0	1	1	0	0	0	1
8	0	0	1	0	1	0	0	1	1	0	1	0	1	0	1	0
9	1	1	1	0	1	0	1	0	0	1	0	1	1	0	1	1
10	0	1	1	0	1	0	0	0	0	1	0	1	0	0	0	1
11	0	0	0	1	0	0	0	0	0	1	0	0	0	1	0	0
12	-	0	1	0	1	1	0	0	1	1	0	1	1	0	1	1
13	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0
14	0	0	0	0	1	0	1	0	0	0	0	1	0	-	-	-
15	0	1	1	0	1	1	0	0	1	0	0	0	1	0	0	1
16	1	1	1	0	0	1	0	0	1	1	0	0	1	0	1	0
17	1	1	1	0	0	1	1	0	0	0	0	1	1	0	0	1
18	0	0	1	0	1	0	1	0	0	1	0	1	0	0	1	1
19	0	1	0	0	1	0	0	1	0	0	1	0	0	1	1	0
20	0	0	0	1	0	0	0	1	0	0	0	0	0	1	0	0
21	0	1	1	0	1	0	-	-	-	-	0	1	1	0	0	1
22	0	0	0	0	1	0	0	1	0	0	0	0	0	1	0	0
23	0	1	0	0	1	0	1	0	0	1	0	1	0	0	1	1
24	0	0	0	1	0	0	0	1	0	0	1	0	0	1	0	0
25	1	1	1	0	1	0	1	1	0	1	0	1	1	0	1	0
26	1	0	1	0	0	1	0	1	0	1	0	0	1	0	0	1
27	1	1	1	0	1	0	-	-	-	-	0	1	0	0	1	1
28	0	0	0	0	1	0	0	1	1	0	1	1	0	0	1	0
29	0	0	0	0	1	0	1	0	0	0	0	1	0	0	0	1

Lernzielstufen und fünf Wissensarten vor. Da dies zu einer unübersichtlichen Zahl teilweise sehr dünn besetzter Kategorien führt, wird anstelle der multiplen Kodierung eine sogenannte Dummy-Kodierung vorgenommen, um die Merkmale mit Mehrfachnennungen den geplanten Analysen zugänglich zu machen. Dabei wird für jede Antwortmöglichkeit eine dichotome Variable mit den Kategorien „trifft zu“ und „trifft nicht zu“ definiert, kodiert mit 1 und 0. Der Anforderungsbereich wird durch die Variablen W (Wiedergabe), A (Anwendung) oder P (Problemlösung) kodiert, der Prozessbereich durch die Variablen MI (Modellieren und Implementieren), BB (Begründen und Bewerten), SV (Strukturieren und Vernetzen) oder DI (Darstellen und Interpretieren), die Lernzielstufen durch V (Verstehen), Aw (Anwenden) oder An (Analysieren), die Wissensarten durch F (Fakten), K (Konzepte) oder P (Prozeduren). Die sechzehn Variablen werden in der neuen Tabelle der Aufgabenmerkmale (Tabelle 7.1, S. 99) zusammengeführt.

7.2 Wettbewerbsergebnisse (Antwortmuster)

Die Teilnahmedaten in Tabelle 6.1 (S. 83) enthalten für jeden Teilnehmer seine Klassenstufe, fünf bis sieben (5-7), acht bis zehn (8-10) oder elf und höher (11+). In jeder Altersgruppe lagen fünfzehn der insgesamt neunundzwanzig Aufgaben vor. Bei den Aufgaben, die in einer Klassenstufe nicht gestellt wurden, lautet das Resultat „lag nicht vor“. Sonst ist das Resultat „korrekt“ bei richtiger Beantwortung, „inkorrekt“ bei falscher Beantwortung und „unbearbeitet“ bei fehlender Beantwortung. Für die Strukturanalysen, die für jede Altersstufe extra erfolgen, werden die Teilnahmedaten in drei Datensätze aufgeteilt.

Die Tabellen 7.2 bis 7.4 enthalten jeweils die Antwortmuster einer Altersgruppe. Jede Teiltabelle enthält nur noch die Aufgaben, die in dieser Stufe gestellt wurden, so dass der Variablenwert „lag nicht vor“ entfällt. Da die Wettbewerbsbedingungen ein Zeitlimit und Punktabzüge für falsche Antworten vorsehen, ist unbekannt, woher die Nichtbeantwortung einer Frage rührt. Eine Erklärung ist, dass der Teilnehmer die Aufgabe tatsächlich nicht lösen konnte. Möglich ist, dass er keine Zeit mehr hatte, die Aufgabe zu bearbeiten, obwohl er sie lösen könnte. Auch möglich ist, dass er seiner Antwort nicht sicher war und keinen Punktabzug riskieren wollte. Daher werden „inkorrekt“ und „unbearbeitet“ zusammengezogen zu „nicht korrekt beantwortet“ und mit 0 kodiert, mit der Bedeutung, dass der Schüler keine korrekte Lösung für die Aufgabe angegeben hat, egal aus welchem Grund. Die Ausprägung „korrekt beantwortet“ wird mit 1 kodiert. Mit den dichotom kodierten Antwortmustern stehen drei neue Datensätze bereit.

Die Wettbewerbsresultate aller Aufgaben werden in den Antwortmustern berücksichtigt, mit Ausnahme der Aufgabe **Ungeschützter PC** (S. 193). Darin sollen die Schüler reflektieren, in welchem Ausmaß Schadprogramme über das Internet verbreitet werden können. Das Anwenderwissen, das vorausgesetzt wird, kann selbst ohne schulischen Informatikunterricht bei Schülern der elften Jahrgang-

stufe erwartet werden. Dennoch scheiterten auffallend viele Schüler dieser Altersgruppe an der Aufgabe, ebenso viele wie in der jüngsten Altersgruppe. Das gab Anlass zu der Vermutung, die Schwierigkeit habe untersuchungsfremde Ursachen, etwa die missverständliche Formulierung der Antwortalternativen. Aus diesem inhaltlichen Grund wird die Aufgabe ausgenommen, um eine Störung zu verhindern.

7.3 Personen-Merkmalsquoten

Die Aufgabenklassifizierung stellt eine Beziehung zwischen Aufgaben und Anforderungsmerkmalen her. Das gestattet, den kognitiven Anspruch der Aufgaben durch die Merkmale zu charakterisieren. Die Antwortmuster stellen den Bezug zwischen Personen und Aufgaben her. Das gestattet, kognitive Fähigkeiten der Personen durch die gelösten Aufgaben zu beschreiben, *indirekt* also auch durch die Merkmale, die die gelösten Aufgaben tragen. Während die Antwortmuster wie die Aufgabenklassifizierung unmittelbar aus dem Wettbewerb und der Expertenbefragung hervorgingen, fehlt noch der *direkte* Bezug zwischen Personen und Schwierigkeitsmerkmalen, der erst durch die Aggregation der Antwortmuster und der Klassifizierung hergestellt wird. Abschnitt 7.3.1 stellt die prinzipielle Vorgehensweise bei der Aggregation dar. In Abschnitt 7.3.2 kommt das schwerwiegende Problem konfundierter Effekte zur Sprache.

7.3.1 Aggregation

Um die Aufgabenmerkmale auf die Personen beziehen zu können, werden die Antwortmuster und die Aufgabenklassifizierung zu Merkmalsmustern der Personen aggregiert. Zuvor wird die Merkmalstabelle (S. 99) in drei Datensätze mit den Aufgabenfällen der Jahrgangsstufen 5-7, 8-10 und 11+ aufgespalten. Da es sich bei den drei neuen Tabellen nur um Auszüge aus der Merkmalstabelle aller Aufgaben handelt, werden sie nicht eigens abgedruckt. Je Teilnehmer wird je Merkmal die Lösungsquote aus der Zahl der korrekt beantworteten und der Zahl der vorliegenden Aufgaben des Merkmals berechnet. Zum Beispiel werden in der Jahrgangsstufe 11+ acht Aufgaben der Lernzielstufe Analysieren gestellt. Beantwortet ein Schüler vier der acht Aufgaben richtig, wird seine Lösungsquote für diese Lernzielstufe mit 50% beziffert. Rechnerisch wird die Zahl der korrekt beantworteten Aufgaben je Teilnehmer und je Merkmal einfach durch eine Matrixmultiplikation der Antwortmustertabellen (Teilnehmer x Aufgabe) mit den Merkmalstabellen (Aufgabe x Merkmal) gewonnen. Die Zahl der vorliegenden Aufgaben je Merkmal ergibt sich aus den Spaltensummen der Merkmalstabellen. Das Ergebnis der Aggregation sind die Personen-Merkmalsquoten in den Tabellen 7.5 bis 7.7.

7.3.2 Konfundierte Effekte

Unter Konfundierung versteht man die Überlagerung von Effekten. Hier ist problematisch, dass es nur fünfzehn Aufgaben je Teilnehmer gibt und sechzehn Merkmale betrachtet werden. Zwar trägt jede Aufgabe mehrere Merkmale, so dass umgekehrt jedes Merkmal mehrfach repräsentiert wird. Wenn aber Merkmalspaare überwiegend kombiniert vorkommen, überlagert der Effekt ihres gemeinsamen Auftretens in den Aufgabenprofilen den zu untersuchenden Effekt ihres gemeinsamen Auftretens im Fähigkeitsprofil der Teilnehmer. Das heißt, wenn bei den Fähigkeitsmerkmalen Korrelationen auftreten, ist nicht unterscheidbar, ob es sich tatsächlich um eine Korrelation in dem Sinn handelt, dass eine Person, die die eine Fähigkeit besitzt, häufig auch die andere Fähigkeit besitzt, oder ob es sich um eine Korrelation in dem Sinn handelt, dass eine Aufgabe, die die eine Fähigkeit erfordert, häufig auch die andere Fähigkeit erfordert.

Das lässt sich gut am Beispiel der Kriterien Formalisierungsgrad und Komplexitätsgrad verdeutlichen. In der Jahrgangsstufe 8-10 sind sieben Aufgaben beiden Merkmalen zugeordnet, also formal und komplex. Zwei Aufgaben sind formal, aber nicht komplex und zwei Aufgaben sind komplex, aber nicht formal. Die restlichen vier Aufgaben sind weder komplex noch formal. Einem Schüler, der Aufgaben bewältigt, die gleichzeitig formal und komplex sind, kann attestiert werden, dass er beiden Schwierigkeitsansprüchen erfolgreich begegnet. Bei einem Schüler, der sie nicht bewältigt, bleibt jedoch unklar, an welchem Aspekt er bei der Aufgabenbearbeitung scheitert. Nur die Aufgaben, die ausschließlich eines der beiden Merkmale tragen, sind dazu geeignet, die beiden Fähigkeitsaspekte voneinander abzuheben.

Theoretisch wäre es möglich, zu jedem Merkmalspaar die unterscheidenden Aufgaben zu bestimmen. Praktisch ist es jedoch aufwändig und unter der gegebenen Fragestellung nicht zielführend, geht es doch gerade um das Zusammenspiel aller Merkmale. So bleibt es dabei, dass die Konfundierung der Merkmale auf der Aufgaben- bzw. Anforderungsebene die geplanten Analysen der Merkmale auf der Personen- bzw. Fähigkeitsebene beeinflusst und verfälscht. Das heißt, der gewünschte Übergang von Aufgaben-Merkmalen zu Personen-Merkmalen durch die beschriebene Aggregation ist mit diesem Aufgabensatz problematisch.

Zusammenfassung

Eingangs stehen die Rohdaten, die durch geeignete Kodierung in Datentabellen für die dimensionalen Analysen überführt werden. Die Datentabellen enthalten die Aufgabenmerkmale, die Antwortmuster und die Merkmalsquoten, aufgeteilt in je drei Teiltabellen für die Klassenstufen 5-7, 8-10 und 11+. Die Merkmals-tabellen enthalten die Aufgaben als Fälle und die Merkmale als Variablen. Die Antwortmuster-Tabellen enthalten die Teilnehmer als Fälle und die Aufgaben als Variablen. Die Tabellen der Personen-Merkmalsquoten werden aus den Aufga-

benmerkmalen und den Antwortmustern berechnet. Sie enthalten die Personen als Fälle und die Merkmale als Variablen. Weil eine Konfundierung der Variablen anzunehmen ist, werden die Merkmalsquoten nur unter starkem Vorbehalt weiter verwendet.

Kapitel 8

Analysen: Von den Merkmalen zum Modell

Nach der Datenaufbereitung stehen die Ergebnisse der Aufgabenklassifizierung, die Antwortmuster der Wettbewerbsteilnehmer und die daraus berechneten Personen-Merkmalsquoten in Tabellen bereit. Als nächster Schritt folgt, mithilfe statistischer Analysen dimensionale Strukturen in den Daten aufzuspüren und Erklärungsmodelle des kognitiven Anspruchs der Aufgaben und der kognitiven Leistungen der Schüler zu formulieren.

Zwei unterschiedliche Perspektiven werden eingenommen. Die Perspektive des Anforderungsmodells ist die Sicht auf die kognitiven Anforderungen, die die Aufgaben an die Bearbeiter stellen. Die Perspektive des Fähigkeitsmodells ist die Sicht auf die Problemlösefähigkeit der Bearbeiter, mit der sie den Anforderungen begegnen.

Im Hinblick auf das *Anforderungsmodell* erfolgt eine Clusteranalyse der Aufgaben nach den Merkmalen. Ein qualitatives Modell des Aufgabenanspruchs wird entwickelt, das die kognitiven Anforderungen der Informatikaufgaben in den Begriffen der Schwierigkeitsmerkmale erklärt (Abschnitt 8.1). Davon zu unterscheiden sind die quantitativen Aufgabenkennzahlen „Empirische Schwierigkeit“, als Lösungsquote berechnet, und „Trennschärfe“, als Korrelationskoeffizient berechnet (Abschnitt 8.2).

Im Hinblick auf das *Fähigkeitsmodell* erfolgen Faktorenanalysen der Aufgaben (A) und der Merkmale (B) nach den Bearbeitungsergebnissen. Auf zwei verschiedenen Wegen werden Modelle berechnet, die die Fähigkeit der Schüler, die Informatikaufgaben zu lösen, in den Begriffen der Schwierigkeitsmerkmale erklären. Beim ersten Ansatz werden die Antwortmuster analysiert und die Aufgaben zusammengefasst, die die gleichen Teilnehmern bewältigten. Anhand der Merkmale, durch die sich die gruppierten Aufgaben auszeichnen, werden Dimensionen abgegrenzt, die die Gruppen repräsentieren (Modell A, Abschnitt 8.3). Beim zweiten Ansatz werden die Merkmalsquoten analysiert und die Merkmale nach der Enge ihres Zusammenhangs zu Dimensionen gebündelt (Modell B, Abschnitt 8.4).

Jede der Analysen liefert ein Dimensionenmuster. In der Interpretation werden daraus Strukturhypothesen der Art abgeleitet, informatische Anforderungen wie Fähigkeiten weisen drei Dimensionen auf, die durch die Merkmale Wiedergabe, Anwendung und Problemlösen gekennzeichnet sind (Kapitel 9).

8.1 Anforderungsmodell

Welche Anforderungen die Aufgaben an den Bearbeiter stellen, wird anhand der Aufgabenklassifizierung analysiert. Bisher liegt zu jeder Aufgabe ein Merkmalsprofil vor, das jedoch durch die Vielzahl der Merkmale schwer überschaubar ist. Angestrebt wird ein Modell des Aufgabenanspruchs, das nur wenige, prägnante Dimensionen besitzt, die durch die Merkmale charakterisiert werden.

Die Datenmatrix der Aufgabenmerkmale (Tab. 7.1, S. 99) enthält in den Zeilen neunundzwanzig Aufgaben als Fälle und in den Spalten sechzehn Merkmale als Variablen. Grundsätzlich bieten sich strukturentdeckende Verfahren an, um die dimensionale Struktur der Datenmatrix zu untersuchen. Eines ist die Faktorenanalyse, die Bündel korrelierter *Merkmalsvariablen* durch Faktoren ersetzt, gemäß der Annahme, die korrelierten Variablen seien jeweils auf einen gemeinsamen Faktor als Ursache zurückführbar. Die Faktoren repräsentieren Dimensionen des Aufgabenanspruchs und werden unmittelbar durch die Merkmalsvariablen beschrieben. Ein anderes Verfahren ist die Clusteranalyse, die Gruppen ähnlicher *Aufgabenfälle* bildet, Cluster genannt. Die Cluster werden durch die Merkmalsprofile der gruppierten Aufgaben charakterisiert und können ebenfalls als Dimensionen des Aufgabenanspruchs interpretiert werden. Die Faktorenanalyse ersetzt Variablenspalten mit ähnlichem Muster durch neue Faktorenspalten, die Clusteranalyse gruppiert Aufgabenzeilen mit ähnlichem Merkmalsmuster zu Aufgabentypen.

Zuerst wird eine Faktorenanalyse der Merkmale durchgeführt. Es gelingt, die sechzehn Merkmale auf drei Faktoren zu reduzieren. Die resultierenden Faktoren als Dimensionen zu interpretieren erweist sich aber als schwierig. Das liegt einerseits an der fehlenden Einfachstruktur, die vorliegt, wenn jedes Merkmal nur mit einem Faktor eine hohe Korrelation aufweist. Dann dient das Merkmal der Spezifikation dieses Faktors. Andererseits liegt das an den hohen negativen Faktorladungen. Eine negative Korrelation zwischen Merkmal und Faktor deutet auf einen gegenläufigen Zusammenhang hin. Der Faktor kann allenfalls durch das Nicht-Vorhandensein des Merkmals spezifiziert werden. Der faktorenanalytische Ansatz wird für das Anforderungsmodell daher nicht weiter verfolgt.

Stattdessen werden mithilfe einer Clusteranalyse die Aufgaben nach den Merkmalen gruppiert (Abschnitt 8.1.2). Die Aufgabencluster repräsentieren die Anspruchsdimensionen. Zur Charakterisierung der Dimensionen werden die Merkmalsprofile der Aufgabencluster betrachtet. In zweierlei Hinsicht ist der Datensatz nicht ideal geeignet für eine Clusteranalyse. Erstens ist die Variablenzahl groß,

dadurch wird es erschwert, prägnante Clusterprofile abzugrenzen. Zweitens stehen die Merkmale teilweise in einer Wechselbeziehung, was das Ergebnis verzerren kann. Beides sind keine Ausschlusskriterien für eine explorative Clusteranalyse. Besonders der zweite Punkt, die Wechselbeziehungen der Aufgabenmerkmale, verdient allerdings Beachtung und wird im nächsten Abschnitt beleuchtet.

8.1.1 Wechselbeziehungen der Merkmale

Einen klärenden Einblick in das Zusammenspiel der Merkmale vermitteln die Korrelationskoeffizienten (Tabelle 8.1). Die Merkmalspaare, die eine positive Korrelation von mindestens 0,5 auf dem Signifikanzniveau 0,01 aufweisen, werden aufgeschlüsselt, siehe unten. In Klammern wird der Korrelationskoeffizient angegeben. Zur Anwendung kommt der Phi-Koeffizient, der sehr einfach aus der Vierfelder-tafel zweier dichotomer Variablen berechnet wird (Bortz 2005, S. 227ff).

1. Abstraktionsgrad, Lernzielstufe Analysieren (0,53)
2. Komplexitätsgrad, Lernzielstufe Analysieren (0,56)
3. Anforderungsbereich Wiedergabe, Faktenwissen (0,81)
4. Anforderungsbereich Anwendung, Lernzielstufe Anwenden (0,59)
5. Anforderungsbereich Problemlösung, Lernzielstufe Analysieren (0,61)
6. Prozessbereich Modellieren/Implementieren, Lernziel Anwenden (0,56)
7. Prozessbereich Begründen/Bewerten, Lernzielstufe Verstehen (0,56)
8. Lernzielstufe Anwenden, Prozedurwissen (0,56)

Die Wechselbeziehungen können auf verschiedenen Ebenen erklärt werden. Zum Beispiel kann die Korrelation *inhaltliche* Gründe haben. Eventuell ist ab einem gewissen Komplexitätsgrad analytisches Vorgehen förderlicher als schematisches Vorgehen (2) oder setzt Begründen und Bewerten Verständnis voraus (7). In diesen Fällen liegen die Merkmale tatsächlich in derselben Anspruchsdimension. Die Korrelation kann auch von *extern* eingebracht werden, bewusst oder unbewusst bei der Aufgabenstellung. Möglicherweise hat es sich eingebürgert, Modellieren und Implementieren stets auf der Lernzielstufe Anwenden zu prüfen (6), obwohl Aufgaben aus diesem Prozessbereich genauso gut auf anderen Stufen denkbar sind, wie **Platzwechsel** (S. 186) zeigt, die Modellieren und Implementieren auf der Lernzielstufe Analysieren anspricht. In diesem Fall liegen die Merkmale nicht unbedingt in derselben Dimension. Drittens können die Merkmale *redundant* sein, wie es beim Anforderungsbereich Anwendung und der Lernzielstufe Anwenden zu vermuten ist (4).

Tabelle 8.1: Merkmalskorrelation. Phi-Koeffizient (Φ). Merkmale Abstraktionsgrad (AG), Formalisierungsgrad (FG), Komplexitätsgrad (KG), Anforderungsbereiche Wiedergabe (W), Anwendung (A), Problemlösung (P), Prozessbereiche Modellieren und Implementieren (MI), Begründen und Bewerten (BB), Strukturieren und Vernetzen (SV), Darstellen und Interpretieren (DI), Lernzielstufen Verstehen (V), Anwenden (Aw), Analysieren (An), Wissensarten Fakten (F), Konzepte (K), Prozeduren (P). Positive Korrelationen auf dem Signifikanzniveau 0,01 sind hervorgehoben.

	Gestaltung			Anforderung				Prozessbereich				Lernzielstufe				Wissensart		
	AG	FG	KG	W	A	P	MI	BB	SV	DI	V	Aw	An	F	K	P		
AG	1	,32	,47	-,34	-,17	,47	,19	-,12	-,10	,51	-,19	,05	,53	-,47	,28	,28		
FG	,32	1	,46	-,40	,29	,08	,22	-,27	-,08	,36	-,28	,37	,27	-,39	,15	,31		
KG	,47	,46	1	-,54	,19	,43	,04	-,30	,39	,39	-,18	,37	,56	-,71	,20	,41		
W	-,34	-,40	-,54	1	-,68	-,23	-,30	,26	-,30	-,26	,17	-,51	-,38	,81	-,46	-,49		
A	-,17	,29	,19	-,68	1	-,39	,22	-,22	,22	,12	-,14	,59	-,04	-,48	,37	,42		
P	,47	,08	,43	-,23	-,39	1	-,08	-,15	,36	,14	-,09	-,22	,61	-,32	,02	,16		
MI	,19	,22	,04	-,30	,22	-,08	1	-,26	-,37	,13	-,33	,56	,05	-,37	,24	,38		
BB	-,12	-,27	-,30	,26	-,22	-,15	-,26	1	-,08	-,35	,56	-,46	-,0,8	,45	-,02	-,60		
SV	-,10	-,08	,39	-,30	,22	,36	-,37	-,08	1	-,21	,28	,04	,23	-,37	,24	,00		
DI	,51	,36	,39	-,26	,12	,14	,13	-,35	-,21	1	-,51	,24	,38	-,46	,42	,33		
V	-,19	-,28	-,18	,17	-,14	-,09	-,33	,56	,28	-,51	1	-,30	-,15	,23	,11	-,44		
Aw	,05	,37	,37	-,51	,59	-,22	,56	-,46	,04	,24	-,30	1	-,09	-,61	,19	,55		
An	,53	,27	,56	-,38	-,04	,61	,05	-,08	,23	,38	-,15	-,09	1	-,53	,33	,12		
F	-,47	-,39	-,71	,81	-,48	-,32	-,37	,45	-,37	-,46	,23	-,61	-,53	1	-,40	-,61		
K	,28	,15	,20	-,46	,37	,02	,24	-,02	,24	,42	,11	,19	,33	-,40	1	-,11		
P	,28	,31	,41	-,49	,42	,16	,38	-,60	,00	,33	-,44	,55	,12	-,61	-,11	1		

In jedem Fall beeinflussen die Wechselbeziehungen die Analysen, weil sie die unabhängige Betrachtung der einzelnen Merkmale erschweren. In der Clusteranalyse führt eine starke Korrelation zweier Merkmalsvariablen dazu, dass diese stärker betont werden, indem beide in die Distanzberechnung zwischen den Aufgaben eingehen. So bedeutet die Korrelation von 0,81 zwischen dem Anforderungsbereich Wiedergabe und der Wissensart Fakten, dass die meisten Wiedergabe-Aufgaben Faktenwissen betreffen.

Ein erster Versuch zur Erkundung des Merkmalsgeflechts ist die folgende, visuelle Auswertung der obigen Relation, bei der direkt und indirekt korrelierte Elemente zu Gruppen zusammengefasst werden. Der Abstraktionsgrad und die Lernzielstufe Analysieren korrelieren (1). Mit der Lernzielstufe Analysieren korrelieren auch der Komplexitätsgrad (2) und der Anforderungsbereich Problemlösung (5). So bilden Abstraktionsgrad, Komplexitätsgrad, Problemlösung und Analysieren eine Gruppe direkt und indirekt korrelierter Merkmale. In der gleichen Art sind über die Lernzielstufe Anwenden die Merkmale Anforderungsbereich Anwendung (4), Prozessbereich Modellieren und Implementieren (6) sowie Prozedurwissen (8) miteinander verbunden. Die Anforderung Wiedergabe und die Wissensart Fakten sind miteinander (3), aber nicht mit weiteren Elementen korreliert, sie bilden eine eigene Gruppe. Auch der Prozessbereich Begründen und Bewerten und die Lernzielstufe Verstehen bilden eine eigene Gruppe (7).

- Abstraktionsgrad, Komplexitätsgrad, Anforderungsbereich Problemlösung, Lernzielstufe Analysieren
- Anforderungsbereich Anwendung, Prozessbereich Modellieren und Implementieren, Lernzielstufe Anwenden, Prozedurwissen
- Anforderungsbereich Wiedergabe, Faktenwissen
- Prozessbereich Begründen und Bewerten, Lernzielstufe Verstehen

Es wird deutlich, dass der Anforderungsbereich Wiedergabe hauptsächlich mit Faktenwissen verknüpft ist. Im Anforderungsbereich Anwendung sind Aufgaben des Prozessbereichs Modellieren und Implementieren vertreten, die Lernzielstufe Anwenden wird angesprochen und Prozedurwissen ist gefragt. Im Anforderungsbereich Problemlösen herrschen abstrakte und komplexe Probleme vor, die auf der Lernzielstufe Analysieren bearbeitet werden. In der letzten Gruppe zeigt sich, dass der Prozessbereich Begründen und Bewerten mit der Lernzielstufe Verstehen einhergeht und diese Merkmale ansonsten unverbunden sind. Die Gruppierung erscheint plausibel, schließt aber erst zwölf der sechzehn Merkmale ein. Gerade von den informatikspezifischen Prozessbereichen fehlen Darstellen und Interpretieren sowie Strukturieren und Vernetzen. In der anschließenden Clusteranalyse, in der sämtliche Merkmale berücksichtigt werden, verdichtet sich das Bild zu vier Aufgabentypen, die als Dimensionen des Aufgabenanspruchs gedeutet werden.

8.1.2 Clusteranalyse der Aufgaben

Die Clusteranalyse ist ein strukturentdeckendes Verfahren, das abhängig von den Merkmalsvariablen nach einem vorher festgelegten Distanzmaß Schritt für Schritt die jeweils ähnlichsten Aufgabenfälle zu Clustern zusammenfasst. Die Aufgaben innerhalb eines Cluster sind möglichst ähnlich bezüglich der Merkmale; zwischen den Clustern liegt ein möglichst großer Abstand. Die Clusteranalyse spürt auf, welche Anforderungstypen unterscheidbar sind.

Die Clusteranalyse der neunundzwanzig Aufgaben wird nach der Ward-Methode vorgenommen, das Distanzmaß ist die binäre quadrierte euklidische Distanz (Backhaus et al. 2008, S. 420f). Bei der Ward-Methode werden in jeder Iteration die Variablenmittelwerte aller Cluster berechnet und immer die Cluster fusioniert, deren Zusammenlegung bewirkt, dass der durchschnittliche Abstand aller Fälle von den Clustermittelwerten den geringsten Zuwachs erfährt. Versuche mit anderen Verfahren zur Clusterfusionierung erbrachten nahezu identische Cluster, wie bei Average Linkage, oder einen sehr großen und weitere sehr kleine Cluster, wie bei der Zentroid-Methode.

Die schrittweise Clusterfusionierung wird beim Stand von vier Clustern beendet, weil die durchschnittliche Distanz an dieser Stelle sprunghaft ansteigt, was im Dendrogramm deutlich wird (Abbildung 8.1). Die Clusteranalyse wird mit der Statistik-Software SPSS durchgeführt, das Ergebnis sind vier Aufgabencluster (Tabelle 8.2, S. 114). In der Interpretation werden die vier resultierenden Aufgabentypen näher beschrieben (siehe Abschnitt 9.2).

8.2 Aufgabenkennzahlen

In dieser Arbeit wird vorrangig und begründet angenommen, dass der Aufgabenanspruch ein mehrdimensionales Konstrukt ist. Kontrastierend wird im folgenden Exkurs der Gedanke aufgegriffen, die Schwierigkeit sei eindimensional und könne entsprechend einfach modelliert werden. Die empirische Schwierigkeit und begleitend die Trennschärfe werden als Aufgabenkennzahlen berechnet (Abschnitt 8.2.1). Wenn es gelingen würde, einen direkten Zusammenhang zwischen den Aufgabenmerkmalen und der Aufgabenschwierigkeit nachzuweisen und zu quantifizieren, dann könnte die Schwierigkeit anhand der Merkmale vorhergesagt werden. Varianzanalysen der Schwierigkeit nach den Merkmalen bestätigen jedoch, dass kein direkter, linearer Zusammenhang besteht (Abschnitt 8.2.2).

8.2.1 Schwierigkeit und Trennschärfe

Im allgemeinen Sprachgebrauch wie in der Pädagogik und der Psychologie wird der Begriff Schwierigkeit mit unterschiedlichen, oft unscharfen Bedeutungen assoziiert. Analog zur empirischen Itemschwierigkeit in der Psychologie wird unter der Aufgabenschwierigkeit hier die relative Lösungshäufigkeit verstanden, das ist

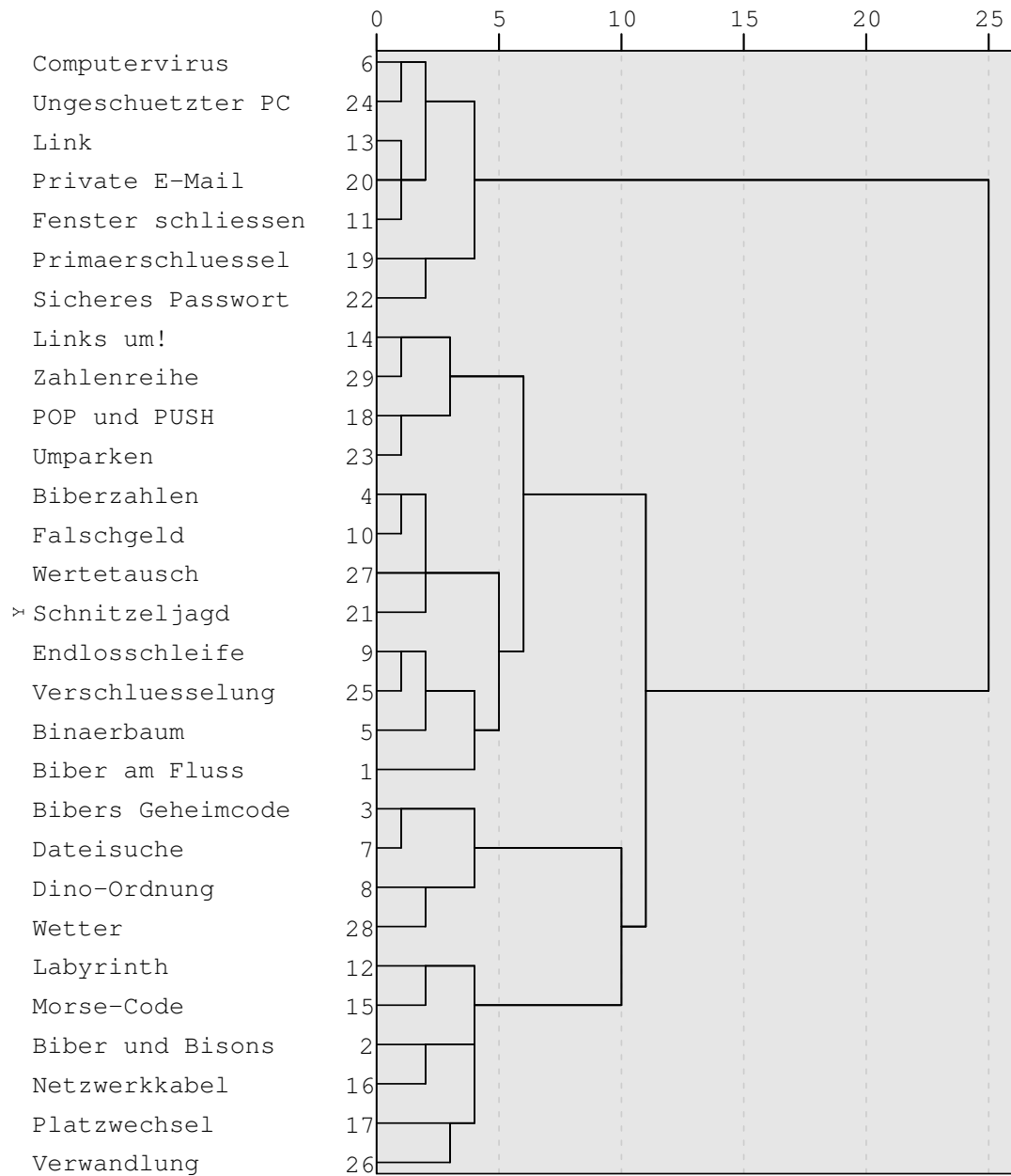


Abbildung 8.1: Dendrogramm zur Clusteranalyse der Aufgaben des Informatikwettbewerbs. Das Baumdiagramm wird von links nach rechts gelesen. Anfangs gibt es 29 Cluster, die jeweils eine Aufgabe enthalten. Die vertikale Verbindung zwischen zwei Clustern zeigt ihre Agglomeration an. Die Länge der horizontalen Linien zeigt an, wie groß die Distanz der agglomerierten Cluster ist (die Distanzwerte wurden auf den Wertebereich 0 bis 25 transformiert). Offenbar steigen die Distanzwerte bis zu einer Clusterzahl von vier nur langsam an und werden danach sprunghaft größer.

Tabelle 8.2: Clusterzuordnung der Aufgaben

Aufgabe	Cluster	Aufgabe	Cluster
1 Biber am Fluss	1	16 Netzwirkabel	2
2 Biber und Bisons	2	17 Platzwechsel	2
3 Bibers Geheimcode	3	18 POP und PUSH	1
4 Biberzahlen	1	19 Primärschlüssel	4
5 Binärbaum	1	20 Private E-Mail	4
6 Computervirus	4	21 Schnitzeljagd	1
7 Dateisuche	3	22 Sicheres Passwort	4
8 Dino-Ordnung	3	23 Umparken	1
9 Endlosschleife	1	24 Ungeschützter PC	4
10 Falschgeld	1	25 Verschlüsselung	1
11 Fenster schließen	4	26 Verwandlung	2
12 Labyrinth	2	27 Wertetausch	1
13 Link	4	28 Wetter	3
14 Links um!	1	29 Zahlenreihe	1
15 Morse-Code	2		

der Anteil der Personen, die die Aufgabe korrekt lösen. Während die Anforderungen, die einer Aufgabe innewohnen, unabhängig vom Bearbeiter sind, bezieht sich die empirische Schwierigkeit auf Personen. Somit werden die Kennzahlen Schwierigkeitsindex wie auch Trennschärfekoeffizient altersgruppenweise berechnet. Für eine Aufgabe, die sowohl in achten bis zehnten (8-10) als auch in elften und höheren Jahrgangsstufen (11+) gestellt wurde, existieren entsprechende Kennzahlen für die Stufen 8-10 und 11+. Erwartungsgemäß ist der Schwierigkeitsgrad derselben Aufgabe in der fünften bis siebten Klassenstufe (5-7) am höchsten, in der oberen Stufe (11+) am niedrigsten.

Empirische Aufgabenschwierigkeit

Der einfachste Schwierigkeitsindex ist der prozentuale Anteil der Personen, die eine Aufgabe lösen (Bühner 2006, S. 83). Der Index wird mit P bezeichnet, die Anzahl der Personen, die die betreffende Aufgabe gelöst haben, mit NR. NR ergibt sich als Spaltensumme in den Antwortmuster-Tabellen 7.2 bis 7.4 (S. 101). N ist die Anzahl der Personen, denen die Aufgabe vorlag, sprich die Anzahl der Teilnehmer in der Altersgruppe.

$$P = (NR / N) * 100$$

Zur Schwierigkeitsmessung der Wettbewerbsaufgaben wird ein Schwierigkeitsindex mit Inangriffnahmekorrektur verwendet (Bühner 2006, S. 85), bezeichnet mit PIK. Dafür setzt man die Anzahl der Personen, die die Aufgabe gelöst haben, NR, ins Verhältnis zur Anzahl der Personen, die sie überhaupt beantwortet haben, NB. NB ergibt sich in der Teilnahmedaten-Tabelle (S. 83), gefiltert nach der Altersgruppe, aus der Anzahl der Resultate „korrekt“ oder „inkorrekt“. Das heißt, außer „lag nicht vor“ wird insbesondere „unbeantwortet“ ausgeblendet.

$$\text{PIK} = (\text{NR} / \text{NB}) * 100$$

Die Variante mit zusätzlicher Inangriffnahmekorrektur berücksichtigt, wie viele Personen eine Aufgabe tatsächlich beantwortet haben, mit oder ohne Erfolg. Die Variante ist üblich, wenn es eine Zeitbeschränkung gibt, weil über die Nichtbearbeitung einer Aufgabe dann nicht bekannt ist, ob die Zeit zu knapp war oder die Aufgabe zu schwierig ist. Außerdem gibt es beim Informatik-Biber für falsch beantwortete Fragen Punktabzüge, was eventuell dazu führt, dass eine Lösung gefunden, aber aus Unsicherheit nicht genannt wird. Systematische Vergleiche beider Varianten und weiterer Indizes wie des „Index of Difficulty“, der nur die beiden Extremgruppen, die besten 27% und die schlechtesten 27% der Testteilnehmer einbezieht (Ebel und Frisbie 1991, S. 225), ergeben nur geringfügige und regelmäßige Verschiebungen der Schwierigkeitswerte. In Tabelle 8.3 (S. 116) sind die Schwierigkeitswerte (*Index der Schwierigkeit*) und deren Unterschiede zwischen den Jahrgangsstufen (Δ *Index*) sowie die Trennschärfen (*Koeffizient*) vollständig aufgelistet.

Zur übersichtlichen Darstellung in allen folgenden Diagrammen können die Aufgaben nach der Schwierigkeit in den Jahrgangsstufen geordnet werden. Die Werte reichen von 92,9% für **Fenster schließen** (S. 182) bis 20% für **Dino-Ordnung** (S. 179) – nur 20% der Teilnehmer, die die Aufgabe in Angriff nahmen, lösten sie korrekt. **Dino-Ordnung** ist also eine schwierige und **Fenster schließen** eine sehr leichte Aufgabe. Werden die stufenübergreifend gestellten Aufgaben zur Verankerung genutzt, gelingt es bis auf wenige Konflikte sogar, eine gemeinsame Ordnung aller Aufgaben herzustellen. Abbildung 8.2 (S. 117) zeigt die geordneten Schwierigkeitswerte aller Aufgaben, wobei mehrfach verwendete Aufgaben durch verbundene Markierungen dargestellt sind.

Konflikte ergeben sich nur an wenigen Stellen. Die Aufgabe **Schnitzeljagd** (S. 190), nach der Schwierigkeit in einer Altersgruppe eingeordnet, verletzt jeweils die Ordnung nach der Schwierigkeit in den beiden anderen Altersgruppen. Mit 35% korrekten Lösungen fällt die Aufgabe in Gruppe 5-7 überdurchschnittlich schwer aus, mit 57% in Gruppe 11+ überdurchschnittlich leicht. **Schnitzeljagd** wird letztendlich nach der Schwierigkeit in Gruppe 8-10 eingereiht, da dann die Abweichung in Summe möglichst gering ist. Genauso wird **Umparken** (S. 192) nach dem Index in 8-10 eingeordnet und fällt bezüglich der Schwierigkeit in 5-7 aus der Reihe. **Verschlüsselung** (S. 193) wird nach dem Index in 11+ eingeordnet, was sich nicht in die Ordnung nach dem Index in 8-10 einfügt.

Tabelle 8.3: Aufgabenkennzahlen. Schwierigkeitsindex, Differenz, Trennschärfe-koeffizient. Die Aufgabe Ungeschützter PC wurde nur in 5-7 und 11+ gestellt, daher wird die Schwierigkeitsdifferenz zwischen diesen Stufen gebildet.

Aufgabe	Index der Schwierigkeit (%)			Differenz: Δ Index (%)		Koeffizient der Trennschärfe (%)		
	5-7	8-10	11+	5-7, 8-10	8-10, 11+	5-7	8-10	11+
1 Biber am Fluss	84,3					46,5		
2 Biber und Bisons		69,7	79,3		9,6		39,4	36,9
3 Bibers Geheimcode	39,9	56,1		16,2		49,7	48,8	
4 Biberzahlen		30,1	46,3		16,2		40,8	46,8
5 Binärbaum		59,6					46,0	
6 Computervirus	55,5	67,3	77,8	11,8	10,5	42,5	38,4	35,4
7 Dateisuche	47,5					43,8		
8 Dino-Ordnung	20,0					42,7		
9 Endlosschleife		47,7	66,1		18,4		39,6	46,8
10 Falschgeld		39,1	58,7		19,6		43,6	51,6
11 Fenster schließen	92,9					37,2		
12 Labyrinth			48,9					36,7
13 Link	77,4					47,4		
14 Links um!	68,8					51,4		
15 Morse-Code			25,7					39,1
16 Netzwerkkabel		51,4	67,3		15,9		50,4	48,4
17 Platzwechsel	82,8					46,2		
18 POP und PUSH			51,7					48,0
19 Primärschlüssel		31,7	50,6		18,9		37,1	45,6
20 Private E-Mail		47,4					39,9	
21 Schnitzeljagd	35,3	42,4	57,4	7,1	15,0	39,0	47,5	46,9
22 Sicheres Passwort	69,5	68,9		-0,6		38,9	28,4	
23 Umparken	62,5	77,2		14,7		52,7	49,1	
24 Ungeschützter PC	21,7		24,7		3,0	26,7		27,1
25 Verschlüsselung		31,5	39,4		7,9		36,4	35,7
26 Verwandlung			43,4					51,5
27 Wertetausch			35,9					46,5
28 Wetter	23,9					34,9		
29 Zahlenreihe	33,5	46,9		13,4		47,5	47,6	
<i>Mittelwert:</i>	54,3	51,1	51,5	10,4	14,7	43,2	42,2	42,9
<i>Standardabweichung:</i>	24,5	15,0	16,7	6,2	4,3	6,9	6,1	7,2

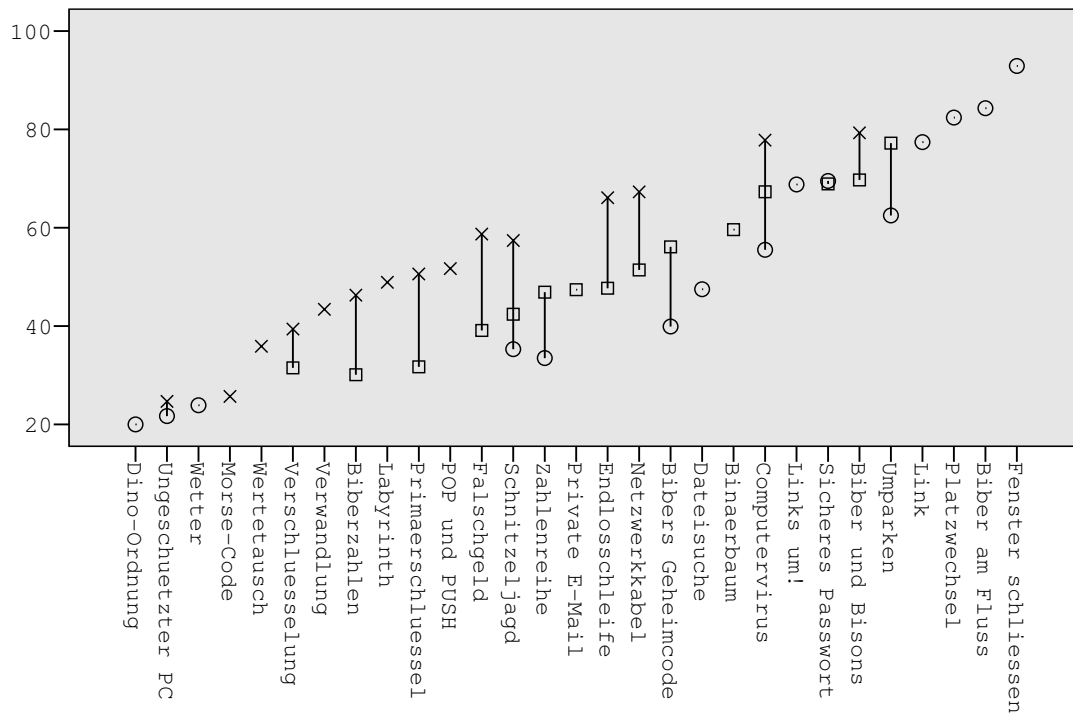


Abbildung 8.2: Anordnung der Aufgaben nach dem Schwierigkeitsindex in den Jahrgangsstufen fünf bis sieben (○), acht bis zehn (□) und elf oder darüber (×)

Die Aufgabe Ungeschützter PC (S. 193) besitzt in der jüngsten wie in der ältesten Gruppe einen konstant hohen Schwierigkeitsgrad. Da Anwenderwissen erfragt wird, das in der elften Jahrgangsstufe längst ohne schulischen Informatikunterricht vorhanden sein sollte, wird vermutet, dass die Schwierigkeit auf Gründe zurückgeht, die nicht Bestandteil dieser Untersuchung sind, etwa die missverständliche Formulierung der Antwortalternativen. Aus diesem Grund wird die Aufgabe von den weiteren Analysen ausgeschlossen (vgl. S. 100).

Trennschärfe

Unter der Trennschärfe wird der Grad verstanden, in dem eine Aufgabe zwischen den Testpersonen differenziert, das heißt, Personen mit hoher Leistung im Gesamttest lösen die Aufgabe deutlich öfter als Personen mit geringer Gesamtleistung. Ein mögliches Maß ist der Korrelationskoeffizient zwischen Aufgabenergebnis und Gesamtergebnis, zum Beispiel die punktbiseriale Korrelation zwischen einer dichotomen und einer intervallskalierten Variable (Bühner 2006, S. 97f). Ein anderes Trennschärfemaß ist die Schwierigkeitsdifferenz, von Ebel und Frisbie (1991) als „Index of Discrimination“ zwischen den beiden Extremgruppen berechnet, den besten 27% und den schlechtesten 27% der Testteilnehmer (Ebel und Frisbie 1991, S. 225). Die Trennschärfe bemisst also, in welchem Aus-

maß die Aufgabe den Gesamttest repräsentiert. Eine trennscharfe Aufgabe trennt möglichst scharf zwischen im Testsinn guten und schlechten Schülern.

Vergleiche verschiedener Trennschärfe-Kennzahlen ergeben für die Wettbewerbsaufgaben lediglich Verhältnisunterschiede, keine Rangunterschiede. Für die Studie wird als Trennschärfekoeffizient einer Aufgabe die punktbiseriale Korrelation zwischen dem Aufgabenergebnis und dem Gesamtergebnis gewählt, also der Anzahl der gelösten Aufgaben einer Person, die sich als Zeilensumme in den Antwortmuster-Tabellen 7.2 bis 7.4 (S. 101) ergibt. Der Vorteil des Korrelationskoeffizienten besteht darin, dass die Variable, die als Testergebnis herangezogen wird, mit geringem Aufwand durch andere Variablen ersetzt werden kann, zum Beispiel durch die Punktzahl, die im Wettbewerb vergeben wurde, und die auf theoretischen Schwierigkeitseinschätzungen der Aufgabensteller beruht. Die Trennschärfekoeffizienten aller Aufgaben sind neben den Schwierigkeitswerten und deren Unterschieden in Tabelle 8.3 (S. 116) aufgelistet.

8.2.2 Exkurs: Varianzanalyse der Merkmale

Wenn auch der Charakter dieser Arbeit strukturentdeckend und nicht strukturprüfend ist, stellt sich doch bei manchen Beobachtungen die Frage, ob ein direkter Zusammenhang zwischen den Aufgabentypen oder einzelnen Anforderungsmerkmalen und der Aufgabenschwierigkeit nachzuweisen ist. Zum Beispiel entstammen dem Themenbereich Kodierung einige der schwierigsten Aufgaben, **Bibers Geheimcode** (S. 175) in Jahrgangsstufe 5-7, **Biberzahlen** (S. 176) und **Verschlüsselung** (S. 193) in Stufe 8-10, **Morse-Code** (S. 185) in Stufe 11+. Dem Zusammenhang zwischen der empirischen Aufgabenschwierigkeit, der Trennschärfe und dem Inhaltsbereich ist der Abschnitt 9.1 gewidmet. Anlass zu dem folgenden Exkurs gibt die Überlegung, dass einfachsten Falles ein quantifizierbarer, linearer Einfluss der Merkmale auf die Schwierigkeit besteht, denn dann könnten durch den geschulten Beobachter die Merkmale einer Aufgabe eingeschätzt und die Schwierigkeit daraus prognostiziert werden.

Zur Relevanzprüfung der Merkmale für die Aufgabenschwierigkeit eignet sich das Verfahren der Varianzanalyse. Bei der einfaktoriellen Varianzanalyse werden die Aufgaben nach den Kategorien eines Merkmals gruppiert, etwa konkret oder abstrakt. Die Nullhypothese „gleiche Mittelwerte des Schwierigkeitsindex in den Kategorien“ wird einem Signifikanztest unterzogen, hier dem F-Test. Muss die Nullhypothese akzeptiert werden, so bedeutet das, durch das Merkmal wird keine sinnvolle Einteilung bezüglich der Schwierigkeit vorgenommen. Kann sie aber verworfen werden, weil zwischen den Kategorien ein deutlicher Schwierigkeitsunterschied besteht, so ist das Merkmal augenscheinlich schwierigkeitsrelevant (Backhaus et al. 2008, S. 154ff).

Die Einflussvariablen Anforderungsbereich, Prozessbereich, Lernzielstufe und Art des Wissens können nicht ohne Weiteres varianzanalytisch untersucht werden, da durch die Mehrfachnennungen keine disjunkte Kategorienbildung möglich

ist. Die Merkmale müssen vorab kodiert werden. Die Analysen werden in zwei Varianten durchgeführt, einmal mit einer multiplen Kodierung, einmal mit einer Dummy-Kodierung (vgl. S. 98f). Bei der *multiplen Kodierung* wird für jede Mehrfachnennung eine neue Kategorie eröffnet, zum Beispiel für die häufige Lernzielkombination Anwenden und Analysieren (siehe Tabelle 6.5). Weder für die solcherart kodierten Variablen Anforderungsbereich, Prozessbereich, Lernzielstufe, Art des Wissens noch für die Variablen Abstraktionsgrad, Formalisierungsgrad, Komplexitätsgrad ist ein signifikanter linearer Zusammenhang mit der Aufgabenschwierigkeit nachweisbar.

Bei der *Dummy-Kodierung* wird für jede Merkmalsausprägung eine eigene Variable definiert, zum Beispiel treten die Variablen Wiedergabe, Anwendung und Problemlösung an die Stelle der Variablen Anforderungsbereich (Tabelle 7.1, S. 99; S. 100 zur dichotomen Kodierung). Die Varianzanalysen mit den nun sechzehn Variablen werden mit der Statistik-Software SPSS als einfaktorielle ANOVA (ANalysis OF VAriance) berechnet. Lediglich in der jüngsten Altersstufe bestätigt sich ein auf dem Niveau 5% signifikanter Zusammenhang der Prozessbereiche Begründen und Bewerten, Strukturieren und Vernetzen, Darstellen und Interpretieren sowie der Lernzielstufe Verstehen mit der Aufgabenschwierigkeit. Für die übrigen Altersgruppen und Merkmale ist kein Zusammenhang nachweisbar.

Das Ergebnis ist unter dem Vorbehalt zu sehen, dass die Aufgaben nicht als Testaufgaben für die Merkmale konstruiert wurden. Die Kategorien sind teilweise sehr dünn oder gar nicht besetzt und vor allem wurde nicht ausgeschlossen, dass Merkmale gehäuft gleichzeitig auftreten oder vorwiegend in leichten oder schwierigen Aufgaben vorkommen. Daher wird die Vermutung aufrechterhalten, dass die Merkmale eine erklärende Rolle für den Kompetenzanspruch spielen, wenn auch nicht unbedingt linear und in einer Dimension modellierbar. Einen Hinweis, dass das durch die Aufgabenschwierigkeit erfasste Konstrukt nicht eindimensional ist, liefert schon die geringe Homogenität der Aufgaben, erfasst durch die Item-Interkorrelation, die in allen Altersgruppen weniger als 20% beträgt (Bortz und Döring 2006, S. 220).

Wenn sich auch aus den Aufgabenanalysen keine Hinweise auf ein eindimensionales Schwierigkeitsmodell und auf eine unmittelbare prognostische Qualität der Merkmale ergeben, so wird der Anspruch einer Aufgabe doch am ehesten durch die Aufgabenmerkmale charakterisiert. Diese bleiben zentrales Thema der Untersuchung, die sich nach diesem Exkurs wieder der These einer mehrdimensionalen Anforderungs- und Fähigkeitsstruktur zuwendet.

8.3 Fähigkeitsmodell A (Antwortmuster)

In den kommenden Überlegungen wird die Perspektive von den Aufgaben-Anforderungen auf die Aufgaben-Bewältigung durch die Personen verlagert. Bisher wurde in Abschnitt 8.1 die Aufgabenklassifizierung betrachtet (welche Merk-

male trägt die Aufgabe?) und in Abschnitt 8.2 wurden die Kennzahlen Aufgabenschwierigkeit und Trennschärfe thematisiert. Die Aufgabenschwierigkeit wird aus den Spaltensummen der Wettbewerbsergebnisse berechnet (wie viele Schüler lösen die Aufgabe?), die Trennschärfe aus den Zeilensummen (wie viele Aufgaben löst der Schüler?). Jetzt rücken die Antwortmuster in das Blickfeld: *welche* Aufgaben löst der Schüler?

8.3.1 Faktorenanalyse der Antwortmuster

Die Antwortmuster stehen in den Tabellen 7.2 bis 7.4 (S. 101) bereit. Jeder Altersgruppe lag ein eigener Aufgabensatz vor und die Analysen werden für jede der Stufen fünf bis sieben (5-7), acht bis zehn (8-10) und elf oder höher (11+) extra ausgeführt. Das spiegelt, dass die Fähigkeit, die Aufgaben zu lösen, sich mit dem Alter des Bearbeiters verändert. Die Tabellen enthalten die Teilnehmer als Fälle und die Aufgaben als Variablen, deren Wert 1 ist, wenn ein Schüler die Aufgabe löste, sonst ist der Wert 0. Mit einer Faktorenanalyse sollen die Aufgaben zusammengefasst werden, die dieselben Schüler lösen konnten, die also bezüglich dieser Schülergruppe in einer gemeinsamen Fähigkeitsdimension liegen. Die so gebündelten Aufgabenvariablen werden durch Faktoren ersetzt, die als Repräsentanten der Dimensionen verstanden werden.

Der faktorenanalytische Ansatz muss für dichotome Daten kritisch hinterfragt werden. Bortz (2005) empfiehlt, nur Variablen zu faktorisieren, bei denen die Enge des linearen Zusammenhangs bestimmt werden kann. Nominal oder ordinal skalierte Merkmale seien weniger geeignet. Allerdings werde die Struktur nicht erheblich verändert, wenn dichotomisierte Merkmale faktorisiert werden, deren Zusammenhänge über den Phi-Koeffizienten oder die punktbiseriale Korrelation ermittelt wurden (Bortz 2005, S. 524). Nach Bühner (2006) sind Normalverteilung und Intervallskalenniveau optimale Bedingungen, jedoch keine zwingenden Voraussetzungen für eine Hauptkomponentenanalyse. Im Falle dichotomer und gleichzeitig unterschiedlich schwieriger Items bilden sich indes leicht *Schwierigkeitsfaktoren*, die Items gleicher Schwierigkeit bündeln (Bühner 2006, S. 196f).

Bartholomew et al. (2008) erklären die Faktorenanalyse binärer Daten prinzipiell für ungeeignet, denn das Modell der Faktorenanalyse, in dem die Ausgangsvariablen durch Linearkombination der Faktoren rekonstruierbar sind, passe generell nicht zu kategorialen, insbesondere binären Variablen. Dem Modell nach könne theoretisch jeder mögliche reelle Wert erreicht werden, was der Voraussetzung binärer Werte widerspreche (Bartholomew et al. 2008, S. 212). Bartholomew et al. empfehlen, auf das Logit-Modell der Item-Response-Theorie auszuweichen, das latente, also nicht beobachtete Variablen unterstellt, die die beobachteten Variablen beeinflussen (ebd., S. 213f). Alternativ spielen sie unter der Bezeichnung „Underlying Variable Approach“ die Vermutung durch, hinter den binären Messwerten stehen kontinuierliche Variablen. Bis zu einer bestimmten Schwelle der kontinuierlichen Variable hat ihre binäre Realisierung den Wert null, oberhalb

davon den Wert eins. Dann kann auf der Basis der tetrachorischen Korrelation, die den Zusammenhang unbeobachteter Variablen erfasst, eine klassische Faktorenanalyse gerechnet werden (Bartholomew et al. 2008, S. 224f; zur tetrachorischen Korrelation siehe auch Bortz 2005, S. 230f).

Hier ist das Ziel, mit klassifizierenden statistischen Methoden die dimensionale Struktur der Fähigkeiten zu erforschen, die die Schüler bei der Lösung der Informatikaufgaben zeigen. Die explorative Faktorenanalyse ist das Verfahren, welches den Ansatz, wenige Dimensionen aus der Merkmalsvielfalt herauszuarbeiten, inhaltlich optimal verwirklicht. Technisch würde die oben skizzierte Idee des Underlying Variable Approach die fragliche Faktorisierung dichotomer Daten umgehen, ist aber kaum vereinbar mit der Lesart der Aufgabenvariablen. Diese sind keine künstlich dichotomisierten Variablen, denen ein kontinuierlicher Wertebereich angemessen wäre. Wenn auch denkbar ist, eine Aufgabe nur zum Teil zu lösen, ist dieser Teil doch nicht kontinuierlich zu beziffern.

Obgleich formal nicht einwandfrei, werden die binären Antwortmuster einer Faktorenanalyse unterzogen. Das geschieht in dem Bewusstsein, dass die explorativen Ergebnisse keine Beweiskraft haben. Der Anspruch an das Verfahren kann nicht sein, Hypothesen zu prüfen, jedoch wird das Ziel verfolgt, Daten zu reduzieren und Hypothesen zu generieren. In der Interpretation wird sich zudem erweisen, dass das Resultat unbefriedigend ausfällt (siehe S. 142ff), wofür das Skalenniveau einer unter mehreren möglichen Gründen ist. Die Arbeitsschritte einer Faktorenanalyse werden im nächsten Abschnitt anschaulich beschrieben, am Beispiel der Untersuchung intervallskalierteter Merkmale (siehe S. 125). Dieser Abschnitt beschränkt sich auf eine kurze Ergebnisdarstellung.

Die Analyse wird als Hauptkomponentenanalyse mit rechtwinkliger Varimaxrotation der Faktoren in der Statistik-Software SPSS implementiert (Backhaus et al. 2008, S. 324ff). Systematische Experimente mit anderen Einstellungen rufen nur geringfügige Änderungen im Ergebnis hervor. Nach dem Eigenwert-Kriterium werden in der unteren und mittleren Altersgruppe drei, in der oberen zwei Faktoren extrahiert. Die Ladungsmatrizen sind am Kapitelende abgebildet (Tab. 8.4, 8.6, 8.8, S. 127ff). Zur Auswertung werden die Aufgaben nach ihren stärksten Ladungen den Faktoren zugeschrieben. In der Theorie werden verschiedene Werte als bedeutsame Ladungen diskutiert. Bortz (2005) nennt 0,4 als unteren Grenzwert für Faktorladungen, die bei der Interpretation berücksichtigt werden (Bortz 2005, S. 551). Backhaus (2008) erachtet 0,5 als hohe Ladung (Backhaus 2008, S. 356). Praxis ist, die Grenzen so festzusetzen, dass die Zuordnung von Variablen zu Faktoren Plausibilitätsbetrachtungen standhält. In dieser Studie erweist es sich als aussagekräftig, Aufgabenvariablen mit einer Ladung ab rund 0,4 einzubeziehen. Die Faktorzuordnung der Aufgaben ist ebenfalls am Kapitelende abgebildet (Tab. 8.5, 8.7, 8.9, S. 127ff). Anhand der Merkmalsprofile der assoziierten Aufgaben werden die Faktoren interpretiert. Das Erklärungsmodell, das durch die Faktorisierung der Antwortmuster entsteht, wird „Fähigkeitsmodell A“ genannt (siehe Abschnitt 9.3).

8.4 Fähigkeitsmodell B (Merkmalsquoten)

Im Fähigkeitsmodell A werden die Dimensionen durch die Aufgaben gekennzeichnet, die die gleichen Schüler lösten beziehungsweise nicht lösten. Die Abstraktion von den Aufgaben erfolgt erst in der Interpretationsphase, wenn die Fähigkeiten durch die Merkmale beschrieben werden. Das Fähigkeitsmodell B, dessen Entwicklung im Folgenden beschrieben wird, abstrahiert schon in der Analysephase von den Aufgaben. Die Dimensionen werden unmittelbar aus den Merkmalen konstruiert.

Aus der Aggregation (S. 102) liegt je Wettbewerbsteilnehmer seine Lösungsquote für jedes Merkmal vor. Das heißt, die Aufgabenmerkmale werden unmittelbar auf Personen bezogen. Die Tabellen 7.5 bis 7.7 (S. 103) enthalten die Personen-Merkmalsprozentage als intervallskalierte Variablen. Eine Faktorisierung der Merkmalsquoten ergibt Dimensionen, die durch die Merkmale gekennzeichnet sind, die bei den gleichen Personen gleich ausgeprägt sind.

Spätestens an diesem Punkt wird die Konfundierung der Effekte offenbar. Wie ausgeführt wurde, sind die Merkmale bereits in den Aufgaben korreliert (vgl. S. 104). Dadurch ist dieser Aufgabensatz im Grunde genommen ungeeignet, das gemeinsame Auftreten von Fähigkeitsmerkmalen zu analysieren – könnte doch jedweder scheinbare Zusammenhang zwischen den Fähigkeiten in Wahrheit eine Wechselbeziehung der Merkmale auf der Aufgabenebene sein. Da die Erprobung des Vorgehens in dieser Arbeit einen hohen Stellenwert innehat, und da das Konstruktionsverfahren mehr noch als das Modell selber exemplarisch ist, wird die Analyse dennoch durchgeführt und kritisch-konstruktiv diskutiert.

8.4.1 Wechselbeziehungen der Merkmale

Das erwartete Fähigkeitsmodell ist personenbezogen und damit altersabhängig. Zur Sondierung des Merkmalsgeflechts werden die Korrelationen innerhalb der drei Aufgabensätze für die Jahrgangsstufen 5-7, 8-10 und 11+ untersucht, analog zum Vorgehen während der Anforderungsmodellierung, wo alle neunundzwanzig Aufgaben gemeinsam analysiert wurden (vgl. S. 109). Der Phi-Koeffizient für dichotome Merkmale wird berechnet (Bortz 2005, S. 227ff) und die Merkmalspaare, deren Korrelation mindestens 0,5 auf dem Signifikanzniveau 0,01 beträgt, werden mit ihrem Phi-Wert¹ aufgelistet.

Jahrgangsstufe 5-7

1. Abstraktionsgrad, Anforderungsbereich Problemlösen (0,68)
2. Anforderungsbereich Wiedergabe, Faktenwissen (0,84)
3. Prozessbereich Begründen und Bewerten, Lernzielstufe Verstehen (0,69)
4. Lernzielstufe Anwenden, Prozedurwissen (0,69)

¹Die vollständigen Korrelationstabellen sind in Anhang C abgedruckt (S. 214ff).

Das sind in der fünften bis siebten Jahrgangsstufe vier klar trennbare Fähigkeitskomponenten, in denen abstrakte Probleme lösen (1), Faktenwissen wiedergeben (2), Verstehen, Begründen und Bewerten (3), prozedurales Wissen anwenden (4) relevant sind.

Jahrgangsstufe 8-10

1. Formalisierungsgrad, Prozessbereich Darstellen und Interpretieren (0,85)
2. Komplexitätsgrad, Prozessbereich Darstellen und Interpretieren (0,69)
3. Komplexitätsgrad, Lernzielstufe Analysieren (0,67)
4. Anforderungsbereich Anwendung, Lernzielstufe Anwenden (0,74)
5. Prozessbereich Begründen und Bewerten, Faktenwissen (0,84)
6. Lernzielstufe Anwenden, Prozedurwissen (0,75)

Werden korrelierte Paare wie (1) und (2), die den Prozessbereich Darstellen und Interpretieren als gemeinsames Element enthalten, zu Gruppen direkt und indirekt verknüpfter Merkmale zusammengezogen (vgl. S. 111), vereinfachen sich die Korrelationsbeziehungen.

- Formalisierungsgrad, Komplexitätsgrad, Prozessbereich Darstellen und Interpretieren, Lernzielstufe Analysieren
- Anforderungsbereich Anwendung, Lernzielstufe Anwenden, Prozedurwissen
- Prozessbereich Begründen und Bewerten, Faktenwissen

So ist in der achten bis zehnten Jahrgangsstufe die erste Fähigkeitskomponente formale und komplexe Sachverhalte analysieren, darstellen und interpretieren. Als zweite Komponente kommt die Anwendung von Prozedurwissen hinzu und als dritte Komponente Faktenwissen begründen und bewerten.

Jahrgangsstufe 11+

1. Komplexitätsgrad, Prozessbereich Darstellen und Interpretieren (0,82)
2. Anforderungsbereich Wiedergabe, Lernzielstufe Verstehen (0,65)
3. Anforderungsbereich Wiedergabe, Faktenwissen (0,78)
4. Anforderungsbereich Anwendung, Lernzielstufe Anwenden (0,76)
5. Anforderungsbereich Problemlösung, Prozessbereich Strukturieren und Vernetzen (0,82)
6. Anforderungsbereich Problemlösung, Lernzielstufe Analysieren (0,66)
7. Lernzielstufe Verstehen, Faktenwissen (0,83)

Aus diesen sieben Paarungen korrelierter Merkmale in Jahrgangsstufe elf und höher ergeben sich vier Merkmalsgruppen, charakterisiert durch Komplexes darstellen und interpretieren (1), die Wiedergabe und das Verstehen von Faktenwissen (2, 3 und 7), Anwendung (4), Probleme lösen durch Analysieren, Strukturieren und Vernetzen (5 und 6).

- Komplexitätsgrad, Prozessbereich Darstellen und Interpretieren
- Anforderungsbereich Wiedergabe, Lernzielstufe Verstehen, Faktenwissen
- Anforderungsbereich Anwendung, Lernzielstufe Anwenden
- Anforderungsbereich Problemlösung, Prozessbereich Strukturieren und Vernetzen, Lernzielstufe Analysieren

Besonders in den beiden Altersgruppen 5-7 und 11+ ist die Grundstruktur wiedererkennbar, die sich schon bei der gemeinsamen Analyse aller neunundzwanzig Aufgaben herauskristallisierte. Vier Fähigkeitskomponenten heben sich ab, die Anforderungsbereiche Wiedergabe, Anwendung und Problemlösung sowie eine Komponente „Argumentieren“, die in 5-7 durch Begründen und Bewerten auf der Lernzielstufe Verstehen, in 8-10 durch Analysieren, Darstellen und Interpretieren formaler und komplexer Sachverhalte und in 11+ durch Darstellen und Interpretieren komplexer Sachverhalte markiert wird.

Es fällt auf, dass die Wissensart Fakten in der niedrigsten Altersgruppe stark mit der Anforderung Wiedergabe, in der mittleren Altersgruppe mit dem Prozess Begründen und Bewerten und in der höchsten Altersgruppe mit Wiedergabe und der Lernzielstufe Verstehen zusammenhängt. Die Einbindung des relevanten Faktenwissens scheint mit der Altersstufe anspruchsvoller zu werden.

Während die Wechselbeziehungen zwischen den Aufgabenmerkmalen *Gegenstand* der Analysen des Aufgabenanspruchs sind, stellen sie für die Analysen der Personenfähigkeit einen *Störfaktor* dar, da die Anspruchsstruktur die Fähigkeitsstruktur unbeabsichtigt überlagert. Die Konfundierung der Effekte kann innerhalb des gegebenen Aufgabenmaterials nicht bereinigt werden. Die geplante Faktorenanalyse der Personen-Merkmalsquoten wird als vielversprechende Vorgehensalternative zur Faktorisierung der Aufgaben zunächst realisiert, muss jedoch unter Berücksichtigung der Konfundierungsproblematik interpretiert und zurückhaltend beurteilt werden.

8.4.2 Faktorenanalyse der Merkmalsquoten

Die Datenmatrix der Merkmalsquoten (wie viel Prozent der Aufgaben eines Merkmals löst der Schüler?) in Tabelle 7.5 bis 7.7 (S. 103) enthält für jeden Teilnehmer die schwer überschaubare Zahl von sechzehn Merkmalen. Die Personenfähigkeiten zu modellieren birgt das Anliegen, ihre Beschreibung zu vereinfachen. Unter der Prämisse, dass die Vielzahl von Merkmalen sich auf wenige Faktoren desselben Konstrukts zurückführen lässt, leistet eine Faktorenanalyse die Schätzung der Faktoren. Die explorative Analyse der sechzehn Merkmale wird als Hauptkomponentenanalyse mit rechtwinkliger Varimaxrotation der resultierenden Faktoren in der Statistik-Software SPSS realisiert (Backhaus et al. 2008, 324ff). Die hier bevorzugte Hauptkomponentenmethode zielt auf eine Datenreduktion ab, weniger auf die Aufdeckung latenter Variablen. Das entspricht dem Anliegen, Dimensionen, nicht Ursachen zu finden (Bühner 2006, S. 194ff).

Anschaulich liefert die Faktorenanalyse die Achsen eines Koordinatensystems, in dem sich jedes Merkmal ausdrücken lässt. Dazu werden von der Korrelationsmatrix ausgehend orthogonale Faktoren geschätzt, aus denen sich die Merkmalsvariablen durch Linearkombination rekonstruieren lassen. Die Faktorenanalyse ist nicht nur ein datenstrukturierendes Verfahren, welches Gruppen hochkorrelierter Variablen aufspürt. Sie wirkt auch dimensionsreduzierend, denn von den bis zu sechzehn Faktoren einer Datenmatrix mit sechzehn Variablen werden nur die wichtigsten Faktoren, die einen Großteil der ursprünglichen Information reproduzieren, beibehalten und interpretiert. Der Auswahlprozess wird mit *Faktorextraktion* bezeichnet.

Nach dem *Eigenwert-Kriterium* werden in der jüngsten Altersstufe vier, in den beiden anderen Gruppen jeweils drei Faktoren extrahiert. Der Eigenwert eines Faktors entspricht dem anteiligen Betrag der Varianz, die der Faktor reproduziert (bei sechzehn standardisierten Variablen beträgt die Gesamtvarianz 16). Nach dem Eigenwert-Kriterium werden nur die Faktoren im Modell belassen, deren Eigenwert mindestens 1 ist, die also mindestens so viel Varianz erklären wie eine einzelne, auf den Mittelwert 0 und die Standardabweichung 1 standardisierte Variable.

Als Rechenergebnis der Faktorenanalysen entstehen sogenannte Ladungsmatrizen, die am Kapitelende abgebildet sind (Tab. 8.10, 8.12, 8.14, S. 130ff). Die Faktorladungen entsprechen den Koeffizienten der Faktoren in der Linearkombination der Merkmalsvariablen. Die Sprechweise ist: „Die Variable lädt hoch/niedrig/mit einer bestimmten Ladung auf den Faktor.“ Zum Beispiel lädt in der Altersgruppe 5-7 das Merkmal Formalisierungsgrad mit ,620 auf den Faktor I. Je höher ein Merkmal auf einen Faktor lädt, desto größer ist die Bedeutung des Faktors für dieses Merkmal. Auf das Bild des Koordinatensystems übertragen heißt das, desto enger liegt das Merkmal in der Fähigkeitsdimension, die der Faktor repräsentiert.

Die Ladungsmatrizen haben auf den ersten Blick noch wenig Aussagekraft. Das Vorgehen zur Interpretation wird sein, Merkmal für Merkmal dem Faktor zuzuordnen, auf den es am höchsten lädt, und Faktor für Faktor anhand der zugeordneten Merkmale zu beschreiben. Die Deutung fällt besonders leicht, wenn eine Einfachstruktur vorliegt, das heißt, die Merkmalsvariablen jeweils auf einen Faktor stark laden und auf alle anderen schwach. Um diese Situation herbeizuführen, wird gegebenenfalls eine *Faktorenrotation* durchgeführt. Das durch die Faktoren dargestellte Koordinatensystem wird im Raum so rotiert, dass jedes Merkmal möglichst nahe bei einer der Achsen und entfernt von den anderen liegt. Als Rotationsmethode wird die rechtwinklige Varimaxrotation gewählt, die unabhängige Dimensionen anstrebt. Die abgebildeten Tabellen enthalten bereits die Faktorladungen nach der Rotation (Tab. 8.10, 8.12, 8.14, S. 130ff).

Den Faktoren werden jeweils die Merkmalsvariablen mit einer Hauptladung ab 0,6 eindeutig zugeordnet. Zusätzlich werden relativ starke Nebenladungen ab 0,4 hervorgehoben, weil sie für die Interpretation von Interesse sind (vgl. S. 121).

Die resultierenden Zuordnungstabellen sind am Kapitelende zusammen mit den Ladungsmatrizen abgedruckt (Tab. 8.11, 8.13, 8.15, S. 130ff). Die hochladenden Merkmale charakterisieren die Faktoren, die als Dimensionen Wiedergabe, Anwendung, Problemlösung interpretiert werden. In der jüngsten Altersgruppe zeichnet sich eine vierte Dimension Verständnis ab. Das Erklärungsmodell, das durch die Faktorisierung der Personen-Merkmalsquoten entsteht, wird „Fähigkeitsmodell B“ genannt (siehe Abschnitt 9.4).

8.5 Zusammenfassung

Als *Anforderungsmodell* wurde durch eine Clusteranalyse der Merkmale auf der Aufgabenebene ein vierdimensionales Erklärungsmodell der Aufgabenschwierigkeit herausgearbeitet. Als *Fähigkeitsmodelle* wurden durch Faktorenanalysen der Aufgaben respektive Merkmalsquoten auf der Personenebene mehrdimensionale Modellalternativen A und B gewonnen. Während das Anforderungsmodell die Verflechtung der *Aufgabenmerkmale* abbildet, gilt die Fähigkeitsmodellierung der Verflechtung der *Personeneigenschaften*. Für letztere ist problematisch, dass im Aufgabensatz nicht alle Merkmalskombinationen vertreten sind. Die Aufgaben wurden nicht systematisch als Merkmalsträger entworfen, so dass das Zusammenspiel der Personenfähigkeiten möglicherweise von den Wechselbeziehungen der Aufgabenanforderungen überlagert wird. Im nächsten Kapitel werden die Strukturmodelle der Anforderungen und Fähigkeiten inhaltlich ausgearbeitet.

Tabelle 8.4: Faktorladung der Aufgaben, Altersstufe 5-7.

Aufgabe	Faktor I	Faktor II	Faktor III
1 Biber am Fluss	,513	,218	,002
3 Bibers Geheimcode	,163	,584	-,067
6 Computervirus	,330	-,043	,404
7 Dateisuche	,084	,317	,274
8 Dino-Ordnung	-,132	,619	,017
11 Fenster schließen	,606	-,014	,030
13 Link	,492	,203	,027
14 Links um!	,381	,406	-,067
17 Platzwechsel	,439	,203	,077
21 Schnitzeljagd	-,098	,249	,514
22 Sicheres Passwort	,457	-,109	,042
23 Umparken	,341	,442	,049
28 Wetter	,038	-,080	,734
29 Zahlenreihe	,095	,468	,187
<i>Eigenwert des Faktors:</i>	2,331	1,074	1,023
<i>Varianzanteil/kumuliert (%):</i>	17 / 17	8 / 24	7 / 32

Tabelle 8.5: Faktorzuordnung der Aufgaben, Altersstufe 5-7. Hauptladungen ab 0,4; kursiv: Hauptladungen unterhalb 0,4 sowie auffallende Nebenladungen.

Aufgabe	Faktor I	Faktor II	Faktor III
11 Fenster schließen	,606		
1 Biber am Fluss	,513		
13 Link	,492		
22 Sicheres Passwort	,457		
17 Platzwechsel	,439		
8 Dino-Ordnung		,619	
3 Bibers Geheimcode		,584	
29 Zahlenreihe		,468	
23 Umparken		,442	
14 Links um!	<i>,381</i>	,406	
7 Dateisuche		<i>,317</i>	
28 Wetter			,734
21 Schnitzeljagd			,514
6 Computervirus			,404

Tabelle 8.6: Faktorladung der Aufgaben, Altersstufe 8-10.

Aufgabe	Faktor I	Faktor II	Faktor III
2 Biber und Bisons	,350	,283	,086
3 Bibers Geheimcode	,382	,299	,008
4 Biberzahlen	,436	,101	-,091
5 Binärbaum	,142	,513	-,031
6 Computervirus	,209	-,018	,636
9 Endlosschleife	,114	,452	,016
10 Falschgeld	,501	,096	,036
16 Netzwerkkabel	,414	,268	,060
19 Primärschlüssel	,495	-,058	,014
20 Private E-Mail	,363	,013	,269
21 Schnitzeljagd	,328	,321	-,024
22 Sicheres Passwort	-,169	,108	,769
23 Umparken	,259	,476	,145
25 Verschlüsselung	-,208	,675	,027
29 Zahlenreihe	,491	,105	,098
<i>Eigenwert des Faktors:</i>	2,396	1,065	1,023
<i>Varianzanteil / kumuliert (%):</i>	16 / 16	7 / 23	7 / 30

Tabelle 8.7: Faktorzuordnung der Aufgaben, Altersstufe 8-10. Hauptladungen ab 0,4; kursiv: Hauptladungen unterhalb 0,4 sowie auffallende Nebenladungen.

Aufgabe	Faktor I	Faktor II	Faktor III
10 Falschgeld	,501		
19 Primärschlüssel	,495		
29 Zahlenreihe	,491		
4 Biberzahlen	,436		
16 Netzwerkkabel	,414		
3 Bibers Geheimcode	,382		
20 Private E-Mail	,363		
2 Biber und Bisons	,350		
21 Schnitzeljagd	,328	,321	
25 Verschlüsselung		,675	
5 Binärbaum		,513	
23 Umparken		,476	
9 Endlosschleife		,452	
22 Sicheres Passwort			,769
6 Computervirus			,636

Tabelle 8.8: Faktorladung der Aufgaben, Altersstufe 11+.

Aufgabe	Faktor I	Faktor II
2 Biber und Bisons	,385	,225
4 Biberzahlen	,517	,106
6 Computervirus	,072	,356
9 Endlosschleife	,092	,599
10 Falschgeld	,489	,287
12 Labyrinth	-,150	,616
15 Morse-Code	,511	-,125
16 Netzwerkkabel	,328	,383
18 POP und PUSH	,186	,517
19 Primärschlüssel	,368	,280
21 Schnitzeljagd	,556	,047
25 Verschlüsselung	,122	,232
26 Verwandlung	,386	,399
27 Wertetausch	,473	,136
<i>Eigenwert des Faktors:</i>	2,647	1,030
<i>Varianzanteil / kumuliert (%):</i>	19 / 19	7 / 26

Tabelle 8.9: Faktorzuordnung der Aufgaben, Altersstufe 11+. Hauptladungen ab 0,4; kursiv: Hauptladungen unterhalb 0,4 sowie auffallende Nebenladungen.

Aufgabe	Faktor I	Faktor II
21 Schnitzeljagd	,556	
4 Biberzahlen	,517	
15 Morse-Code	,511	
10 Falschgeld	,489	
27 Wertetausch	,473	
2 Biber und Bisons	,385	
19 Primärschlüssel	,368	
12 Labyrinth		,616
9 Endlosschleife		,599
18 POP und PUSH		,517
26 Verwandlung	,386	,399
16 Netzwerkkabel		,383
6 Computervirus		,356
25 Verschlüsselung		

Tabelle 8.10: Faktorladung der Merkmale, Altersstufe 5-7.

Merkmal	Faktor I	Faktor II	Faktor III	Faktor IV
Abstraktionsgrad	,374	,021	,777	,165
Formalisierungsgrad	,620	,163	,587	,034
Komplexitätsgrad	,260	,724	,531	-,007
Wiedergabe	,171	,074	,126	,896
Anwendung	,675	,631	,197	,172
Problemlösung	,096	,028	,922	,093
Modellieren/Implementieren	,758	,142	,379	,111
Begründen/Bewerten	,069	,583	,022	,669
Strukturieren/Vernetzen	,222	,921	,065	,044
Darstellen/Interpretieren	,880	-,005	,146	,260
Verstehen	,099	,808	,028	,443
Anwenden	,654	,564	,345	,058
Analysieren	,344	,388	,693	,083
Fakten	,173	,102	,116	,930
Konzepte	,683	,474	,082	,173
Prozeduren	,711	,366	,469	,068
<i>Eigenwert des Faktors:</i>	8,335	2,259	1,634	1,120
<i>Varianzanteil/kumuliert (%):</i>	52 / 52	14 / 66	10 / 76	7 / 83

Tabelle 8.11: Faktorzuordnung der Merkmale, Altersstufe 5-7.

Merkmal	I	II	III	IV
Darstellen/Interpretieren	,880			
Modellieren/Implementieren	,758			
Prozeduren	,711		,469	
Konzepte	,683	,474		
Anwendung	,675	,631		
Anwenden	,654	,564		
Formalisierungsgrad	,620		,587	
Strukturieren/Vernetzen		,921		
Verstehen		,808		,443
Komplexitätsgrad		,724	,531	
Problemlösung			,922	
Abstraktionsgrad			,777	
Analysieren			,693	
Fakten				,930
Wiedergabe				,896
Begründen/Bewerten		,583		,669

Tabelle 8.12: Faktorladung der Merkmale, Altersstufe 8-10.

Merkmal	Faktor I	Faktor II	Faktor III
Abstraktionsgrad	,614	,653	,121
Formalisierungsgrad	,843	,401	,213
Komplexitätsgrad	,766	,567	,133
Wiedergabe	,073	,046	,852
Anwendung	,889	,219	,291
Problemlösung	,152	,935	,134
Modellieren/Implementieren	,835	,101	,131
Begründen/Bewerten	,298	,075	,892
Strukturieren/Vernetzen	,281	,702	,066
Darstellen/Interpretieren	,842	,412	,105
Verstehen	,115	,439	,699
Anwenden	,934	,234	,131
Analysieren	,662	,624	,130
Fakten	,145	,047	,956
Konzepte	,653	,598	,264
Prozeduren	,896	,125	,111
<i>Eigenwert des Faktors:</i>	9,596	2,445	1,378
<i>Varianzanteil / kumuliert (%):</i>	60 / 60	15 / 75	9 / 84

Tabelle 8.13: Faktorzuordnung der Merkmale, Altersstufe 8-10.

Merkmal	I	II	III
Anwenden	,934		
Prozeduren	,896		
Anwendung	,889		
Formalisierungsgrad	,843	,401	
Darstellen/Interpretieren	,842	,412	
Modellieren/Implementieren	,835		
Komplexitätsgrad	,766	,567	
Analysieren	,662	,624	
Konzepte	,653	,598	
Problemlösung		,935	
Strukturieren/Vernetzen		,702	
Abstraktionsgrad	,614	,653	
Fakten			,956
Begründen/Bewerten			,892
Wiedergabe			,852
Verstehen		,439	,699

Tabelle 8.14: Faktorladung der Merkmale, Altersstufe 11+.

Merkmal	Faktor I	Faktor II	Faktor III
Abstraktionsgrad	,789	,189	,336
Formalisierungsgrad	,806	,268	,350
Komplexitätsgrad	,833	,177	,501
Wiedergabe	-,016	,822	,038
Anwendung	,868	,241	,328
Problemlösung	,385	,182	,874
Modellieren/Implementieren	,862	,103	-,065
Begründen/Bewerten	,509	,710	,135
Strukturieren/Vernetzen	,269	,050	,898
Darstellen/Interpretieren	,839	,145	,405
Verstehen	,238	,873	,199
Anwenden	,923	,116	,249
Analysieren	,667	,189	,640
Fakten	,186	,941	,052
Konzepte	,789	,305	,375
Prozeduren	,808	,130	,449
<i>Eigenwert des Faktors:</i>	10,343	2,274	1,152
<i>Varianzanteil/kumuliert (%):</i>	65 / 65	14 / 79	7 / 86

Tabelle 8.15: Faktorzuordnung der Merkmale, Altersstufe 11+.

Merkmal	I	II	III
Anwenden	,923		
Anwendung	,868		
Modellieren/Implementieren	,862		
Darstellen/Interpretieren	,839		,405
Komplexitätsgrad	,833		,501
Prozeduren	,808		,449
Formalisierungsgrad	,806		
Abstraktionsgrad	,789		
Konzepte	,789		
Analysieren	,667		,640
Fakten		,941	
Verstehen		,873	
Wiedergabe		,822	
Begründen/Bewerten	,509	,710	
Strukturieren/Vernetzen			,898
Problemlösung			,874

Kapitel 9

Interpretation: Strukturmodelle

Das Ziel der vorangehenden statistischen Analysen war, rechnerisch einige wenige, möglichst unabhängige Dimensionen zu bestimmen, die sich im untersuchten Aufgabenrepertoire abzeichnen. In diesem Kapitel wird als fünfter Schritt im Forschungsprozess das entstehende Gerüst informatischer Anforderungen und Fähigkeiten in den Begriffen der Aufgabenmerkmale beschrieben.

Die unmittelbaren Analyseergebnisse bestehen in den Aufgabenkennzahlen empirische Schwierigkeit und Trennschärfe, den Aufgabenclustern und den Faktorensätzen. Die Aufgabenkennzahlen werden in Abschnitt 9.1 betrachtet. Die Aufgabentypen aus der Clusteranalyse werden in Abschnitt 9.2 als *Anforderungsdimensionen* interpretiert. Die Faktorensätze aus den Faktorenanalysen der Antwortmuster und der Merkmalsquoten werden in den Abschnitten 9.3 und 9.4 als *Fähigkeitsdimensionen* zweier alternativer Modellansätze interpretiert.

9.1 Empirische Aufgabenschwierigkeit

Die Tabelle 8.3 (S. 116) enthält den Schwierigkeitsindex und den Trennschärfe-koeffizienten, die als Maße der Aufgabengüte dienen und das Aufgabenmaterial erschließen. Die Kennzahlen erfassen leichte, schwierige und trennscharfe Aufgaben, die erfolgreiche Schüler herausheben. Vor allem die Anforderungsprofile hochgradig schwieriger Aufgaben können aufschlussreich sein, denn ihre Eigenschaften erzeugen offensichtlich Schwierigkeit. Zusätzlich wird in der Kennzahl Δ *Index* die Schwierigkeitsentwicklung zwischen den Jahrgangsstufen registriert, da Unregelmäßigkeiten auf inhaltliche Besonderheiten hindeuten können. So fällt die Aufgabe *Ungeschützter PC* (S. 193) durch ihren hohen Schwierigkeitsgrad für jüngere und unverändert auch für ältere Schüler auf, was sich nicht durch ihre Merkmalscharakteristik erklären lässt. Die Aufgabe wurde aus der Untersuchung herausgenommen (vgl. S. 100, S. 117).

Die Diagrammskizzen in Abbildung 9.1 vermitteln einen Überblick über die Schwierigkeit und Trennschärfe der Aufgaben in den drei Altersgruppen. Der

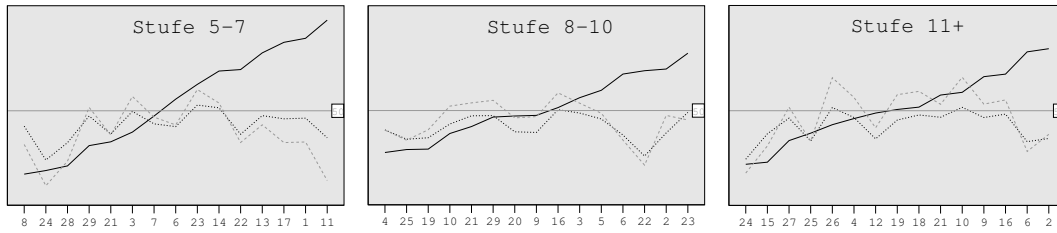


Abbildung 9.1: Aufgabenkennzahlen Schwierigkeitsindex (durchgezogene Linie), Trennschärfekoeffizient (schwarz gepunktete Linie), Index of Discrimination (grau gestrichelte Linie) für die Jahrgangsstufen 5-7, 8-10, 11+

Schwierigkeitsindex, berechnet als Prozentanteil der Lösungserfolge an den Lösungsversuchen, streut in der jüngsten Altersstufe von 20 % bis 92,9 %, in den anderen beiden Stufen nur von 30,1 % bis 77,2 % beziehungsweise von 25,7 % bis 79,3 % (schwarz durchgezogene Linie, siehe auch Tabelle 8.3, S. 116). Erwartungsgemäß weisen Aufgaben mittlerer Schwierigkeit eine höhere Trennschärfe auf als leichte und schwierige Aufgaben, die von sehr vielen Personen oder von sehr wenigen Personen gelöst werden. Dies wird durch den Index of Discrimination (grau gestrichelte Linie) stärker betont als durch den Trennschärfekoeffizienten (schwarz gepunktete Linie). Der *Discrimination-Index* berechnet die Trennschärfe aus der Differenz zwischen der Aufgabenschwierigkeit in der Gruppe der besten und in der Gruppe der schlechtesten Testteilnehmer. Der *Trennschärfekoeffizient* berechnet die Trennschärfe als punktbiseriale Korrelation zwischen dem Aufgabenergebnis und der Gesamtzahl gelöster Aufgaben. Korrelationstypisch gelten Werte im Bereich von 30 % bis 50 % (graue Bezugslinie) als mittlere, oberhalb davon als hohe Trennschärfe.

In den Streudiagrammen in Abbildung 9.2 werden Schwierigkeit und Trennschärfe je Wettbewerbsaufgabe gemeinsam dargestellt. Die Symbole tragen die Information, in welchen der drei Altersgruppen eine Aufgabe verwendet wurde. Die „Verwendung“ zeigt lediglich an, ob die Datenlage einen stufenübergreifenden Vergleich ermöglicht. Sie ist nicht als Aufgabenmerkmal „Eignung“ zu verstehen, weil eine Aufgabe nicht unbedingt in jeder Altersgruppe zur Verwendung kommt, für die sie geeignet erscheint.

In der fünften bis siebten Jahrgangsstufe fallen die Aufgaben *Dino-Ordnung* (S. 179), in der mehrere Prämissen – zweibeinig, vierbeinig, fing ein Tier – ausgewertet werden, um die Dinosaurierart zu bestimmen, und *Wetter* (S. 195), die umgekehrt nach einer Prämisse fragt, durch einen sehr hohen Schwierigkeitsgrad auf. Es stellt sich die Frage, inwieweit die Schwierigkeit durch den Inhaltsbereich, Logisches Schließen, erzeugt wird. Nach Jean Piaget (Lefrancois 1986) wird die Stufe der formalen Operationen, die durch abstraktes Denken und durch logisches Schlussfolgern gekennzeichnet ist, erst zu Beginn der Sekundarstufe erreicht (vgl. S. 70), was die hohe Schwierigkeit in der jüngsten Altersgruppe erklären kann.

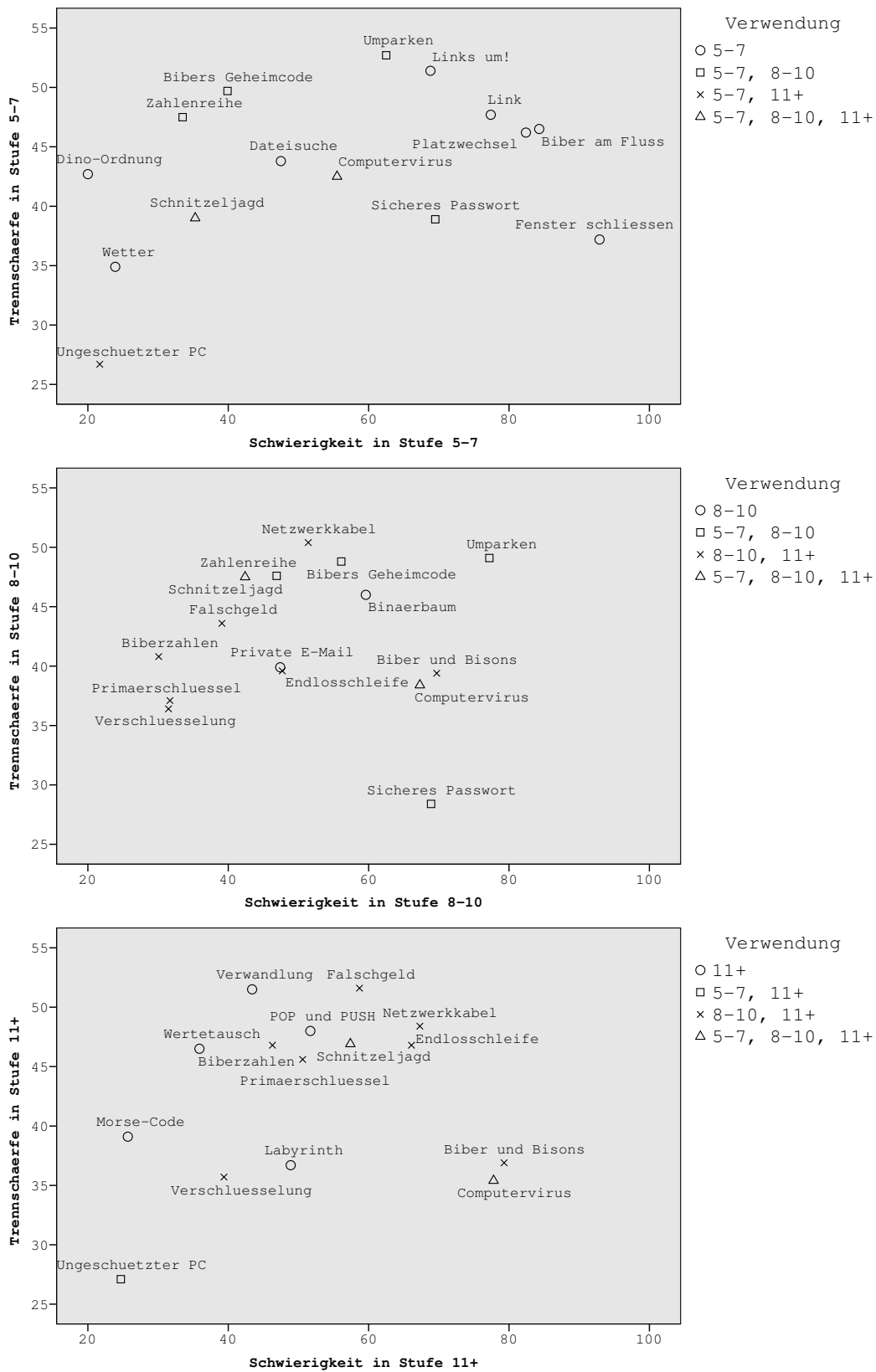


Abbildung 9.2: Streudiagramme der Schwierigkeit und Trennschärfe

In der mittleren und oberen Jahrgangsstufe stellt der Themenbereich Kodierung mit den Aufgaben **Biberzahlen** (S. 176) zur Binärcodierung, **Verschlüsselung** (S. 193) zur Eindeutigkeit einer Kodierung und **Morse-Code** (S. 185) zur Effizienz einer Kodierung einige der schwierigsten Aufgaben. Ähnlich schwierig wie die Aufgaben zur Kodierung fällt in der Jahrgangsstufe acht bis zehn die Datenbank-Aufgabe **Primärschlüssel** (S. 188) aus und in Stufe elf und höher die Algorithmen-Aufgabe **Wertetausch** (S. 195), die den Wertetausch zweier Variablen ohne Zwischenspeicherung in einer dritten Variable durchspielt.

Umparken (S. 192), zum Thema Kellerspeicher und damit dem Inhaltsbereich Datenstrukturen zugeordnet, und **Links um!** (S. 184), aus dem Bereich Algorithmen, sind die trennschärfsten Aufgaben in der Stufe fünf bis sieben. **Umparken** wird auch in der Stufe acht bis zehn gestellt, wiederum mit hoher Trennschärfe, aber mit geringer Schwierigkeit. Ähnlich verhält es sich mit der Aufgabe **Netzwerkkabel** (S. 186), in der redundante Kanten aus einem Graphen entfernt werden sollen. Sie weist die höchste Trennschärfe in der mittleren Altersgruppe auf und rangiert auch in der höchsten Altersgruppe unter den Aufgaben mit guter Trennschärfe bei moderater Schwierigkeit. Die höchste Trennschärfe besitzt in Stufe elf und höher wieder eine Graphen-Aufgabe, **Verwandlung** (S. 194), neben einer Algorithmen-Aufgabe, **Falschgeld** (S. 181), in der ein Ablaufplan zu deuten ist.

Die Aufgaben **Computervirus** (S. 178), zum Thema Benutzerverhalten, und **Schnitzeljagd** (S. 190), zum Thema Automaten, werden in allen Altersgruppen eingesetzt und erweisen sich als exemplarische Aufgaben mittlerer Schwierigkeit, die in der untersten Jahrgangsstufe am schwierigsten, in der oberen Jahrgangsstufe am leichtesten gelöst werden und deren Trennschärfe im mittleren Schwierigkeitsbereich am deutlichsten ausgeprägt ist.

Derartige Beobachtungen ergänzen die Interpretation der drei Modellansätze und tragen zur Erklärung der Ergebnisse bei. Obgleich sie interessante Ansatzpunkte für differenzierte Analysen informatischer Kompetenz bieten, nehmen diese hier keinen weiteren Raum ein, da sie nicht auf das Forschungsziel, eine Vorgehensweise zur Strukturmodellierung zu erproben, hinsteuern. Ausführlich werden Analysen, die sich auf die Aufgabenkennzahlen und die Merkmalsklassifizierung stützen, in vier assoziierten Examensarbeiten beschrieben.

Meinl (2012) analysiert zum Thema „Merkmale mädchen- bzw. jugungeeigneter Informatikaufgaben im Schülerwettbewerb Informatik-Biber“ die Differenz der Aufgaben-Erfolgsquoten von Mädchen und Jungen und stellt dar, dass die Merkmale Textrelevanz, Sprachliche Anforderungen, Sozialer Kontext und Anrede des Lesers typisch für Aufgaben sind, in denen Mädchen besser abschneiden. Umgekehrt sind Mathematik-Nähe und ein hoher Informationsgehalt der Bilder typisch für Aufgaben, in denen Jungen besser abschneiden.

Schmidt (2012) führt in der Arbeit „Merkmale altersgemäßer Informatikaufgaben am Beispiel eines Schülerwettbewerbs“ die Unterschiede der Erfolgsquoten zwischen den Jahrgangsstufen auf die Schwierigkeitsmerkmale Textmenge, Abstraktionsgrad, Erfahrungsweltnähe, Komplexität, Mathematischer Kalkül, Ver-

gleichsoperationen, Geometrie, Räumliches Vorstellungsvermögen, Umstrukturierung und Anordnung, Ineinanderschachtelung und Abweichung von der Norm einer Gesetzmäßigkeit zurück.

Krusch (2012) untersucht in der „Analyse des Zusammenhangs zwischen Informatikunterricht und dem Erfolg in einem Informatikwettbewerb“ die Situation des Informatikunterrichts – ob Pflichtfach oder Wahlfach, Stundenzahl, frühester Beginn etc. – in den Bundesländern. Für das Abschneiden im Wettbewerb belegt er, dass schulischer Informatikunterricht in der Vermittlung von Fachinhalten wie Algorithmen oder Datenschutz und Datensicherheit stärker wirksam ist als in der Vermittlung von Bedienerfertigkeiten, etwa in Computer-Grundkenntnissen oder Standardsoftware.

Hansky (2010) wertet in ihrer Arbeit „Typische Schülerfehler bei Informatikaufgaben“ die Falschantworten aus. Sie schöpft die Information aus, welche der je drei falschen Antworten gegebenenfalls zu der Bewertung „inkorrekt“ führte, indem sie die klassische Fehlertypologie nach Weimer (1925) auf Informatikaufgaben überträgt und um typische Informatikfehler ergänzt.

9.2 Anforderungsmodell

Aus der Clusteranalyse liegt in Tabelle 8.2 (S. 114) eine Zuordnung der Aufgaben aller Altersgruppen zu vier Clustern vor. Zur Profilbildung werden die Aufgaben in Tabelle 9.1 und 9.2 (S. 140f) clusterweise mit ihrer Merkmalsklassifizierung dargestellt. Als zusätzliche Information wird für jede Aufgabe ihr Inhaltsbereich und ihre Rangnummer in der altersstufenübergreifenden Anordnung nach der Schwierigkeit aufgeführt (vgl. S. 112). Da die Schwierigkeit je Altersstufe extra berechnet wird, also je Aufgabe gegebenenfalls mehr als ein Wert existiert, ist die Rangnummer hier die prägnantere Kennzahl. Mit in die Übersicht aufgenommen werden der Prozessbereich Kommunizieren und Kooperieren sowie die Lernzielstufen Erinnern und Bewerten, die nicht in die Clusteranalyse einbezogen wurden, da es in dem Aufgabensatz nicht mehr als drei Merkmalsträger gibt (vgl. S. 98). Auch diese Merkmale tragen zur Charakterisierung der Cluster bei.

Der erste ist mit zwölf Aufgaben zugleich der größte Cluster (Tab. 9.1, S. 140). Die Aufgaben, meist formal und komplex, gehören alle dem Anforderungsbereich Anwendung an und sind überwiegend den Prozessbereichen Modellieren und Implementieren oder Darstellen und Interpretieren zugeordnet. Die Lernzielstufe ist Anwenden, teilweise mit Tendenz zur Lernzielstufe Analysieren. Die relevanten Wissensarten sind gleichermaßen Konzept- wie Prozedurwissen. Bis auf **Schnitzeljagd** (S. 190), die dem Bereich Sprachen und Automaten entstammt, gehören die Aufgaben des Clusters entweder dem Bereich Algorithmen oder dem Bereich Information und Daten an. Die Aufgabe **Binärbaum** (S. 177) weist das clustertypische Merkmalsprofil auf, das nach der durchgängigen Anforderung mit *Anwendung* bezeichnet wird.

Der zweite Cluster enthält abstrakte und komplexe Aufgaben aus dem Anforderungsbereich Problemlösen (Tab. 9.1, S. 140). Die Prozessbereiche Strukturieren und Vernetzen sowie Darstellen und Interpretieren werden angesprochen. Sämtliche Aufgaben sind der Lernzielstufe Analysieren zugeordnet. Faktenwissen kommt nicht vor, aber Wissen über Konzepte und Prozeduren spielt eine Rolle. Ein typischer Vertreter dieses Clusters ist *Biber und Bisons* (S. 175), in der kein Lösungsweg vorgezeichnet ist. Vielmehr führt die logische Analyse der Problemstellung zum Ergebnis: Biber sagen immer die Wahrheit, die anderen Tiere lügen immer. Fragt ein Blinder eine Gruppe Tiere, wie viele unter ihnen Biber sind, wird die Biber-Zahl genau von den Bibern, das heißt genau in ihrer Anzahl genannt werden. Im Beispiel sind es drei Biber, denn genau drei Tiere sagen „3“. Der Cluster repräsentiert Aufgaben des Typs *Problemlösung*.

Der dritte und kleinste Cluster enthält komplexere Aufgaben, konkret und informell gestaltet, die dem Anforderungsbereich Anwendung entstammen (Tab. 9.2, S. 141). Alle vier Aufgaben sind dem Prozessbereich Strukturieren und Vernetzen zugeordnet, teils auch dem Prozessbereich Begründen und Bewerten, überwiegend auf den Lernzielstufen Verstehen oder Anwenden, zweimal auf Konzeptwissen und einmal auf Prozedurwissen bezogen. Bemerkenswert ist, dass drei der vier Aufgaben dem Inhaltsbereich Sprachen und Automaten angehören, zwei davon, *Wetter* (S. 195) und *Dino-Ordnung* (S. 179), verlangen Logisches Schließen. Das Vorherrschen des Merkmals Komplexität, der Prozessanforderung Strukturieren und Vernetzen und der Lernzielstufe Verstehen legt nahe, den Aufgabentyp des dritten Clusters als *Verständnis-Typ* zu charakterisieren.

Der vierte Cluster enthält einfache, konkrete, informelle Aufgaben (Tab. 9.2, S. 141). Sie entstammen überwiegend dem Anforderungsbereich Wiedergabe und sprechen den Prozessbereich Begründen und Bewerten an. Mit *Link* (S. 184) und *Private E-Mail* (S. 189) sind dem Cluster die einzigen vorkommenden Aufgaben aus dem Prozessbereich Kommunizieren und Kooperieren zugeordnet. Diese und *Fenster schließen* (S. 182) sind auch die einzigen Aufgaben der Lernzielstufe Erinnern. *Sicheres Passwort* (S. 191) ist die einzige Aufgabe der Lernzielstufe Bewerten. Sonst wird die Lernzielstufe Verstehen angesprochen. Alle Aufgaben sind mit der Wissensart Faktenwissen verknüpft. Auffallend ist, dass im Cluster alle Aufgaben der Inhaltsbereiche Informatiksysteme sowie Informatik, Mensch und Gesellschaft vereint sind, genauer, sämtliche Aufgaben zu den Themen Anwenderwissen und Benutzerverhalten. Die Aufgabe *Computervirus* (S. 178), in der es um die Haftung bei einem Virenbefall geht, ist ein typisches Beispiel für diesen Cluster. Nach dem prägenden Anforderungsbereich wird der Aufgabentyp des vierten Clusters *Wiedergabe-Typ* genannt.

Zusammenfassend fällt auf, dass die Clusterung nach den Merkmalen die Wettbewerbsaufgaben auch bezüglich ihres Inhaltsbereiches sinnvoll gruppiert. Dagegen scheint kein Zusammenhang mit der Schwierigkeit zu bestehen. In jedem Cluster sind gleichermaßen Aufgaben niedrigerer und höherer Rangstufen zu finden. Die Cluster repräsentieren vier Aufgabentypen.

- *Wiedergabe*: einfache Sachverhalte, Wiedergabe, Begründung und Bewertung, Erinnern und Verstehen von Faktenwissen
- *Anwendung*: formale und komplexe Problemstellung, Anwendung, Modellieren und Implementieren, Darstellen und Interpretieren, Anwenden von Konzept- und Prozedurwissen
- *Verständnis*: komplexe Sachverhalte, Anwendung, Strukturieren und Vernetzen, Anwenden von Konzept- und Prozedurwissen
- *Problemlösung*: abstrakte und komplexe Problemstellung, Strukturieren und Vernetzen, Darstellen und Interpretieren, Analysieren von Konzept- und Prozedurwissen

Damit hat die Clusteranalyse der Aufgaben vier charakteristische Profile herausgearbeitet. Die Profile erinnern an die bekannten Lernstufen nach Roth (1971), Reproduktion, Reorganisation, Transfer, Problemlösen, Kreativität (vgl. S. 20).

Der *Wiedergabe-Typ* (Cluster 4) entspricht mit der Wiedergabe einfacher Sachverhalte und dem Begründen, Bewerten, Erinnern und Verstehen von Faktenwissen der Stufe „Reproduktion des Gelernten“ (Roth 1971, S. 475).

In der Spezifikation des *Anwendungs-Typs* (Cluster 1) und des *Verständnis-Typs* (Cluster 3) sind die Stufen „Reorganisation unter neuen Aspekten“ und „produktive Transformation des Gelernten auf neue Probleme“ (ebd.) zu erkennen. Beide enthalten die Anwendung von Konzept- und Prozedurwissen, der Anwendungs-Typ mehr im Prozessbereich Modellieren und Implementieren, was der Reorganisation unter neuen Aspekten nahe kommt, der Verständnis-Typ mehr im Prozessbereich Strukturieren und Vernetzen, was eher mit dem Transfer des Gelernten vergleichbar ist.

Der Aufgabentyp *Problemlösung* (Cluster 2), der sich durch das Merkmal Analysieren abhebt, lässt sich mit der Rothschen Stufe Problemlösen in Einklang bringen. Die Stufe Kreativität wird durch die untersuchten Aufgaben nicht explizit abgebildet. Unbestritten ist Kreativität als Erfindungsgabe und schöpferische Kraft aber eine Voraussetzung zur Problemlösung.

Anhand der Merkmalsprofile ist festzustellen, dass es sich nicht um hierarchische Stufen in dem Sinn handelt, dass eine Aufgabe, die zum Beispiel vom Typ Verständnis ist, Fähigkeiten vom Typ Wiedergabe voraussetzt. Anhand der Schwierigkeitsränge ist festzustellen, dass es sich auch nicht um hierarchische Stufen in dem Sinn handelt, dass die Schwierigkeit mit der Stufe steigt. Die Typen werden vielmehr als Dimensionen begriffen. Auf die bildliche Vorstellung übertragen gibt es in dem Raum, der durch die vier Achsen Wiedergabe, Verständnis, Anwendung und Problemlösung aufgespannt wird, Idealtypen, die jeweils nur in einer Dimension liegen, aber auch Mischtypen, die kognitive Merkmale mehrerer Dimensionen tragen und demgemäß im Raum verortet sind.

Tabelle 9.1: Merkmalsprofile der Cluster 1 und 2. Abstraktionsgrad (AG), Formalisierungsgrad (FG), Komplexitätsgrad (KG), Anforderungsbereich Wiedergabe (W), Anwendung (A), Problemlösung (P), Prozessbereich Modellieren/Implementieren (MI), Begründen/Bewerten (BB), Strukturieren/Vernetzen (SV), Darstellen/Interpretieren (DI), Lernzielstufe Verstehen (V), Anwenden (Aw), Analysieren (An), Wissensart Fakten (F), Konzepte (K), Prozeduren (P). Gegebenenfalls keine Einschätzung (-).

Aufgabe	Gestaltung			Anforderg.			Prozess					Lernziel			Wissen			Rang,		
	AG	FG	KG	W	A	P	MI	BB	SV	KK	DI	E	V	Aw	An	B	F	K	P	Inhalt
Cluster 1																				
Biber am Fluss	AG				A					DI				An				K	P	2 ID
Biberzahlen	AG	FG	KG	A						DI			Aw						P	23 ID
Binärbaum		FG	KG	A						DI			Aw	An				K		11 ID
Endlosschleife	AG	FG	KG	A			MI			DI			Aw	An				K	P	14 A
Falschgeld		FG	KG	A						DI			Aw						P	18 A
Links um!				A			MI						Aw					-	-	6 A
POP und PUSH			KG	A			MI			DI			Aw					K	P	19 ID
Schnitzeljagd		FG	KG	A			-	-	-	-			Aw	An					P	16 SA
Umparken		FG		A			MI			DI			Aw					K	P	7 ID
Verschlüsselung	AG	FG	KG	A			MI	BB		DI			Aw	An				K		22 ID
Wertetausch	AG	FG	KG	A			-	-	-	-			Aw					K	P	25 A
Zahlenreihe				A			MI						Aw						P	17 A
Cluster 2																				
Biber und Bisons	AG		KG			P	-	-	-	-			V		An			K		8 SA
Labyrinth	-		KG	A					SV	DI				Aw	An			K	P	21 A
Morse-Code		FG	KG	A					SV						An				P	27 ID
Netzwerkkabel	AG	FG	KG						SV	DI					An			K		13 ID
Platzwechsel	AG	FG	KG				MI							Aw	An				P	3 A
Verwandlung	AG		KG					BB		DI					An				P	24 ID

Tabelle 9.2: Merkmalsprofile der Cluster 3 und 4. Abstraktionsgrad (AG), Formalisierungsgrad (FG), Komplexitätsgrad (KG), Anforderungsbereich Wiedergabe (W), Anwendung (A), Problemlösung (P), Prozessbereich Modellieren/Implementieren (MI), Begründen/Bewerten (BB), Strukturieren/Vernetzen (SV), Darstellen/Interpretieren (DI), Lernzielstufe Verstehen (V), Anwenden (Aw), Analysieren (An), Wissensart Fakten (F), Konzepte (K), Prozeduren (P). Gegebenenfalls keine Einschätzung (-).

Aufgabe	Gestaltung			Anforderg.			Prozess					Lernziel			Wissen			Rang, Inhalt			
	AG	FG	KG	W	A	P	MI	BB	SV	KK	DI	E	V	Aw	An	B	F		K	P	
Cluster 3																					
Bibers Geheimcode			KG		A				SV					Aw			-	-	-	12	ID
Dateisuche	-	-	KG		A				SV				V	Aw					P	10	SA
Dino-Ordnung			KG		A			BB	SV				V		An				K	29	SA
Wetter					A			BB	SV				V	Aw					K	26	SA
Cluster 4																					
Computervirus				W				BB					V						F	9	MG
Fenster schließen				W						DI		E							F	1	IS
Link				W							KK	E							F	4	IS
Primärschlüssel					A			BB					V						F	20	IS
Private E-Mail				W				BB		KK		E							F	15	IS
Sicheres Passwort					A			BB								B			F	5	MG
Ungeschützter PC				W				BB					V						F	28	IS

9.3 Fähigkeitsmodell A (Antwortmuster)

Im Gegensatz zur Clusteranalyse der Aufgaben nach den kognitiven Merkmalen, die alle Aufgaben gemeinsam erfasste, erfolgte die Faktorenanalyse der Aufgaben nach den Antwortmustern altersgruppenweise. Die klassifizierenden Merkmale hängen zwar nur von den Aufgaben ab (vgl. S. 67f), ungeachtet ihrer Verwendung in den Altersgruppen. Die Faktorenanalyse aber bündelt die Aufgaben nach ihrer Resonanz bei den Schülern, die im Normalfall mit deren Alter variiert. Für jede Jahrgangsstufe wurde ein Satz von zwei oder drei Faktoren geschätzt. Die Ladungsmatrizen und die Faktorzuordnung der Aufgaben befinden sich am Ende des Analysekapitels (Tab. 8.4 bis 8.9, S. 127ff). Für die Interpretation der Faktoren werden die zugeordneten Aufgaben mit ihrem Merkmalsprofil, Schwierigkeitsrang und Themenbereich dargestellt (Tab. 9.3 bis 9.5, S. 152ff).

Jahrgangsstufe 5-7

Die Aufgaben in der jüngsten Altersgruppe ließen sich meist sehr deutlich einem der drei extrahierten Faktoren zuschreiben (Tab. 8.4, 8.5, S. 127). Die Aufgabe *Links um!* (S. 184) weist eine Hauptladung von 0,406 auf Faktor II und eine kaum geringere Nebenladung von 0,381 auf Faktor I auf. Sie wurde beiden Faktoren zugeordnet. Die Aufgabe *Dateisuche* (S. 178) weist als höchste Ladung 0,317 auf den zweiten Faktor auf und wurde diesem zugeordnet, weil sie sich harmonisch in das Merkmalsprofil des Faktors einfügt. Wäre durch diese Aufgabe das Merkmalsprofil deutlich verändert worden, hätte man die niedrigere Ladung auch zum Anlass nehmen können, sie nicht in die Interpretation einzubeziehen.

Der erste Faktor (Tab. 9.3, S. 152) zeigt kein charakteristisches Merkmalsprofil. Er vereint Merkmale, die in der Altersgruppe 5-7 selten vertreten sind: *Platzwechsel* (S. 186) und *Biber am Fluss* (S. 174) sind abstrakt. *Link* (S. 184) gehört zum Prozessbereich Kommunizieren und Kooperieren, *Link und Fenster schließen* (S. 182) sind der Lernzielstufe Erinnern zugeordnet, *Sicheres Passwort* (S. 191) der Stufe Bewerten. *Links um!* (S. 184) gehört zum Anforderungsbereich Anwendung, zum Prozessbereich Modellieren und Implementieren und zur Lernzielstufe Anwenden. Meist ist Faktenwissen gefragt. Ein spezifisches Profil ergibt sich daraus nicht. Auch der Inhaltsbereich ist breit gestreut. Erst der Schwierigkeitsindex offenbart die Gemeinsamkeit der sechs Aufgaben. Es sind diejenigen Aufgaben, die in dieser Altersgruppe von den meisten Schülern erfolgreich bearbeitet werden konnten. Es handelt sich also um einen Schwierigkeitsfaktor. Faktorenanalysen dichotomer Variablen, die nicht gleichverteilt sind, bergen stets die Gefahr der Bildung solcher Schwierigkeitsfaktoren (vgl. S. 120).

Die übrigen beiden Faktoren (Tab. 9.3, S. 152) versammeln Aufgaben des Anforderungsbereichs Anwendung. Erst der Prozessbereich grenzt die Faktoren voneinander ab. Die Aufgaben des zweiten Faktors sprechen den Anforderungsbereich Anwendung an und verlangen Modellieren und Implementieren oder Struk-

turieren und Vernetzen auf der Lernzielstufe Anwenden, meist ist Prozedurwissen gefragt. Dieser Faktor repräsentiert eine Dimension, in der die Anwendung der Konzepte und Prozeduren, die in der Aufgabenstellung angeboten werden, im Vordergrund steht. Mit dem Merkmalsprofil des Clusters Anwendung aus der Anforderungsanalyse verglichen (S. 137), sind die Aufgaben eher konkret und informell und der Prozess des Strukturierens und Vernetzens ist häufiger vertreten.

Neben dem Anforderungsbereich Anwendung und der Lernzielstufe Anwenden kommt in den Aufgaben des dritten Faktors der Prozessbereich Begründen und Bewerten zum Tragen. **Wetter** (S. 195) und **Computervirus** (S. 178) verlangen Begründen und Bewerten auf der Lernzielstufe Verstehen. Die sehr formale Aufgabe **Schnitzeljagd** (S. 190), die die Lernzielstufe Analysieren anspricht, konnte keinem Prozessbereich eindeutig zugeordnet werden. Drei Aufgaben reichen nicht aus, um ein scharfes Bild des Faktors zu gewinnen. Würde man ihn mit den Clustern des Anforderungsmodells vergleichen (S. 137), käme der Faktor aufgrund der Prozessbereiche Strukturieren und Vernetzen, Begründen und Bewerten und der Lernzielstufen Verstehen und Anwenden dem Cluster Verständnis am nächsten.

Zusammengefasst ergaben sich rechnerisch drei Faktoren, darunter ein sogenannter Schwierigkeitsfaktor. Eine Interpretation fällt äußerst schwer, da die Merkmalsprofile eine ungefähre Deutung, aber keine klare Trennung erlauben.

Jahrgangsstufe 8-10

Für die achte bis zehnte Jahrgangsstufe ergaben sich drei Faktoren (Tab. 8.6, 8.7, S. 128). Mit Ladungen von 0,350, 0,363 und 0,382, die geringfügig unterhalb 0,4 liegen, wurden die Aufgaben **Biber und Bisons** (S. 175), **Private E-Mail** (S. 189) und **Bibers Geheimcode** (S. 175) Faktor I zugeordnet. **Schnitzeljagd** (S. 190) wurde mit Ladungen von 0,328 und 0,321 sowohl Faktor I als auch Faktor II zugeordnet. Diese Aufgaben fallen durch Merkmale auf, die insgesamt selten vertreten sind, so **Private E-Mail** mit dem Prozessbereich Kommunizieren und Kooperieren und der Lernzielstufe Erinnern. Versuchsweise wurden diese Aufgaben von der Profilbildung ausgeschlossen. Das Bild änderte sich nicht wesentlich.

Faktor I (Tab. 9.4, S. 153) nimmt neben den stark vertretenen Anwendungsaufgaben auch Aufgaben des Problemlösen auf. Mit fünf bzw. sechs von neun Aufgaben kommen formal und komplex relativ häufig vor. Die Prozessbereiche sind breit gestreut. Die Aufgaben **Biber und Bisons** (S. 175) und **Schnitzeljagd** (S. 190) konnten keinem Prozessbereich zugeordnet werden. **Biber und Bisons** (S. 175) und **Netzwerkkabel** (S. 186) sind die einzigen Aufgaben im Bereich Strukturieren und Vernetzen. Die Lernzielstufe reicht von Erinnern bis Analysieren, meist trifft Anwenden zu. Ebenso kommen alle Wissensarten vor. Mit Blick auf die höher ladenden Aufgaben **Biberzahlen** (S. 176), **Falschgeld** (S. 181), **Primärschlüssel** (S. 188) und **Zahlenreihe** (S. 196), in der zwei Rechenregeln wiederholt auf eine Startzahl angewendet werden, solange bis die Zielzahl Eins erreicht ist, könnte der Faktor als Dimension schematischer Wissensanwendung interpretiert werden.

Faktor II (Tab. 9.4, S. 153) repräsentiert formale und komplexe Aufgaben wie zum Beispiel **Endlosschleife** (S. 180). Alle Aufgaben entstammen dem Anforderungsbereich Anwendung. Die Prozessbereiche sind in den meisten Fällen Darstellen und Interpretieren sowie Modellieren und Implementieren, die Lernzielstufen Anwenden und Analysieren, in erster Linie von Konzeptwissen, aber auch von Prozedurwissen. Der Faktor könnte interpretiert werden als Dimension der Analyse komplexer, formaler Problemstellungen und Umsetzung beziehungsweise Anwendung geeigneter Konzepte und Prozeduren, wie in der Aufgabe **Verschlüsselung** (S. 193), in der ein eindeutiger Code gesucht wird.

Faktor III (Tab. 9.4, S. 153) repräsentiert konkrete, informelle, einfache Aufgaben aus dem Inhaltsbereich Informatik, Mensch und Gesellschaft, die den Prozessbereich Begründen und Bewerten und die Lernzielstufen Verstehen und Bewerten von Faktenwissen ansprechen. Mit nur zwei Aufgaben, **Computervirus** (S. 178) und **Sicheres Passwort** (S. 191), beide zum Thema Benutzerverhalten, ist der Faktor extrem dünn besetzt. Die Aufgabe **Private E-Mail** (S. 189) zum Thema Anwenderwissen wurde mit einer vergleichsweise geringen Hauptladung von 0,363 dem ersten Faktor zugeordnet, weist aber eine deutliche Nebenladung von 0,269 auf dem dritten Faktor auf. Die Berücksichtigung dieser Aufgabe mit den Merkmalen Wiedergabe, Begründen und Bewerten, Kommunizieren und Kooperieren, Erinnern von Faktenwissen würde die Auslegung des Faktors als Wiedergabedimension bestätigen.

Vor allem die Abgrenzung der ersten beiden Faktoren fällt schwer. Auf der Ebene der Aufgabenmerkmale liefern die Prozessbereiche die deutlichste Trennung. Die Prozesskompetenz Strukturieren und Vernetzen, wie in der Graphen-Aufgabe **Netzwerkkabel** (S. 186), ist nur in der ersten Dimension relevant. Die Prozesskompetenz Modellieren und Implementieren, wie in **Umparken** (S. 192), ist nur in der zweiten Dimension relevant. Auf der Ebene der Personeneigenschaften, die allerdings nicht in die Berechnung der Faktoren einbezogen wurden, sind in der ersten Dimension, die die detailgetreue Befolgung eines vorgegebenen Schemas erfordert, Genauigkeit und Ausdauer wichtig. In der zweiten Dimension ist es wichtig, aufgeschlossen gegenüber unvertrauten Notationen zu sein, wie dem Flussdiagramm in **Endlosschleife** (S. 180), dem Baumdiagramm in **Binärbaum** (S. 177) oder der prozeduralen Schreibweise in **Umparken** (S. 192). Wie schon in der unteren Jahrgangsstufe lässt diese Charakterisierung keine separaten, gar unabhängigen Dimensionen erkennen.

Jahrgangsstufe 11+

Für die Aufgaben der Schüler ab der elften Klasse wurden nur zwei Faktoren extrahiert (Tab. 8.8, 8.9, S. 129). Die Aufgabe **Verschlüsselung** (S. 193) ist dabei mit keinem der beiden Faktoren ausreichend korreliert. Die Aufgabe **Verwandlung** (S. 194) trägt mit Ladungen von 0,386 und 0,399 zur Spezifikation beider Dimensionen bei.

Das Profil des ersten Faktors (Tab. 9.5, S. 154) zeigt komplexe, formale Aufgaben, überwiegend aus dem Anforderungsbereich Anwendung. Die Prozessbereiche sind auffallend dünn besetzt. Für die Aufgaben **Wertetausch** (S. 195), **Schnitzeljagd** (S. 190), **Biber und Bisons** (S. 175) wurde in der Expertenbefragung zu den Prozessbereichen keine ausreichende Übereinstimmung erzielt. Anwenden und Analysieren sind als Lernzielstufen gleich häufig vertreten, vier der acht Aufgaben erfordern Anwenden, vier erfordern Analysieren, nur zwei erfordern Verstehen. Prozedurwissen ist die überwiegende Wissensart. Kodierung wie in **Morse-Code** (S. 185) und Logisches Schließen wie in **Biber und Bisons** (S. 175) sind typische Aufgabeninhalte.

Der zweite Faktor (Tab. 9.5, S. 154) wird durch weniger formale, aber durchaus komplexe Aufgaben aus den Anforderungsbereichen Anwendung und Problemlösung repräsentiert, etwa **Labyrinth** (S. 183). Der Prozessbereich ist durchgängig Darstellen und Interpretieren, lediglich die Aufgabe **Computervirus** (S. 178), die auch die einzige Aufgabe des Anforderungsbereichs Wiedergabe ist, wird stattdessen Begründen und Bewerten zugeordnet. Modellieren und Implementieren kommt allein in diesem Faktor vor, bei **POP und PUSH** (S. 187) und **Endlosschleife** (S. 180). Die häufigste Lernzielstufe ist Analysieren, vor Anwenden, meist geht es um Konzept- oder Prozedurwissen. Die Aufgabeninhalte sind Algorithmen und Datenstrukturen, insbesondere Graphen wie in **Verwandlung** (S. 194) und **Netzwerkkabel** (S. 186). Der Faktor zeigt deutliche Ähnlichkeit mit dem Cluster Problemlösung aus der Analyse der Aufgabenanforderungen (S. 137).

Auch in der oberen Jahrgangsstufe bleibt die Spezifikation unscharf. Gleicht man etwa die Aufgabe **Verschlüsselung** (S. 193) mit den Beschreibungen ab, entspricht sie beiden gut. In allen drei Altersstufen fällt es schwer, die Faktoren anhand der Merkmale prägnant zu beschreiben. Die Charakterisierung genügt dem Anspruch, die Dimensionen zusammenfassend zu beschreiben, trennt sie jedoch nicht deutlich genug, um jede Aufgabe zweifelsfrei einzuordnen. Insgesamt ist das Resultat der Faktorenanalyse der Aufgaben unbefriedigend. Mehrere Gründe sind denkbar.

- (a) Die Faktorenanalyse ist für dichotome und nicht gleichverteilte Variablen nicht geeignet. Bekanntermaßen neigt sie bei dieser Konstellation zur Bildung von Schwierigkeitsfaktoren, wie in der fünften bis siebten Jahrgangsstufe geschehen. Mit dieser Methode gelingt es nicht, die existierenden Dimensionen sichtbar zu machen.
- (b) Die Qualität des theoretischen oder empirischen Materials, der Merkmale und der Aufgaben, reicht für die Analyse nicht aus. Die Fähigkeitsstruktur, die aus den Merkmalen modelliert wird, kann nur die Kompetenzfacetten abbilden, die durch die Merkmale zum Ausdruck gebracht und bei der Aufgabenbearbeitung gefordert werden. Kurz gesagt kann das Modell höchstens so treffend sein wie die Aufgabenmerkmale.

- (c) Die Zahl der Aufgaben, die die Dimensionen belegen, ist für die Interpretation zu gering, so dass die modellimmanente Struktur sich nicht abzeichnet.
- (d) Das Konstrukt, informatische Anforderungen und Fähigkeiten, weist die vermutete mehrdimensionale Struktur nicht auf.

Ob die unscharfe Spezifikation der Faktoren auf Schwächen der Analyse (a), des theoretischen oder empirischen Materials (b), der Interpretation (c) oder der Prämisse (d) zurückgeht, kann nicht geklärt werden. Da eine dimensionale Struktur der Personenfähigkeiten schwach erkennbar ist – wenn auch nicht annähernd so deutlich wie die Anforderungsstruktur – wird das Modell als Alternative A eines Fähigkeitsmodells fortgeführt. In Kapitel 10 wird das Fähigkeitsmodell A einer Gültigkeitsprüfung im Vergleich zur Modellalternative B und zum Anforderungsmodell unterzogen.

9.4 Fähigkeitsmodell B (Merkmalsquoten)

Im Fähigkeitsmodell A charakterisieren *Aufgabenbündel* die Faktoren, die als Dimensionen interpretiert werden. Um die Dimensionen in den Begriffen der Fähigkeitsmerkmale zu beschreiben, werden die Anforderungsprofile der Aufgaben herangezogen. Für das Fähigkeitsmodell B wird ein alternativer Weg beschritten, der die Aufgaben ausblendet und direkt über die Personen-Merkmalsquoten (wie viel Prozent der Aufgaben, die das Merkmal tragen, löst die Person?) zu Dimensionen führt. Die Merkmalsquoten werden als Personenvariablen faktorenanalytisch ausgewertet und es ergeben sich *Merkmalsbündel*, die für die Dimensionen stehen, vier in der unteren Altersgruppe und jeweils drei in der mittleren und oberen Altersgruppe. Anders als bisher bezieht sich die Interpretation unmittelbar auf die Ladungsmatrizen und die Faktorzuordnung (Tab. 8.10 bis 8.15, S. 130ff).

Jahrgangsstufe 5-7

In der unteren Stufe werden vier Faktoren berechnet (Tab. 8.10, 8.11, S. 130). Der erste weist hohe Ladungen der Prozessmerkmale Darstellen und Interpretieren (0,880), Modellieren und Implementieren (0,758) und der Wissensarten Prozeduren (0,711) und Konzepte (0,683) auf. Weiter sind mit Hauptladungen über 0,6 der Anforderungsbereich Anwendung und das Lernziel Anwenden sowie die formale Gestaltung wichtig.

Der zweite Faktor ist hochgradig mit der Prozesskompetenz Strukturieren und Vernetzen (0,921) und der Lernzielstufe Verstehen (0,808) sowie dem Komplexitätsgrad (0,724) verbunden. Mit relativ hohen Nebenladungen um 0,6 tragen auch der Anforderungsbereich Anwendung und das Lernziel Anwenden sowie das Prozessmerkmal Begründen und Bewerten zum Faktorprofil bei.

Die Faktoren I und II ähneln sich darin, dass die Anforderung Anwendung und das Lernziel Anwenden sie prägen. Zur Abgrenzung legt vor allem der starke Zusammenhang des ersten Faktors mit Modellieren, Implementieren und Prozedurwissen nahe, ihn als *Anwendungskomponente* bei der Aufgabenbearbeitung zu interpretieren. Der starke Zusammenhang des zweiten Faktors mit Strukturieren, Vernetzen, Verstehen und dem Komplexitätsgrad legt nahe, ihn als *Verständnis-komponente* bei der Aufgabenbearbeitung zu interpretieren.

Der dritte Faktor wird durch einen sehr starken Zusammenhang mit dem Anforderungsbereich *Problemlösung* (0,922), dem Abstraktionsgrad (0,777) und der Lernzielstufe Analysieren (0,693) charakterisiert. Nebenladungen von 0,587 und 0,531 bedeuten, dass auch der Formalisierungsgrad und der Komplexitätsgrad in die Fähigkeitsdimension III hineinspielen.

Dem vierten Faktor werden der Anforderungsbereich *Wiedergabe* (0,896) und die Wissensart Fakten mit sehr hoher Ladung (0,930) und der Prozessbereich Begründen und Bewerten (0,669) zugeordnet. Relevante Nebenladungen treten nicht auf, bis auf die Lernzielstufe Verstehen, die hoch auf Faktor II lädt (0,808) und eine Nebenladung von 0,443 auf Faktor IV aufweist. Das heißt, im Gegensatz zu den übrigen Faktoren bildet Faktor IV eine Dimension ohne Überschneidungen mit den Merkmalen der anderen Dimensionen.

Jahrgangsstufe 8-10

In der mittleren Jahrgangsstufe werden drei Faktoren extrahiert (Tab. 8.12, 8.13, S. 131). Der erste Faktor steht mit sehr hoher Korrelation für die Lernzielstufe Anwenden (0,934) und die Anforderung *Anwendung* (0,889), hauptsächlich von Prozedurwissen (0,896), aber auch von Konzeptwissen (0,653), um formale und komplexe Aufgabenstellungen zu bearbeiten (0,843 und 0,766). Unter den Prozessbereichen korrelieren Darstellen und Interpretieren, Modellieren und Implementieren hoch mit Faktor I (0,842 und 0,835). Die Nebenladung von 0,614 spricht dafür, dass auch der Abstraktionsgrad von Bedeutung ist.

Der zweite Faktor wird in hohem Maß durch den Anforderungsbereich *Problemlösung* charakterisiert (0,935), weiterhin durch den Prozessbereich Strukturieren und Vernetzen (0,702) und den Abstraktionsgrad (0,653). Auch das Lernziel Analysieren, die Wissensart Konzepte und der Komplexitätsgrad tragen mit Nebenladungen um 0,6 zur Beschreibung des Faktors II bei.

Der dritte Faktor vereint mit sehr hoher Ladung die Wissensart Fakten (0,956) mit der Anforderung *Wiedergabe* (0,852), dem Prozess des Begründens und Bewertens (0,892) und dem Lernziel Verstehen (0,699). Wie auch in der Jahrgangsstufe fünf bis sieben sind keine Nebenladungen weiterer Merkmale zu beobachten. Die Dimension der Wiedergabe, des Verstehens, Begründens und Bewertens von Faktenwissen ist damit schärfer abgegrenzt als die übrigen Dimensionen.

Jahrgangsstufe 11+

In der oberen Jahrgangsstufe ergeben sich drei Faktoren (Tab. 8.14, 8.15, S. 132). Der erste Faktor wird durch die große Zahl von zehn Merkmalen gebildet, meist mit hoher Hauptladung. Am höchsten laden die Anforderung *Anwendung* und die Lernzielstufe Anwenden (0,932 und 0,868), dicht gefolgt von den Prozessbereichen Modellieren und Implementieren, Darstellen und Interpretieren. Die drei Gestaltungsmerkmale Abstraktionsgrad, Formalisierungsgrad und Komplexitätsgrad werden dem Faktor zugeordnet, ebenso die Wissensarten Konzepte und Prozeduren. Schließlich trägt die Lernzielstufe Analysieren zum Profil des Faktors bei. Dabei überschneiden sich die Profile von Faktor I und III, denn Analysieren lädt mit einer Hauptladung von 0,667 auf Faktor I und mit einer kaum geringeren Nebenladung von 0,640 auf Faktor III.

Der zweite Faktor ist besonders eng mit der Wissensart Fakten (0,941), dem Anforderungsbereich *Wiedergabe*, der Lernzielstufe Verstehen und dem Prozessbereich Begründen und Bewerten verbunden. Weitere Merkmale laden nicht bemerkenswert hoch auf den Faktor.

Der dritte Faktor schließlich repräsentiert zuvorderst den Anforderungsbereich *Problemlösung* (0,874) und den Prozessbereich Strukturieren und Vernetzen (0,898). Durch eine vergleichsweise hohe Nebenladung von 0,640 ist er auch mit der Lernzielstufe Analysieren verknüpft.

Gemeinsamer Modellansatz

In der Zusammenschau der drei Altersgruppen ist ein gemeinsamer Modellkern erkennbar. Das Fähigkeitsgerüst jeder Altersgruppe enthält mindestens die drei Dimensionen Wiedergabe, Anwendung und Problemlösung (Abbildung 9.3).

Stets tritt ein Faktor auf, dem der Anforderungsbereich Wiedergabe und die Wissensart Fakten mit hoher Ladung sowie der Prozessbereich Begründen und Bewerten und die Lernzielstufe Verstehen zugeordnet werden. Das ist Faktor IV in Stufe 5-7, Faktor III in Stufe 8-10 und Faktor II in Stufe 11+. Der Faktor repräsentiert die *Wiedergabedimension*, das Vermögen, Wissen zu verstehen und wiederzugeben, nicht nur als bloße Wiederholung, sondern auch innerhalb von Begründungen und Bewertungen.

In allen Jahrgangsstufen weisen die Analysen einen starken ersten Faktor aus, auf den viele Variablen hoch laden, die Anforderung Anwendung, die Prozesse Modellieren und Implementieren, Darstellen und Interpretieren, die Lernzielstufe Anwenden, die Wissensarten Konzepte und Prozeduren, der Formalisierungsgrad. Dieser Faktor wird als *Anwendungsdimension* verstanden. Wesentlicher Bestandteil ist nicht nur die prozedurale Fertigkeit des Modellierens und Implementierens, sondern auch die Fähigkeit, Konzepte zu erfassen und auszulegen. Dazu gehört Formalisieren und mit formalen Darstellungen umgehen. Ab Stufe 8-10 gewinnt die Fähigkeit, Abstraktion und Komplexität zu bewältigen, an Gewicht.

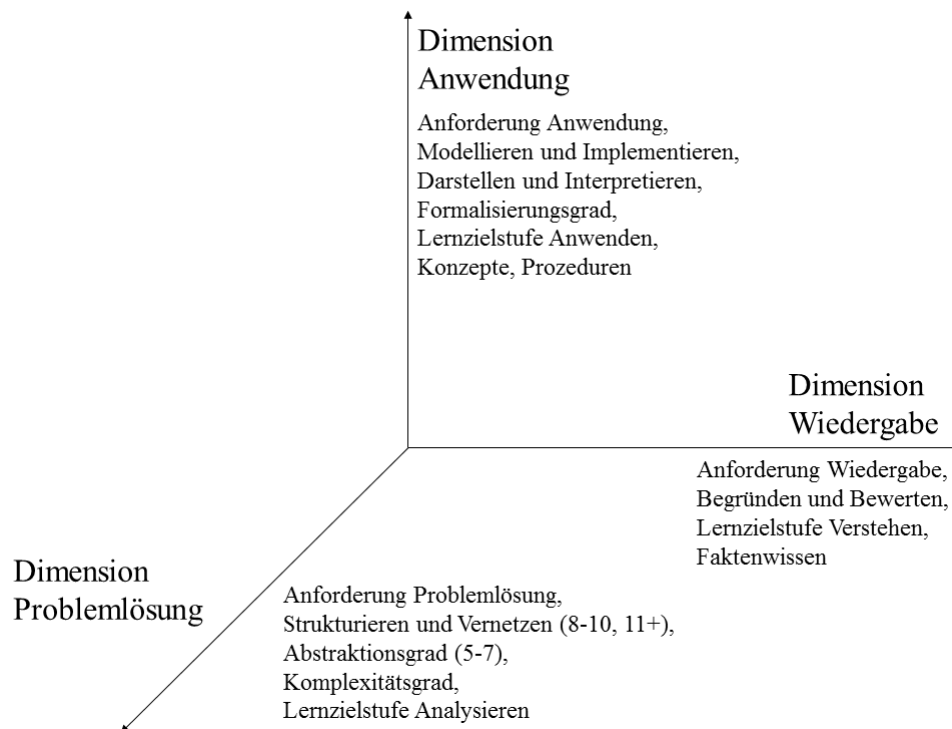


Abbildung 9.3: Fähigkeitsmodell B als dreidimensionales Strukturmodell

Einem weiteren Faktor, der allen Stufen gemeinsam ist, sind der Anforderungsbereich Problemlösung, die Lernzielstufe Analysieren und der Komplexitätsgrad zugeordnet. Es handelt sich um Faktor III in Stufe 5-7, Faktor II in Stufe 8-10 und Faktor III in Stufe 11+. Der Faktor wird als *Problemlösungsdimension* interpretiert und repräsentiert Analyse- und Problemlösekompetenzen. Der Umgang mit Komplexität spielt eine wichtige Rolle. Insgesamt stellt sich dieser Faktor weniger homogen dar als die Faktoren Wiedergabe und Anwendung. So lädt in den höheren Jahrgangsstufen, acht bis zehn und ab elf, der Prozessbereich Strukturieren und Vernetzen auf den Problemlösefaktor, während in der Jahrgangsstufe fünf bis sieben Strukturieren und Vernetzen eher im Verständniserwerb zum Tragen kommen. Dafür prägt in der jüngsten Altersgruppe besonders das Abstraktionsvermögen die Problemlösungsdimension.

Nur in der in der jüngsten Altersgruppe tritt ein vierter Faktor auf, Faktor II, der hoch mit dem Komplexitätsgrad, dem Prozessbereich Strukturieren und Vernetzen und der Lernzielstufe Verstehen korreliert. Der Faktor wird als *Verständnisdimension* interpretiert, die für die spezifische Fähigkeit steht, Anwendungswissen zu erwerben, zu strukturieren und zu vernetzen. Zumindest anhand der Wettbewerbsdaten lässt sich die Verständnisdimension in der mittleren und oberen Altersgruppe nicht mehr identifizieren.

9.5 Zusammenfassung

Mit verschiedenen statistischen Ansätzen wurden im vergangenen Analysekapitel Strukturen aus den merkmalsklassifizierten Aufgaben und den Schülerantworten herausgearbeitet. Das rechnerische Ergebnis besteht in den Zuordnungstabellen von Aufgaben oder Merkmalen zu Clustern beziehungsweise Faktoren. Unter der Prämisse, dass die Struktur informatischer Anforderungen und Fähigkeiten als System möglichst unabhängiger Dimensionen darstellbar ist, werden die Cluster und Faktoren als Modelldimensionen interpretiert. Anhand der zugeordneten Aufgaben, ihrer Schwierigkeit und ihrer kognitiven Merkmale und Inhalte werden Profile der Cluster und Faktoren erstellt, die die Dimensionen charakterisieren.

- Die Clusteranalyse der Aufgaben nach den Merkmalen führt zu vier Aufgabentypen, die als Repräsentanten der Dimensionen Wiedergabe, Verständnis, Anwendung, Problemlösung interpretiert werden. Ein vierdimensionales Anforderungsmodell entsteht, das den kognitiven Anspruch der Aufgaben beschreibt.
- Die Faktorenanalyse der Aufgaben nach den Antwortmustern ergibt je nach Altersstufe zwei bis vier Faktoren eines Fähigkeitsmodells, Modell A genannt. Da sie sich nicht deutlich genug unterscheiden, gelingt die Interpretation anhand der Merkmalsprofile nicht.
- Die Faktorenanalyse der Merkmale nach den Lösungsquoten berechnet drei bis vier Faktoren eines Fähigkeitsmodells, Modell B genannt. Die Profile der vier Faktoren in der Jahrgangsstufe fünf bis sieben decken sich größtenteils mit den Clusterprofilen des Anforderungsmodells. Entsprechend werden die Dimensionen für diese Stufe Wiedergabe, Verständnis, Anwendung und Problemlösung genannt. In den Stufen acht bis zehn und ab elf werden die drei Dimensionen übereinstimmend als Wiedergabe, Anwendung und Problemlösung charakterisiert (vgl. Abb. 9.3, S. 149). Die Verständnisdimension entfällt in diesen Altersgruppen.

In allen Analysen leisten die Anforderungsbereiche den größten Beitrag zur Definition der Dimensionen, indem sie jeweils auf einen eigenen Faktor laden. Sie sind daher namensgebend für die Dimensionen. So bildet sich ein Grundgerüst heraus, vierdimensional im Falle des Anforderungsmodells und zwei-, drei- oder vierdimensional im Falle der Fähigkeitsmodelle A und B. Das Grundgerüst, bestehend aus den Dimensionen Wiedergabe, eventuell Verständnis, Anwendung und Problemlösung, wird nach den übrigen charakterisierenden Merkmalen ausdifferenziert.

Der Vorbehalt der Variablenkonfundierung scheint sich zu bestätigen (zu den konfundierten Effekten vgl. S. 104). Ein Vergleich der Merkmalspaare, die bereits in den Aufgaben korreliert sind (S. 122), mit den Faktorenspezifikationen zeigt,

dass die Faktorenanalyse genau diese Merkmale bündelt. Bestenfalls liegt das daran, dass die Aufgabenkonstrukteure bewusst oder unbewusst die Merkmalsstruktur antizipieren. Wahrscheinlicher ist, dass die Struktur der kognitiven Aufgabenanforderungen die Struktur der kognitiven Schülerfähigkeiten überlagert.

Das Ergebnis der Interpretation ist eine Modellvorstellung informatischer Anforderungen und Fähigkeiten, wie sie durch den Schülerwettbewerb Informatik-Biber aktiviert werden. Die Analyse der merkmalsklassifizierten Aufgaben liefert die Grundlage für ein vierdimensionales Anforderungsmodell, das mit theoretischen Kompetenzannahmen harmoniert. Die Analyse der Schülerantworten liefert die Grundlage für ein drei- bis vierdimensionales Fähigkeitsmodell. Dieses lässt sich ebenfalls mit theoretischen Kompetenzvorstellungen zur Deckung bringen. Es ist allerdings mit dem Verdacht behaftet, nicht die realisierte, sondern die erwartete Schülerfähigkeit abzubilden, da die Merkmalsvariablen konfundiert sind.

Tabelle 9.3: Merkmalsprofile der Faktoren in Jahrgangsstufe 5-7. Abstraktionsgrad (AG), Formalisierungsgrad (FG), Komplexitätsgrad (KG), Anforderungsbereich Wiedergabe (W), Anwendung (A), Problemlösung (P), Prozessbereich Modellieren/Implementieren (MI), Begründen/Bewerten (BB), Strukturieren/Vernetzen (SV), Darstellen/Interpretieren (DI), Lernzielstufe Verstehen (V), Anwenden (Aw), Analysieren (An), Wissensart Fakten (F), Konzepte (K), Prozeduren (P). Gegebenenfalls keine Einschätzung (-).

Aufgabe	Gestaltung			Anforderg.			Prozess			Lernziel			Wissen			Rang, Inhalt				
	AG	FG	KG	W	A	P	MI	BB	SV	KK	DI	E	V	Aw	An		B	F	K	P
Faktor I																				
Fenster schließen				W							DI	E					F			1 IS
Biber am Fluss	AG				A						DI				An			K	P	2 ID
Link				W						KK		E					F			4 IS
Sicheres Passwort					A			BB								B		F		5 MG
Platzwechsel	AG	FG	KG			P	MI							Aw	An				P	3 A
Links um!					A		MI							Aw					-	6 A
Faktor II																				
Dino-Ordnung			KG		A			BB	SV				V		An			K		29 SA
Bibers Geheimcode			KG		A				SV					Aw				-	-	12 ID
Zahlenreihe					A		MI							Aw					P	17 A
Umparken		FG			A		MI				DI			Aw				K	P	7 ID
Links um!					A		MI							Aw				-	-	6 A
Dateisuche	-	-	KG		A				SV				V	Aw					P	10 SA
Faktor III																				
Wetter					A			BB	SV				V	Aw				K		26 SA
Schnitzeljagd		FG	KG		A		-	-	-	-	-			Aw	An				P	16 SA
Computervirus				W				BB					V				F			9 MG

Tabelle 9.4: Merkmalsprofile der Faktoren in Jahrgangsstufe 8-10. Abstraktionsgrad (AG), Formalisierungsgrad (FG), Komplexitätsgrad (KG), Anforderungsbereich Wiedergabe (W), Anwendung (A), Problemlösung (P), Prozessbereich Modellieren/Implementieren (MI), Begründen/Bewerten (BB), Strukturieren/Vernetzen (SV), Darstellen/Interpretieren (DI), Lernzielstufe Verstehen (V), Anwenden (Aw), Analysieren (An), Wissensart Fakten (F), Konzepte (K), Prozeduren (P). Gegebenenfalls keine Einschätzung (-).

Aufgabe	Gestaltung			Anforderg.			Prozess					Lernziel			Wissen			Rang, Inhalt		
	AG	FG	KG	W	A	P	MI	BB	SV	KK	DI	E	V	Aw	An	B	F		K	P
Faktor I																				
Falschgeld	FG	KG		A						DI			Aw						P	18 A
Primärschlüssel	FG			A				BB				V					F	K		20 IS
Zahlenreihe				A			MI						Aw						P	17 A
Biberzahlen	AG	FG	KG	A						DI			Aw						P	23 ID
Netzwerkkabel	AG	FG	KG			P			SV	DI				An				K		13 ID
Bibers Geheimcode			KG	A					SV				Aw					-	-	12 ID
Private E-Mail					W			BB	KK		E						F			15 IS
Biber und Bisons	AG		KG			P	-	-	-	-		V		An				K		8 SA
Schnitzeljagd	FG	KG		A			-	-	-	-			Aw	An					P	16 SA
Faktor II																				
Verschlüsselung	AG	FG	KG	A			MI	BB		DI			Aw	An				K		22 ID
Binärbaum		FG	KG	A						DI			Aw	An				K		11 ID
Umparken		FG		A			MI			DI			Aw					K	P	7 ID
Endlosschleife	AG	FG	KG	A			MI			DI			Aw	An				K	P	14 A
Schnitzeljagd		FG	KG	A			-	-	-	-			Aw	An					P	16 SA
Faktor III																				
Sicheres Passwort				A				BB							B		F			5 MG
Computervirus				W				BB				V					F			9 MG

Tabelle 9.5: Merkmalsprofile der Faktoren in Jahrgangsstufe 11+. Abstraktionsgrad (AG), Formalisierungsgrad (FG), Komplexitätsgrad (KG), Anforderungsbereich Wiedergabe (W), Anwendung (A), Problemlösung (P), Prozessbereich Modellieren/Implementieren (MI), Begründen/Bewerten (BB), Strukturieren/Vernetzen (SV), Darstellen/Interpretieren (DI), Lernzielstufe Verstehen (V), Anwenden (Aw), Analysieren (An), Wissensart Fakten (F), Konzepte (K), Prozeduren (P). Gegebenenfalls keine Einschätzung (-).

Aufgabe	Gestaltung			Anforderg.			Prozess					Lernziel			Wissen			Rang, Inhalt				
	AG	FG	KG	W	A	P	MI	BB	SV	KK	DI	E	V	Aw	An	B	F		K	P		
Faktor I																						
Schnitzeljagd		FG	KG		A		-	-	-	-	-			Aw	An				P	16	SA	
Biberrzahlen	AG	FG	KG		A					DI				Aw					P	23	ID	
Morse-Code		FG	KG		A	P			SV						An				P	27	ID	
Falschgeld		FG	KG		A					DI				Aw					P	18	A	
Wertetausch	AG	FG	KG		A		-	-	-	-	-			Aw				K	P	25	A	
Verwandlung	AG		KG			P		BB		DI					An				P	24	ID	
Biber und Bisons	AG		KG			P	-	-	-	-	-	V			An					8	SA	
Primärschlüssel		FG			A			BB				V					F	K		20	IS	
Faktor II																						
Labyrinth	-		KG		A	P			SV		DI			Aw	An				K	P	21	A
Endlosschleife	AG	FG	KG		A			MI		DI				Aw	An				K	P	14	A
POP und PUSH			KG		A			MI		DI				Aw					K	P	19	ID
Verwandlung	AG		KG			P		BB		DI					An				P	24	ID	
Netzwerkkabel	AG	FG	KG			P			SV	DI					An				K		13	ID
Computervirus				W				BB				V					F				9	MG
-																						
Verschlüsselung	AG	FG	KG		A			MI	BB		DI			Aw	An				K		22	ID

Kapitel 10

Modellgeltung

Bis zu diesem Punkt war das Vorgehen ausschließlich exploratorisch. Die Aufgaben wurden von Experten nach ihren kognitiven Anforderungen klassifiziert und clusteranalytisch untersucht, um Anforderungsdimensionen offenzulegen. Die Antworten wurden mit explorativen Faktorenanalysen untersucht, um Fähigkeitsdimensionen aufzuspüren. Zuletzt erfolgte die Interpretation der resultierenden Strukturmodelle. In diesem Kapitel kommt zur Modellgeltungsprüfung erstmals ein konfirmatorisches Verfahren zum Einsatz. Die Strukturprüfung schließt als sechster Vorgehensschritt die Modellkonstruktion ab.

Drei Modellvarianten liegen vor, das vierdimensionale Anforderungsmodell aus der Clusteranalyse der Aufgaben sowie die Fähigkeitsmodelle A und B aus den Faktorenanalysen der Antwortmuster und der Merkmalsquoten. Mit dem multidimensionalen Rasch-Testmodell MultiRa wird die Geltung der Modellvarianten im Sinne ihrer Passung zu den Daten empirisch überprüft (Abschnitt 10.1). Das Hauptgütekriterium bleibt jedoch der Erklärungswert der Modelle, das heißt ob sie theoretisch begründbar, interpretierbar und plausibel sind und weiter, ob sie die Daten im Sinne der Ausgangsfrage erklären: „Welche Dimensionen weisen informatikspezifische Aufgabenanforderungen und Schülerfähigkeiten auf?“ Abschließend werden die wichtigsten Kritikpunkte an der Datengrundlage und der Modellbildung diskutiert (Abschnitt 10.2).

10.1 Strukturprüfung

Zur Beurteilung der Modellgüte wird die Datenpassung betrachtet, das heißt, die Geltung der Modelle für die Stichprobendaten wird vergleichend untersucht. Dazu wird das strukturprüfende MultiRa-Modell verwendet, das auf einer multidimensionalen Erweiterung des Rasch-Modells basiert (Carstensen 2000). Das MultiRa-Modell gestattet, eine mehrdimensionale Strukturhypothese, die sich durch die Zuordnung der Testaufgaben zu den vermuteten Dimensionen ausdrücken lässt, auf ihre Passung zu einem Testergebnis zu überprüfen.

Allgemein begründet die Testtheorie die Konstruktion von Testinstrumenten, z. B. Fragebögen, die psychische Merkmale erfassen. Dazu werden Skalen wie der Intelligenzquotient IQ in Intelligenztests konstruiert. Die *klassische* Testtheorie fragt nach dem wahren Wert einer Personeneigenschaft, wenn ein messfehlerbehaftetes Testergebnis vorliegt. Die *probabilistische* Testtheorie dagegen formalisiert die Wahrscheinlichkeit, mit der eine Person ein Testitem löst, wenn Personenfähigkeit und Itemschwierigkeit bekannt sind. Die bekanntesten Modelle der probabilistischen Testtheorie sind die Item-Response-Modelle von Rasch (Rost 2004a), die auch der PISA-Studie zugrundeliegen. Auf dieselben Daten angewendet, liefern klassische und probabilistische Ansätze ähnliche Ergebnisse. Im Lehrbuch zur Testtheorie und Testkonstruktion von Rost werden die Theorien als „komplementäre, nicht als konkurrierende Theorien behandelt“ (ebd., S. 9).

Bei den Wettbewerbsaufgaben handelt es sich nicht um Testaufgaben, die testtheoretische Ansprüche erfüllen und zu diagnostischen Zwecken konstruiert wurden. Die Intention ist vielmehr, das Interesse an Informatikinhalten zu wecken und ein attraktives Bild der informatischen Themenvielfalt zu zeichnen (Pohl et al. 2009, S. 39). Was die Wettbewerbsdaten für diese Untersuchung auszeichnet, ist die Bandbreite der Informatikinhalte und die Größe der Stichprobe, die Schülerinnen und Schüler aller Jahrgangsstufen der Sekundarschulen umfasst.

Die Studie zielt nicht darauf ab, individuelle Schülerleistungen zu beziffern, wie es in der Testtheorie geschieht. Vielmehr werden charakterisierende Aufgabenmerkmale und die dimensionale Struktur, die sich in den Wettbewerbsdaten abzeichnet, herausgearbeitet. Die Kenntnis der Schwierigkeitsmerkmale und Anforderungsdimensionen schafft die Grundlage der fundierten Aufgabenkonstruktion zur Kompetenzmessung. So gesehen leistet die Studie mit klassischen *struktur-entdeckenden* Methoden die modellbildende Vorarbeit zur Testkonstruktion. Die Motivation für eine Verknüpfung mit dem probabilistischen Rasch-Modell ist, die Passung der Modellvarianten zu den Daten empirisch untersuchen zu können. Der letzte Schritt ist also *strukturprüfend*.

Das ursprüngliche, dichotome logistische Modell von Rasch definiert die Lösewahrscheinlichkeit für ein dichotomes Testitem durch eine logistische Funktion, abhängig von der Personenfähigkeit und der Itemschwierigkeit (Carstensen 2000, S. 48ff). Dieses Modell ist jedoch sehr restriktiv, vor allem was die Gültigkeitsbedingungen betrifft. Üblicherweise entstehen raschkonforme Tests erst durch eine strenge Selektion homogener Items aus einem größeren Item-Grundschatz. Es ist nicht anzunehmen, dass die Aufgaben mit einem eindimensionalen Rasch-Modell konform sind. Würden sie dessen Voraussetzung der Item-Homogenität erfüllen, nach der alle Items die gleiche Schwierigkeit aufweisen, würden sich die Dimensionen aufspürenden Analysen erübrigen, denn dann wäre der Informationsgehalt der Antwortmuster mit der Anzahl gelöster Aufgaben ausgeschöpft. Welche Aufgaben mit welchen Merkmalen eine Person löst, würde nicht zur Differenzierung der Fähigkeit beitragen, die durch den Wettbewerb erfasst wird. Gerade die Differenzierung nach Merkmalen ist aber das Thema dieser Arbeit.

Aus dem strengen eindimensionalen Rasch-Modell wird das mehrdimensionale Itemkomponenten-Rasch-Modell MultiRa abgeleitet, das weniger restriktiv ist (Carstensen 2000, S. 68ff). Es formalisiert die Annahme, dass mehrere Fähigkeitskomponenten sich zur Erfolgswahrscheinlichkeit addieren. Das stimmt mit der Modellvorstellung mehrerer Dimensionen der Aufgabenschwierigkeit beziehungsweise der Personenfähigkeit überein. Das multidimensionale Rasch-Modell wird hier verwendet, um die Modellgeltung der cluster- und faktorenanalytisch gewonnenen Modellvarianten zu überprüfen.

10.1.1 MultiRa-Analyse

Für die MultiRa-Analyse (Carstensen und Rost 2000) wird eine mehrdimensionale Strukturhypothese durch einen Satz Items und eine sogenannte Designmatrix spezifiziert, die die Items den postulierten Dimensionen zuordnet. Ein Item kann mehreren Dimensionen zugeordnet sein. In Tabelle 10.1 bis 10.3 (S. 159ff) sind die Designmatrizen der Modellvarianten dargestellt. Die Matrizen enthalten Werte 1, wenn die Aufgabe in dieser Zeile der Dimension in dieser Spalte angehört, oder 0, wenn sie ihr nicht angehört. Je Modellvariante und je Altersgruppe werden von der Designmatrix und den binär kodierten Antwortmustern der Stichprobe ausgehend die Parameter eines multidimensionalen Rasch-Modells nach dem Maximum-Likelihood-Prinzip geschätzt. Mittels Bootstrapping wird die Modellanpassung an die Daten überprüft.

Das Maximum-Likelihood-Verfahren maximiert eine Likelihoodfunktion, die die Wahrscheinlichkeit, dass die Antwortmuster der Stichprobe zustande kommen, als Produkt von Antwortwahrscheinlichkeiten berechnet. Die Antwortwahrscheinlichkeit einer Person auf eine Frage hängt davon ab, mit welchem Gewicht die Frage zu einer Dimension beiträgt, wie fähig die Person in der Dimension ist und wie schwierig die Frage ist. Die Gewichte ergeben sich aus den Einträgen in der Designmatrix, einfachsten Falles 1 oder 0. Wie fähig eine Person in einer Dimension ist, beschreibt der Personenparameter, wie schwierig eine Frage ist, beschreibt der Itemparameter. Für die Parameter werden iterativ Schätzwerte bestimmt, die die Likelihoodfunktion maximieren. Die Suche wird durch die Vorgabe einer Iterationenzahl beschränkt oder abgebrochen, sobald die Likelihood sich in einer Iteration nur noch geringfügig verbessert (Carstensen 2000, S. 95ff).

Zur Kontrolle der Datenanpassung werden nach der Bootstrap-Methode anhand der geschätzten Parameter Stichproben simuliert. Je Stichprobe werden die Häufigkeiten aller Antwortmuster (Pattern) erfasst und eine Chi-Quadrat-Statistik berechnet, die die Abweichungen der *simulierten* Patternhäufigkeiten von den unter Modellgeltung zu erwartenden Patternhäufigkeiten aggregiert. So erhält man eine empirische Näherung der Häufigkeitsverteilung der Abweichungstatistik. Liegt der Chi-Quadrat-Wert der *echten* Daten am Rand der Verteilung, ist dieser Wert unter den modellkonform simulierten Daten unwahrscheinlich. Das widerspricht der Annahme, Modell und Daten passen zusammen (ebd., S. 188ff).

Der MultiRa-Algorithmus wurde von Carstensen und Rost (2000) implementiert und als Software MULTIRA verfügbar gemacht. MULTIRA wird hier mit folgenden Modelleinstellungen zur Prüfung der Modellgeltung ausgeführt. Die Itemparameter werden per Conditional Maximum Likelihood-Schätzung bestimmt (Carstensen 2000, S. 96f). Die Personenparameter werden durch Weighted Likelihood-Schätzer nach Warn berechnet (Carstensen 2000, S. 103). Für den Bootstrap wird 100 als Zahl der Resimulationen eingestellt.

10.1.2 MultiRa-Designmatrizen

Als Eingabe der MultiRa-Analyse werden die Aufgaben für jede Modellvariante den Dimensionen zugeordnet. Im Fall des Anforderungsmodells (Tab. 8.2, S. 114) wird die Clusterzuordnung der Aufgaben direkt in die MultiRa-Designmatrix übertragen (Tab. 10.1 bis 10.3, S. 159ff). Die letzte Zeile der Designmatrix des Anforderungsmodells lautet „1 0 0 0“, denn die Aufgabe **Zahlenreihe** (S. 196) spricht die erste Dimension an. Auch im Fall des Fähigkeitsmodells A (Tab. 8.5, 8.7, 8.9, S. 127ff) wird die Faktorzuordnung der Aufgaben direkt in die MultiRa-Designmatrix übertragen. Die letzte Zeile der Designmatrix des Fähigkeitsmodells A für die jüngste Altersgruppe lautet „0 1 0“, denn die Aufgabe **Zahlenreihe** wird in diesem Modell der zweiten Dimension zugeordnet.

Im Fall des Fähigkeitsmodells B werden die Dimensionen unmittelbar durch Merkmale charakterisiert, was im Hinblick auf die Zielstellung erstrebenswert ist. Für die vergleichende MultiRa-Analyse aber müssen die Dimensionen durch Aufgaben vermittelt werden. Die Merkmalsdimensionen werden in Aufgabendimensionen übersetzt, indem für jede Aufgabe ihre Merkmale mit den relevanten Merkmalen einer Dimension abgeglichen werden. Wenn die Zahl der anwesenden und in der Dimension relevanten Merkmale die Zahl der abwesenden relevanten Merkmale und der dimensionsfremden Merkmale übersteigt, wird die Aufgabe dieser Dimension zugeordnet.

Zum Beispiel trägt **Zahlenreihe** (Aufgabe 29 in Tabelle 7.1, S. 99) die Merkmale Anforderungsbereich Anwendung, Prozessbereich Modellieren und Implementieren, Lernzielstufe Anwenden und Wissensart Prozeduren. Damit sind vier Merkmale anwesend, die in der Dimension Anwendung für die jüngste Altersgruppe relevant sind (Tab. 8.11, S. 130). Abwesend sind drei Kriterien, Formalisierungsgrad, Prozessbereich Darstellen und Interpretieren und Konzeptwissen. Die Aufgabe trägt keine Merkmale anderer Dimensionen. Sie wird in die Dimension Anwendung (I) eingeordnet, entsprechend lautet die zugehörige letzte Zeile der Designmatrix des Fähigkeitsmodells B „1 0 0 0“ (Tabelle 10.1).

Bei einigen Aufgaben ist die Zuordnung unscharf. Zum Beispiel trägt **Biber am Fluss** (Aufgabe 1 in Tabelle 7.1, S. 99), die nur in Gruppe 5-7 vorlag, die Merkmale Abstraktion, Anforderungsbereich Anwendung, Prozessbereich Darstellen und Interpretieren, Lernzielstufe Analysieren, Wissensarten Konzepte und Prozeduren. Die Kriterien Anforderung, Prozessbereich und Wissensart sprechen

Tabelle 10.1: MultiRa-Designmatrizen der Modelle, Jahrgangsstufe 5-7.

Aufgabe	Anforderungsmodell				Fähigkeitsmodell A			Fähigkeitsmodell B			
	I	II	III	IV	I	II	III	I	II	III	IV
1 Biber am Fluss	1	0	0	0	1	0	0	1	0	0	0
3 Bibers Geheimcode	0	0	1	0	0	1	0	0	1	0	0
6 Computervirus	0	0	0	1	0	0	1	0	0	0	1
7 Dateisuche	0	0	1	0	0	1	0	0	1	0	0
8 Dino-Ordnung	0	0	1	0	0	1	0	0	1	0	0
11 Fenster schließen	0	0	0	1	1	0	0	0	0	0	1
13 Link	0	0	0	1	1	0	0	0	0	0	1
14 Links um!	1	0	0	0	1	1	0	1	0	0	0
17 Platzwechsel	0	1	0	0	1	0	0	0	0	1	0
21 Schnitzeljagd	1	0	0	0	0	0	1	1	0	0	0
22 Sicheres Passwort	0	0	0	1	1	0	0	0	0	0	1
23 Umparken	1	0	0	0	0	1	0	1	0	0	0
28 Wetter	0	0	1	0	0	0	1	0	1	0	0
29 Zahlenreihe	1	0	0	0	0	1	0	1	0	0	0

für die Dimension Anwendung. Dagegen spricht, dass Formalisieren, Modellieren und Implementieren sowie Anwenden, die die Dimension prägen, für die Aufgabe nicht zutreffen. Das Vorhandensein von Abstraktion, Analysieren und Prozedurwissen weist auch auf die Dimension Problemlösung hin. Für die Problemlösungsdimension sind drei Kriterien nicht erfüllt, dazu kommen drei dimensionsfremde Merkmale, so dass **Biber am Fluss** in Summe der Anwendungsdimension (I) näher ist als der Problemlösungsdimension (Tabelle 10.1).

Die Aufgabenzuordnung, die sich solcherart aus dem Fähigkeitsmodell ergibt, ist nahezu identisch mit der Aufgabenzuordnung durch die Clusteranalyse. Das kann daran liegen, dass die Fähigkeitsstruktur, die bei den Schülerinnen und Schülern tatsächlich erkennbar ist, die Anforderungsstruktur bestätigt, die von den Aufgabenstellern vorweggenommen wird. Wahrscheinlicher ist, dass die Wechselbeziehung der Aufgabenmerkmale das Zusammenspiel der Personenmerkmale überdeckt (vgl. S. 104 zur Konfundierung der Effekte).

10.1.3 MultiRa-Ergebnisse

Die Tabellen 10.1 bis 10.3 enthalten die Designmatrizen für die drei Altersgruppen. Die Modellvarianten werden jeweils nebeneinander dargestellt. Die Bezeichnung der Dimensionen entspricht den ursprünglichen Nummern der Cluster und

Tabelle 10.2: MultiRa-Designmatrizen der Modelle, Jahrgangsstufe 8-10.

Aufgabe	Anforderungsmodell				Fähigkeitsmodell A			Fähigkeitsmodell B		
	I	II	III	IV	I	II	III	I	II	III
2 Biber und Bisons	0	1	0	0	1	0	0	0	1	0
3 Bibers Geheimcode	0	0	1	0	1	0	0	1	0	0
4 Biberzahlen	1	0	0	0	1	0	0	1	0	0
5 Binärbaum	1	0	0	0	0	1	0	1	0	0
6 Computervirus	0	0	0	1	0	0	1	0	0	1
9 Endlosschleife	1	0	0	0	0	1	0	1	0	0
10 Falschgeld	1	0	0	0	1	0	0	1	0	0
16 Netzwerkkabel	0	1	0	0	1	0	0	0	1	0
19 Primärschlüssel	0	0	0	1	1	0	0	0	0	1
20 Private E-Mail	0	0	0	1	1	0	0	0	0	1
21 Schnitzeljagd	1	0	0	0	1	1	0	1	0	0
22 Sicheres Passwort	0	0	0	1	0	0	1	0	0	1
23 Umparken	1	0	0	0	0	1	0	1	0	0
25 Verschlüsselung	1	0	0	0	0	1	0	1	0	0
29 Zahlenreihe	1	0	0	0	1	0	0	1	0	0

Faktoren in den Analysen. Jede Matrix repräsentiert die Hypothese, das mehrdimensionale Strukturmodell, dessen Dimensionen durch die zugeordneten Aufgaben charakterisiert werden, passe zu den vorliegenden Daten.

Als Ausgabe der MultiRa-Berechnungen sind in Tabelle 10.4 (S. 162) die Resultate des Bootstraps für zwei verschiedene Prüfstatistiken zu sehen, χ^2 (Chi-Quadrat) und CR , nach Cressie und Read benannt (Carstensen 2000, S. 189; Bühner 2006, S. 346f). Sie kommen zustande, indem unter der Annahme der Modellgeltung 100 Stichproben simuliert und die Prüfstatistiken berechnet werden. Eine empirische Näherung der Häufigkeitsverteilungen von χ^2 und CR entsteht und der Rangplatz der Prüfstatistiken der echten Daten unter den resimulierten Stichproben kann ermittelt werden. Die häufigen Werte belegen die ersten Plätze, selten vorkommende Werte die letzten. Nimmt etwa χ^2 für das Anforderungsmodell in der Stufe 5-7 den Rang 101 ein, liegt der Wert der Prüfgröße am äußersten Verteilungsrand und gilt als unwahrscheinlich. Genauer: es ist unwahrscheinlich, dass bei einer Stichprobe, für die das Modell gilt, dieser Wert auftritt.

Zum Rangplatz 101 gehört die Wahrscheinlichkeit $p = 0,00$, dass die echten Daten gleich gut oder besser zum Modell passen als die resimulierten Daten. p ist nicht als Irrtumswahrscheinlichkeit bei der Ablehnung einer Negativhypo-

Tabelle 10.3: MultiRa-Designmatrizen der Modelle, Altersstufe 11+.

Aufgabe	Anforderungsmodell			Fähigkeitsmodell A		Fähigkeitsmodell B		
	I	II	III	I	II	I	II	III
2 Biber und Bisons	0	1	0	1	0	0	1	0
4 Biberzahlen	1	0	0	1	0	1	0	0
6 Computervirus	0	0	1	0	1	0	0	1
9 Endlosschleife	1	0	0	0	1	1	0	0
10 Falschgeld	1	0	0	1	0	1	0	0
12 Labyrinth	0	1	0	0	1	1	1	0
15 Morse-Code	0	1	0	1	0	0	1	0
16 Netzwerkkabel	0	1	0	0	1	0	1	0
18 POP und PUSH	1	0	0	0	1	1	0	0
19 Primärschlüssel	0	0	1	1	0	0	0	1
21 Schnitzeljagd	1	0	0	1	0	1	0	0
25 Verschlüsselung	1	0	0	0	0	1	0	0
26 Verwandlung	0	1	0	1	1	0	1	0
27 Wertetausch	1	0	0	1	0	1	0	0

these zu interpretieren (Bortz 2005, S. 107ff). Die Voraussetzung der bekannten Verteilung einer Prüfgröße ist nicht erfüllt. In dieser Lage liefert MultiRa „eine empirisch erzeugte Prüfverteilung als zweitbeste Wahl nach theoretisch ableitbaren Prüfverteilungen“ (Carstensen 2000, S. 188). Innerhalb der Verteilung, die durch Bootstrapping ermittelt wurde, ist p die „Wahrscheinlichkeit für einen gleichen oder besseren Modell-Fit der echten Daten gegenüber den resimulierten Datensätzen“ (Carstensen 2000, S. 192). Das erlaubt nur, die Hypothese der Modellpassung bei geringer Wahrscheinlichkeit zu verwerfen, nicht, sie zu bestätigen.

In der unteren Altersgruppe (5-7) erzielt χ^2 für alle drei Modellvarianten nur Rangplatz 101 bei 100 Resimulationen. Die Wahrscheinlichkeit p einer gleichen oder besseren Modellanpassung der echten Daten liegt entsprechend bei null. Demnach müssten die Modelle abgelehnt werden. CR erreicht allerdings für das Anforderungsmodell den Rang 34 und damit eine hohe Anpassungswahrscheinlichkeit von 0,67. Ähnlich sehen die Ergebnisse für die Fähigkeitsmodelle A und B aus. Demnach könnte die Annahme aufrecht erhalten werden, dass die Modelle die beobachteten Daten erklären.

In der mittleren Altersgruppe (8-10) deuten umgekehrt die p -Werte, die auf der Cressie-Read-Statistik basieren, auf die Ablehnung der Anpassungshypothese, während die p -Werte, die auf der Chi-Quadrat-Statistik basieren, die Beibehaltung der Hypothese erlauben. Im Modellvergleich zeigt das vierdimensionale

Tabelle 10.4: Güte der Datenanpassung der Modelle. Rang der echten Daten und Wahrscheinlichkeit p für eine gleiche oder bessere Modellanpassung der echten Daten unter den resimulierten Datensätzen.

	Anforderungsmodell		Fähigkeitsmodell A		Fähigkeitsmodell B	
5-7	χ^2	CR	χ^2	CR	χ^2	CR
<i>Rang</i>	101	34	101	30	101	35
<i>p</i>	0,00	0,67	0,00	0,71	0,00	0,66

	Anforderungsmodell		Fähigkeitsmodell A		Fähigkeitsmodell B	
8-10	χ^2	CR	χ^2	CR	χ^2	CR
<i>Rang</i>	63	99	83	101	82	100
<i>p</i>	0,38	0,02	0,18	0,00	0,19	0,01

	Anforderungsmodell		Fähigkeitsmodell A		Fähigkeitsmodell B	
11+	χ^2	CR	χ^2	CR	χ^2	CR
<i>Rang</i>	101	101	101	101	101	101
<i>p</i>	0,00	0,00	0,00	0,00	0,00	0,00

Anforderungsmodell mit $p = 0,38$ gegenüber den dreidimensionalen Fähigkeitsmodellen A und B mit $p = 0,18$ und $p = 0,19$ die höchste Anpassungsgüte. Es leuchtet ein, dass die größere Dimensionenzahl eine bessere Erklärung der Daten ermöglicht – zu Lasten der Einfachheit des Modells.

In der oberen Altersgruppe (11+) sprechen beide Prüfstatistiken, χ^2 und CR , gegen die Geltungsannahme des vierdimensionalen Anforderungsmodells, des zweidimensionalen Fähigkeitsmodells A sowie des dreidimensionalen Fähigkeitsmodells B.

Auffallend ist, dass die Ergebnisse des Bootstraps für die beiden Prüfgrößen sich nicht zur Deckung bringen lassen. Gemäß Cressie-Read-Statistik wäre die Anpassungshypothese in der jüngsten Altersgruppe für alle drei Modelle beizubehalten. Gemäß Chi-Quadrat-Statistik wäre die Hypothese in der mittleren Altersgruppe für alle drei Modelle beizubehalten. In der ältesten Gruppe wäre die Anpassungshypothese nach beiden Teststatistiken zu verwerfen. Weiter fällt auf, dass die Ergebnisse altersgruppenweise und nicht modellweise übereinstimmen. Das deutet darauf hin, dass die generelle Eignung des Aufgabensatzes, ein mehrdimensionales Modell abzubilden, sich stark auf das Ergebnis auswirkt.

Insgesamt wurden drei Modellvarianten in drei Altersstufen auf ihre Gültigkeit getestet. Der Schluss daraus ist nicht, einige der Varianten zu akzeptieren und andere zu verwerfen. Zum einen widersprechen sich die Wahrscheinlichkeitswerte bezüglich der Prüfgrößen Chi-Quadrat und Cressie-Read. Zum anderen hängen die Ergebnisse offenbar stark vom Aufgabensatz ab. Eine zielgerichtete Vorgehensweise wäre, anhand der itembezogenen MultiRa-Statistik die Aufgaben zu selektieren, die am wenigsten zum Modell passen, und mit den restlichen Aufgaben eine erneute Parameterschätzung durchzuführen. Das verbietet sich jedoch, weil die Aufgabenzahl zu klein ist. Bereits jetzt kommt es vor, dass eine Dimension nur von einer einzelnen Aufgabe repräsentiert wird. Dieser und weitere Kritikpunkte an den Daten und der Modellbildung, die sich im Verlauf der Untersuchung ergaben, werden im nächsten Abschnitt zusammengeführt.

10.2 Reflexion

Aus theoretischen Merkmalen und empirischen Daten wurden rechnerisch drei unterschiedliche Modellansätze informatischer Anforderungen und Fähigkeiten entwickelt. Die mehrdimensionalen Modelle sind plausibel und sehr gut vereinbar mit den vorherrschenden Kompetenzvorstellungen, wie sie etwa in den Einheitlichen Prüfungsanforderungen in der Abiturprüfung Informatik (KMK 2004b) zum Ausdruck kommen. Im Verlauf der Modellentwicklung wurden jedoch auch Beschränkungen und Probleme offenbar, die keine Auflösung fanden. Die kritischen Punkte werden im Folgenden reflektiert.

Merkmalsauswahl

Die Auswahl der Merkmale wirkt sich unmittelbar auf die Aussagekraft des Modells aus. Nur solche Kompetenzaspekte, die durch eine Eingangsvariable repräsentiert sind, können sich im Modell niederschlagen. In den Katalog der vorliegenden Arbeit wurden informatikrelevante Aufgabenmerkmale aus dem kognitiven Bereich aufgenommen, die Gebieten wie der traditionellen Didaktik und der aktuellen Bildungsforschung entstammen. So umfassend der Merkmalskatalog ist, kann er nicht den Anspruch auf Vollständigkeit erheben. Als Beispiel sei das Kriterium „Offenheit der Aufgabenstellung“ genannt. Obwohl es höchstwahrscheinlich in engem Bezug zur Aufgabenschwierigkeit steht, wird es nicht berücksichtigt, denn im Multiple-Choice-Format der Wettbewerbsaufgaben ist die Offenheit der Aufgabenstellung nicht abbildbar.

Datengrundlage

Mit 21.746 Datensätzen des Schülerwettbewerbs Informatik-Biber, den Einreichungen von Mädchen und Jungen aller Jahrgangsstufen von der fünften bis zur dreizehnten Klasse und aus allen Bundesländern, steht eine umfangreiche

empirische Datenbasis zur Verfügung. Die Wettbewerbsaufgaben wurden innerhalb einer Expertenbefragung nach den Merkmalen eingeschätzt, so dass eine fundierte Klassifizierung vorliegt. Kritisch ist die geringe Zahl von fünfzehn Aufgaben je Altersstufe. In den Aufgaben, die nicht für die Untersuchung konstruiert wurden, treten die Merkmale in unsystematischer Kombination auf. Einige Merkmale kommen nur gleichzeitig vor und können nicht isoliert betrachtet werden. Diese Variablenkonfundierung beeinflusst die Faktorenanalysen, indem durch Korrelationsbeziehungen der Anforderungsmerkmale Korrelationen zwischen den Fähigkeitsmerkmalen vorgetäuscht werden.

Modellbildung

Zur Identifizierung der Anforderungs- und Fähigkeitsdimensionen wurden die klassifizierenden statistischen Verfahren der Cluster- und der Faktorenanalyse eingesetzt. Für die Clusteranalyse wäre empfehlenswert, mit unkorrelierten Variablen zu arbeiten, um eine ungewollte Gewichtung zu vermeiden. Da es dem exploratorischen Charakter der Studie entgegensteht, aufgabeninhärente und potentiell schwierigkeitsgenerierende Merkmale im Vorhinein auszuschließen, wurde die Clusteranalyse mit abhängigen Variablen durchgeführt. Der Einfluss der korrelierten Merkmale auf das *Anforderungsmodell* als Ergebnis der Clusteranalyse ist kalkulierbar gering.

Gravierender ist die Wirkung auf das Ergebnis der Faktorenanalyse der Merkmalsquoten, die zum *Fähigkeitsmodell B* führte. Die Wechselbeziehung der Anforderungsmerkmale lässt sich nicht von der Wechselbeziehung der Fähigkeitsmerkmale trennen (siehe auch die Kritik an der Datengrundlage). Das Modell B, wenn gleich es sehr gut interpretierbar, inhaltlich schlüssig und erklärungs mächtig erscheint, bleibt mit dem Verdacht behaftet, die Merkmalsstruktur der Aufgabenanforderungen zu spiegeln, nicht die gesuchte Merkmalsstruktur der Schülerfähigkeit.

Das *Fähigkeitsmodell A* als Vergleichsmodell ist unberührt von der Konfundierungsproblematik. Es wurde durch Faktorisieren der Aufgaben nach den Antwortmustern gewonnen. Dabei wurde außer Acht gelassen, dass die Faktorenanalyse für metrische Variablen konzipiert ist und mit dichotomen Variablen eventuell nicht optimal arbeitet. Das Ergebnis ist unscharf, die Interpretation unsicher. Zudem wurden Schwierigkeitsfaktoren gebildet, was ein bekannter, unerwünschter Effekt ist, wenn dichotome anstelle metrischer Daten faktorisiert werden.

Zusammenfassung

Für alle drei Modelle ist die Auswahl der Merkmale von zentraler Bedeutung, da die Modelle nur die Schwierigkeitsfacetten abbilden können, die durch die Merkmale ausgedrückt werden. In diesem Rahmen bildet das Anforderungsmodell die dimensionale Struktur der Aufgabenanforderungen prägnant und plausibel ab. Für die Entwicklung eines Fähigkeitsmodells stellt die geringe Aufga-

benzahl und die fehlende systematische Variation der Merkmale in den Aufgaben eine große Beschränkung dar. Vor allem das Modell B, direkt aus den Personen-Merkmalquoten ermittelt, ist stimmig und theoriekonform, steht aber im Verdacht, dass es durch die Konfundierung der Merkmalsvariablen verzerrt wird.

Das konstruktive Vorgehen, mit statistischen Methoden aus den Aufgabenmerkmalen und der Resonanz bei den Schülern die Modelldimensionen herauszuarbeiten, hat sich bewährt und erbrachte verschiedene Modellansätze der informatischen Anforderungs- und Fähigkeitsstruktur. Auf den Modellansätzen und den Erfahrungen aus ihrer Erarbeitung aufbauend sollte eine weitere Iteration der Modellierung im Unterschied zu dieser Studie einen geplanten Versuchsaufbau mit systematischer Aufgabenerstellung beinhalten.

Kapitel 11

Resümee

Diese Arbeit entstand, als die Kompetenzmodellierung in Deutschland im Zusammenhang mit länderweiten und länderübergreifenden Bildungsstudien in den Forschungsfokus rückte. Die kompetenzorientierte Sichtweise birgt für die Informatik-Didaktik die Aufforderung und zugleich Chance, das Selbstverständnis des Schulfachs Informatik mit neuen Erkenntnissen über spezifisch informatische Anforderungen und Fähigkeiten, Arbeits- und Lernweisen, Strategien und Fehlkonzepte zu ergänzen. Bisher gibt es für die Informatik meist normative Kompetenzmodelle, die den Sollzustand beschreiben. Diese unterstellen eine Struktur, die selten empirisch überprüft wird. Es existieren kaum Modelle des Istzustands. Vor allem fehlen bewährte Verfahren, die dimensionale Struktur der Anforderungen und Fähigkeiten zu ergründen und Modelle zu entwickeln. Auch Datenmaterial, wie es in den Fächern vorliegt, die an Schulleistungsstudien wie PISA teilnehmen, existiert noch nicht.

11.1 Ziel und Ergebnis

In der beschriebenen Situation schien das Ziel ehrgeizig, aber lohnend, einen empirisch-statistischen Weg abzustecken, der von bekannten und vermuteten Schwierigkeitsmerkmalen zu einem Strukturmodell informatischer Anforderungen und Fähigkeiten führt. Das Unterfangen war nur möglich, weil die Teilnahmedaten von über zwanzigtausend Schülern am Online-Wettbewerb Informatik-Biber zur Verfügung standen. Die pragmatische Entscheidung für die Verwendung dieser Daten brachte nicht nur Einschränkungen des Geltungsbereichs mit sich, etwa die Restriktion auf die Sekundarstufe. Sie war auch mit methodischen Problemen verbunden, die aus der Variablenkonfundierung im Aufgabensatz herrühren.

Die Auswertung von Publikationen zu Lernzieltaxonomien, Bildungsstandards, Kompetenzmodellen und Schwierigkeitsmerkmalen in den benachbarten Fächern, etwa Mathematik, ergab einen Katalog kognitiver Anforderungsmerkmale wie den Abstraktionsgrad, der den Umgang mit Abstraktion und das Vermögen zur Ab-

straktion anspricht. In einer Expertenbefragung, an der sich 45 Lehrende an Universitäten und Schulen beteiligten, wurden die Aufgaben des Schülerwettbewerbs Informatik-Biber 2007 nach den kognitiven Anforderungsmerkmalen eingeschätzt.

Auf der Basis von 21.746 Teilnehmerlösungen zu den klassifizierten Aufgaben wurden die Merkmalseffekte auf die Aufgabenschwierigkeit analysiert. Vorwiegend die strukturentdeckenden statistischen Methoden der Cluster- und Faktorenanalyse kamen zur Anwendung, um aus der Merkmalsvielfalt die Dimensionen der Aufgabenanforderungen und Personenfähigkeiten herauszuarbeiten. Die Analysen führten zu drei Strukturmodellen, dem sogenannten Anforderungsmodell und zwei alternativen Fähigkeitsmodellen. Strukturprüfend wurde das mehrdimensionale Rasch-Modell MultiRa zur Überprüfung der Datenpassung der Modellalternativen eingesetzt.

11.2 Modellbildung

Das Anforderungsmodell fächert die Aufgabenschwierigkeit in die Dimensionen Wiedergabe, Verständnis, Anwendung und Problemlösung auf. Durch die Eingangsmerkmale werden die Dimensionen anschaulich charakterisiert und deutlich voneinander abgegrenzt. Der Wesenskern der Dimensionen gleicht den bekannten Anforderungsbereichen Wiedergabe, Anwendung und Problemlösung, die den Bildungsstandards und Prüfungsanforderungen der Kultusministerkonferenz (KMK 2004b, KMK 2005) zugrundeliegen. Die Anforderungsbereiche gehen bereits auf die Lernstufen nach Roth (1971; siehe auch S. 20) zurück, Reproduktion – Reorganisation – Transfer – Problemlösen – Kreativität. Im Unterschied zu den Lernstufen handelt es sich im Modell ausdrücklich nicht um Kompetenzstufen, sondern um unabhängige Kompetenzkomponenten. Idealerweise lassen sich die kognitiven Anforderungen von Aufgaben wie auch die Problemlösefähigkeiten von Personen additiv aus den Teilkompetenzen in den einzelnen Dimensionen zusammensetzen. Das erfordert allerdings ihre Quantifizierung, die in dieser Phase noch nicht geleistet wurde.

Die Modellierung der kognitiven Aufgabenanforderungen ist gelungen. Die Fähigkeitsmodellierung jedoch stieß auf Hindernisse. Die erste Faktorenanalyse, die die Aufgaben nach den Antwortmustern bündelt (Welche Aufgaben löste der Schüler?), erbrachte unscharfe Faktoren, deren Interpretation Schwierigkeit bereitet. Die zweite Faktorenanalyse, die die Merkmale nach den Lösungsquoten bündelt (Wie viele der Aufgaben eines Merkmals löste der Schüler?) erbrachte zwar charakteristische Faktoren, die aber möglicherweise verfälscht sind. Gemessen an der Zahl von sechzehn interessierenden Merkmalen ist die Zahl von fünfzehn Aufgaben, die ein Schüler bearbeitet, sehr gering. Es ist zu befürchten, dass Merkmale konfundiert sind, das heißt, sie treten gehäuft in Kombination auf. Dieser Störeffekt überlagert die Fähigkeitsstruktur. Äußerstenfalls modellieren die Faktoren nicht wie erwünscht *Fähigkeitsdimensionen*, indem sie Merkmale

nach Erfolgsquoten zusammenfassen, sondern *Anforderungsdimensionen*, indem sie Merkmale nach ihrem Vorkommen in den Aufgaben zusammenfassen. Im Sinne der Weitergabe von Erfahrungen wurden die Fähigkeitsmodelle nicht verworfen, sondern als Modellalternativen beibehalten, um an ihrem Beispiel die vergleichende Modellgeltungsprüfung auszuführen. Bewusst wurden stets die methodischen Probleme thematisiert, die sich aus der geringen Aufgabenzahl ergeben.

11.3 Einschränkungen

Erwünscht ist die Fokussierung auf kognitive Aspekte der Informatikkompetenz, die sich aus der Kompetenzdefinition als „kontextspezifische kognitive Leistungsdispositionen“ von Klieme und Leutner (2006, S. 879) ergibt.

Ebenso gewollt ist die Beschränkung auf ein statisches Modell, das auf einer Momentaufnahme der kognitiven Fähigkeiten beruht. Weder gehen Prozessinformationen zum Begreifen und Bearbeiten der Aufgaben ein, noch wird die zeitliche Entwicklung des Leistungsvermögens der Schüler berücksichtigt.

Eine ungewollte Einschränkung resultiert daraus, dass es nicht möglich war, Einfluss auf den Versuchsaufbau zu nehmen – die Daten wurden nicht zweckgerichtet erhoben, sondern sind ein Nebenprodukt des Schülerwettbewerbs. Die geringe Zahl der Aufgaben und die Tatsache, dass sie nicht systematisch als Merkmalsträger konstruiert wurden, bringt methodische Schwierigkeiten mit sich, die die Validität des resultierenden Modells in Frage stellen.

11.4 Beitrag zur Informatik-Didaktik

Das Strukturmodell, das die kognitiven *Anforderungen* von Informatikaufgaben in die vier Dimensionen Wiedergabe, Verständnis, Anwendung und Problemlösung aufschlüsselt, bereichert die Informatik-Didaktik um ein deskriptives Kompetenzmodell. Ein deskriptives Modell bildet die Wirklichkeit ab, im Gegensatz zu einem normativen Modell, das eine Erwartung ausdrückt. Die Ausformung der Dimensionen überrascht nicht, sie harmonisiert mit klassischen Lernstufen und etablierten Anforderungsbereichen. Neu ist, die Anforderungsbereiche überhaupt als Dimensionen zu begreifen, die einen Kompetenzraum aufspannen, nicht als Stufen. Weder setzt eine Dimension die andere voraus noch wächst von einer zur anderen die Schwierigkeit. Das Dimensionenmodell liefert ein Bezugssystem, in das Aufgaben anhand ihrer Merkmale eingeordnet werden können. Dagegen unterliegen die Modelle A und B, die die kognitiven *Fähigkeiten* der Schüler abbilden, noch einigen Einschränkungen und sind nicht ohne Weiteres übertragbar.

Von großem Wert ist die Erprobung eines Herangehens an die Kompetenzmodellierung, welches in der Informatik-Didaktik noch nicht beschrieben wurde: Mit statistischen Methoden aus theoretischen Kriterien, Aufgaben und empirischen

Schülerdaten ein Modell zu konstruieren, das die Kriterien zu Dimensionen verdichtet. Diese Bottom-up-Herangehensweise zeigte sich gangbar und zielführend. Die erste Forschungsfrage wird positiv beantwortet: aus kognitiven Merkmalen von Aufgaben lässt sich mittels einer Clusteranalyse ein Strukturmodell ableiten, welches die Dimensionen informatischer Anforderungen abbildet (vgl. S. 59, Forschungsfrage 1). Gewinnbringend ist, dass die empirisch-statistische Herangehensweise zu theoretisch nachvollziehbaren Dimensionen führt, die empirisch auffindbar sind. Diese können die informatischen Anforderungen existierender Aufgaben und somit Bearbeitungserfolge und Schwierigkeiten erklären.

Bei der Faktorenanalyse der Personen-Merkmalsquoten stieß der Weg an eine Grenze, die Variablenkonfundierung, so dass fraglich bleibt, ob die resultierenden Dimensionen wirklich das Konstrukt „Fähigkeit“ modellieren. Die methodischen Probleme, auch wenn sie durch die Daten und nicht durch das Verfahren bedingt sind, erlauben nicht, die zweite Forschungsfrage ebenfalls positiv zu beantworten: ob sich theoretisch mittels Faktorenanalysen ein Strukturmodell ableiten lässt, das die Dimensionen informatikspezifischer Schülerfähigkeiten abbildet, wenn zusätzlich zu den Aufgabenmerkmalen die Bearbeitungsergebnisse einbezogen werden (vgl. S. 59, Forschungsfrage 2). Mit dem gegebenen Datenmaterial lässt sich das nicht zweifelsfrei praktisch umsetzen.

Ein wichtiger Beitrag zur Informatik-Didaktik ist aber die Dokumentation der exemplarischen Modellbildung – einschließlich der auftretenden Schwierigkeiten. Im Methodikkapitel wurden sechs Vorgehensschritte formuliert (S. 60).

1. Theoretische Begründung: Aufgabenmerkmale
2. Datenerhebung: Expertenbefragung, Wettbewerb
3. Datenaufbereitung: Klassifizierung, Antwortmuster, Merkmalsquoten
4. Statistische Analysen: Dimensionale Struktur
5. Interpretation: Anforderungen und Fähigkeiten
6. Modellprüfung: Anpassungsgüte

Eine Erkenntnis aus dieser Forschungsarbeit ist, dass für die Analyse der Fähigkeiten eine systematische und gezielte Entwicklung von Aufgaben unumgänglich ist. Dies berücksichtigend wird als weiterer Schritt die Aufgabenkonstruktion eingefügt. Da die Aufgaben gezielt als Merkmalsträger konstruiert werden, entfällt die Expertenbefragung zur Klassifizierung. Ein Leitfaden für die empirisch-statistische Modellkonstruktion umfasst nun sieben Schritte.

1. Theoretische Begründung der Aufgabenmerkmale
2. Systematische Aufgabenkonstruktion

3. Aufgabenbearbeitung durch Versuchspersonen
4. Aufbereitung der Antwortmuster, Berechnung der Merkmalsquoten
5. Faktorenanalyse der dimensionalen Struktur
6. Interpretation der Fähigkeitsdimensionen
7. Prüfung der Anpassungsgüte des Modells

11.5 Ausblick

Exemplarisch wurde ein Anfangsmodell informatischer Anforderungen und Fähigkeiten entwickelt, das zunächst deren dimensionale Struktur erklärt. Auf diesem theoretischen Fundament können nun Aufgaben entwickelt werden, um das Anfangsmodell auszubauen und zu verfeinern. Die Aufgabenkonstruktion ist dabei von zentraler Bedeutung. Aufgaben zu konstruieren, in denen die Schwierigkeitsmerkmale systematisch variiert werden, ist extrem aufwändig. Wenn Qualitätsaspekte wie die Attraktivität oder auch die Authentizität von Aufgaben nicht vernachlässigt werden, ist es theoretisch und praktisch sehr anspruchsvoll, gezielt einzelne Merkmale zu variieren und andere Merkmale konstant zu halten, wie Goldin und Caldwell (1979) unverändert aktuell in ihrer ausführlichen Studie der Aufgabenvariablen mathematischen Problemlösens feststellen: „But this instance points out the extreme sensitivity which may exist towards minor wording changes in a verbal problem, and the extraordinary difficulty of controlling for every task variable which may affect an experimental outcome“ (Goldin und Caldwell 1979, S. 272). Es empfiehlt sich, der Aufgabenentwicklung Raum in einer eigenen Forschungsarbeit zu geben.

Liegt erst ein neuer Aufgabensatz vor, ist es ratsam, mit konfirmatorischen Faktorenanalysen oder zum Beispiel MultiRa-Analysen zu prüfen, inwieweit die dimensionale Struktur durch beobachtete Daten bestätigt wird. Die leitende Frage lautet: Sind die Dimensionen der Aufgabenanforderungen auf die Personenfähigkeiten übertragbar? Das würde die Modellhypothese des Fähigkeitsmodells B bestätigen, informatische Fähigkeit sei durch drei Dimensionen strukturiert, nämlich Wiedergabe, Anwendung und Problemlösung, zusätzlich durch eine vierte Dimension Verständnis in der jüngsten Altersgruppe. Zur Überprüfung werden die neuen Aufgaben in das Anforderungsmodell eingeordnet und untersucht, ob die Antwortmuster zu den postulierten Dimensionen passen. Oder die Aufgaben werden in das Fähigkeitsmodell eingeordnet und wiederum wird untersucht, ob es die Antwortmuster erklärt.

Eine vielversprechende Anwendung der Fähigkeitsmodelle ist die Erstellung typischer Schülerprofile entlang der postulierten Dimensionen, was mit Datenmaterial wie dem vorhandenen zu bewerkstelligen ist. Dazu werden die Faktor-

werte für die Schülerfälle berechnet. Die Faktorwerte dürfen nicht mit den Faktorladungen verwechselt werden, die die Enge des Zusammenhangs zwischen den Aufgabenmerkmalen und den Faktoren beschreiben. Vielmehr geben die Faktorwerte die Lage der Schülerfälle im Koordinatensystem an, das durch die Faktoren gebildet wird. Interpretiert werden sie als Stärken und Schwächen in den entsprechenden Dimensionen. Eine Clusteranalyse der Faktorwerte bildet dann Gruppen ähnlicher Stärken-Schwächen-Profile, die Schülertypen repräsentieren.

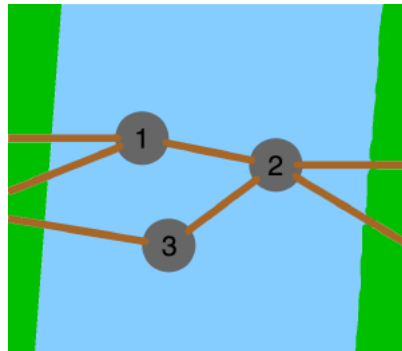
So erwachsen aus den Ergebnissen dieser Arbeit weitere Forschungsansätze. Sie zielen zum einen auf die Verfeinerung des Anforderungsmodells mit den Dimensionen Wiedergabe, eventuell Verständnis, Anwendung und Problemlösung ab. Zum anderen haben sie die Überarbeitung des Fähigkeitsmodells zum Ziel. Darüber hinaus zeigt die beispielhafte Anwendung, Schülertypen durch Merkmalsprofile zu beschreiben, dass der Wert der Modelle sich nicht darin erschöpft, theoretische Stütze weiterer Modellentwicklung zu sein. Sie liefern ein Bezugssystem, in das sich Aufgaben nach ihren kognitiven Merkmalen und Schüler nach ihren kognitiven Fähigkeiten einordnen lassen.

Anhang A

Aufgaben

Biber am Fluss

Eine Biberfamilie lebt am Fluss. Der Fluss ist zu breit, um einen Baumstamm darüber legen zu können. Glücklicherweise ragen ein paar große Geröllsteine aus dem Wasser. So können die Biber mit mehreren Baumstämmen eine Flussüberquerung bauen:



Die Steine rollen auf sandigem Flussboden immer wieder weg. Deswegen ist ein Stein besonders wichtig: wenn der wegrollt, gibt es keinen Weg mehr über den Fluss.

Welcher Stein ist das?

- A) Stein 1
- B) Stein 3
- C) Stein 2
- D) Alle Steine sind gleich wichtig.

Abbildung A.1: Biber am Fluss (Aufgabe 1)

Biber und Bisons

Biber sagen immer die Wahrheit, und Bisons lügen immer. Im Biber-und-Bison-Zeltlager wohnen insgesamt zehn Tiere. Ein blinder Maulwurf kommt vorbei und möchte wissen, wie viele Biber und wie viele Bisons anwesend sind. Darum fragt er jedes Tier: „Wie viele Biber gibt es hier?“ Die zehn Antworten sind:

3, 4, 1, 4, 1, 1, 3, 4, 3, 2

Jetzt weiß der blinde Maulwurf genau Bescheid! Du auch?

Wie viele Biber sind im Biber-und-Bison-Zeltlager?

A) 1 B) 2 C) 3 D) 4

Abbildung A.2: Biber und Bisons (Aufgabe 2)

Bibers Geheimcode

Biber möchte seinem Freund, dem Hasen, geheime Nachrichten zukommen lassen. Die beiden haben sich dafür einen Geheimcode ausgedacht. Mit dem werden ihre Nachrichten verschlüsselt, damit niemand mitlesen kann.

Bei ihrem Geheimcode bleiben die Vokale (A, E, I, O, U) und die Satzzeichen unverändert. Die Konsonanten werden durch den jeweils folgenden Konsonanten im Alphabet ersetzt. Z wird dabei durch B ersetzt.

Wie lautet Bibers Nachricht „HALB ACHT IM WALD“ im Geheimcode?

- A) HELB ECHT OM WELD
- B) JEMC EDJV ON XEMF
- C) GAKZ ABGS IL VAKC
- D) JAMC ADJV IN XAMF

Abbildung A.3: Bibers Geheimcode (Aufgabe 3)

Biberzahlen

Du weißt, wie man unsere gewohnten Dezimalzahlen als Binärzahlen aufschreibt? In der Tabelle wird es noch einmal gezeigt.

Biber hat eine weitere Schreibweise für Zahlen entwickelt. Er benutzt auch nur die Ziffern 1 und 0. Jedoch darf die Ziffer 1 höchstens so oft in einer Biberzahl vorkommen, wie der Biber Schneidezähne hat - also höchstens zweimal. Natürlich müssen alle Zahlen wieder unterschiedlich sein. Die Tabelle zeigt die Biberzahlen für die Dezimalzahlen 0 bis 10.

Beispiele:

1000100110 ist keine Biberzahl (zu viele Schneidezähne)

0000100100 ist keine Biberzahl (führende Nullen sind nicht erlaubt)

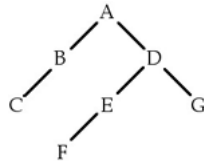
Dezimalzahl	Binärzahl	Biberzahl
0	0	0
1	1	1
2	10	10
3	11	11
4	100	100
5	101	101
6	110	110
7	111	1000
8	1000	1001
9	1001	1010
10	1010	1100

Wie lautet die Biberzahl für die Dezimalzahl 20?

- A) 10100 B) 101000 C) 100100 D) Die Zahl gibt es nicht.

Abbildung A.4: Biberzahlen (Aufgabe 4)

Binärbaum



Dies ist ein Beispiel für einen Binärbaum.

Ein Binärbaum hat eine Wurzel (die ist oben – hier 'A'), von der maximal zwei Äste abgehen. Am Ende eines Asts ist immer genau ein Knoten (hier 'B' – 'G'). Von jedem Knoten gehen wiederum maximal zwei Äste ab.

Der obige Binärbaum kann auch durch folgende Zeichenkette beschrieben werden:

$$(A(B(C))(D(E(F))(G)))$$

Welcher der unten gezeigten Binärbäume wird durch folgende Zeichenkette beschrieben?

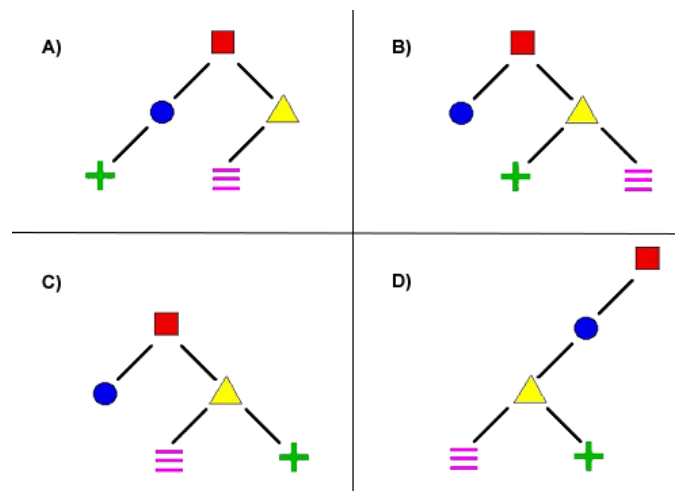
$$(\blacksquare(\bullet)(\blacktriangle(\equiv)(+)))$$


Abbildung A.5: Binärbaum (Aufgabe 5)

Computervirus

Ein Computervirus breitet sich im Betriebssystem deines Computers aus und verursacht einen solchen Schaden, dass der Computer nicht mehr gestartet werden kann. Der Computer ist nagelneu, so dass die Verkaufsgarantie noch gültig ist. Der Computervirus war allerdings nicht von Anfang an da, sondern ist erst nach ein paar Tagen aufgetaucht.

Wer übernimmt die Reparaturkosten?

- A) Der Laden, in dem der Computer gekauft wurde.
- B) Der Hersteller der Festplatte.
- C) Der Hersteller des Betriebssystems.
- D) Niemand. Du musst die Reparatur selbst bezahlen und dafür sorgen, dass das nicht noch einmal geschieht!

Abbildung A.6: Computervirus (Aufgabe 6)

Dateisuche

Im Computer kann man nach Dateien suchen, auch wenn man nur einen Teil ihres Namens benutzt. Nimm an, folgende vier Dateien sind vorhanden:

nmas.jpg
astmp.jpg
mdmtexas.png
nmtast.jpg

Wenn du mit „*.jpg“ suchst, erhältst du eine Liste mit den Dateien nmas.jpg, astmp.jpg und nmtast.jpg.

Eine Suche mit „?????.jpg“ liefert die Datei astmp.jpg.

Die Suche „*s??.*“ findet keine Datei mit einem passenden Namen.

Welche Datei findest du mit „*???as.*“?

- A) nmas.jpg
- B) astmp.jpg
- C) nmtast.jpg
- D) mdmtexas.png

Abbildung A.7: Dateisuche (Aufgabe 7)

Dino-Ordnung

In einem Buch über Dinosaurier gibt es drei Bilder, in denen Folgendes zu sehen ist:

Bild 1: Ein Dinosaurier mit zwei Beinen.

Bild 2: Ein Dinosaurier mit vier Beinen.

Bild 3: Ein Dinosaurier mit zwei Beinen, der gerade ein Tier gefangen hat.

Du sollst nun jeden dieser Dinosaurier einer der folgenden Dinosaurierarten zuordnen:

Hypsilophodon: Das ist ein zweibeiniger Pflanzenfresser.

Triceratops: Das ist ein vierbeiniger Pflanzenfresser.

Allosaurus: Das ist ein zweibeiniger Fleischfresser.

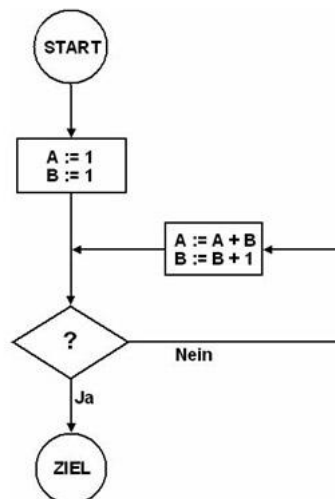
Ist diese Zuordnung eindeutig möglich?

- A) Ja, in Bild 1 ist ein Hypsilophodon, in Bild 2 ein Triceratops und in Bild 3 ein Allosaurus zu sehen.
- B) Ja, in Bild 1 und in Bild 3 sind Allosaurier zu sehen, und in Bild 2 ist ein Triceratops zu sehen.
- C) Nein, keiner der Dinosaurier kann eindeutig zugeordnet werden.
- D) Nein, der Dinosaurier in Bild 1 könnte ein Hypsilophodon oder ein Allosaurus sein.

Abbildung A.8: Dino-Ordnung (Aufgabe 8)

Endlosschleife

Ein Fluss-Diagramm beschreibt einen Algorithmus. Die verschiedenen Wege vom Start zum Ziel repräsentieren alle möglichen Wege, die der Algorithmus einschlagen kann. In den rechteckigen Kästen steht jeweils eine Folge von Befehlen. In der Raute soll eine Frage stehen; die Antwort auf die Frage bestimmt die Richtung, in die der Algorithmus weiter läuft. Der Operator $:=$ weist der Variablen auf der linken Seite den Wert auf der rechten Seite zu.



Welche Frage in der Raute führt zu einer endlosen Schleife?

- A) $A = 16$?
- B) $B = 100$?
- C) $A < B$?
- D) $A > B * 100$?

Abbildung A.9: Endlosschleife (Aufgabe 9)

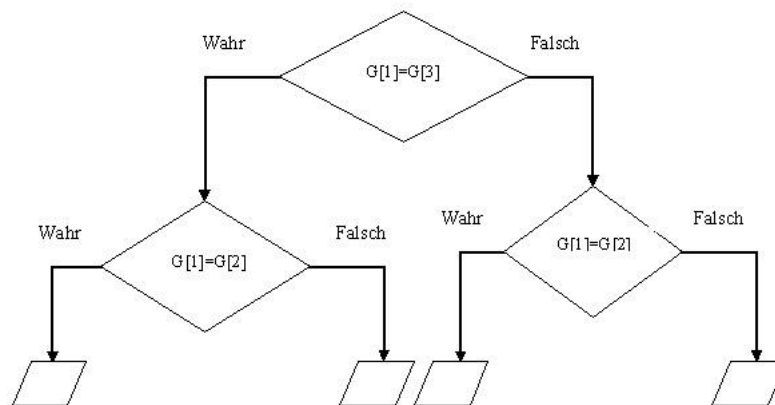
Falschgeld

Du hast 4 Münzen geschenkt bekommen, doch eine davon ist leider falsch. Die falsche Münze hat ein anderes Gewicht.



Wenn du eine Balkenwaage benutzt, kannst du mit nur zwei Vergleichen herausfinden, welche die falsche Münze ist. Gehe dazu nach dem unten abgebildeten Plan vor.

In dem Plan steht $G[1]$ für das Gewicht von Münze 1, $G[2]$ für das Gewicht von Münze 2, usw. Zuerst wiegst du also Münze 1 gegen Münze 3 ab (obere Raute), im zweiten Schritt Münze 1 gegen Münze 2. Die unterste Zeile nennt dir dann die Nummer der falschen Münze, die sich ergibt, wenn die Vergleiche nach den Beschriftungen der Pfeile ausfallen. Jedoch sind die Felder gerade leer.



Welche Nummern müssen in den leeren Feldern stehen (von links nach rechts)?

- A) 1, 2, 3, 4 B) 4, 3, 2, 1 C) 4, 2, 3, 1 D) 3, 2, 1, 4

Abbildung A.10: Falschgeld (Aufgabe 10)

Fenster schließen

Viele Programmfenster haben in einer Ecke Schaltflächen, die so (oder ähnlich) aussehen:



Du möchtest das aktive Fenster schließen.

Auf welche Schaltfläche klickst du?

- A) Auf die linke Schaltfläche.
- B) Auf die mittlere Schaltfläche.
- C) Auf die rechte Schaltfläche.
- D) Ich mache gar nichts. Irgendwann wird das Fenster von alleine zugehen.

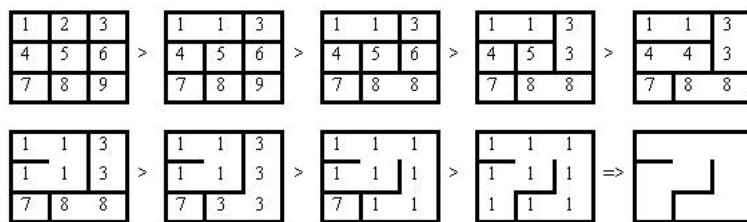
Abbildung A.11: Fenster schließen (Aufgabe 11)

Labyrinth

Im Folgenden wird eine Methode erklärt, um ein Labyrinth aus einem Rechteck mit den Seitenlängen m und n zu erstellen (m, n beide größer als 2). Dafür werden zunächst Trennwände wie ein Gitter in das Rechteck gesetzt. Jedes Feld ist nun ein Raum und bekommt eine Nummer.

Erstellen eines Labyrinths:

- Schritt 1: Verbinde zwei Räume, indem du eine Trennwand entfernst.
- Schritt 2: Nummeriere den neuen, verbundenen Raum mit der niedrigeren Nummer der beiden gerade verbundenen Räume.
- Schritt 3: Wiederhole Schritt 1 und Schritt 2 solange, bis nur noch ein Raum übrig ist. Dieser hat die Nummer 1.



Nur einer der folgenden Raumpläne wurde mit der gerade erklärten Labyrinth-Methode erstellt. Welcher?

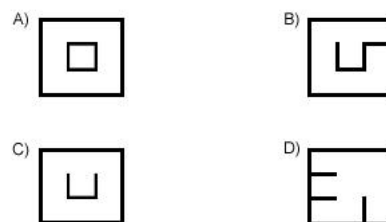


Abbildung A.12: Labyrinth (Aufgabe 12)

Link

Wenn du im Internet unterwegs bist, triffst du häufig auf den Satz:
Klicke diesen Link, um ...

Aber was ist ein Link?

- A) Eine Verknüpfung zu einer anderen Internet-Seite.
- B) Das Kabel, das den Drucker mit dem Computer verbindet.
- C) Ein anderes Wort für Internet-Seite.
- D) Ein anderes Wort für E-Mail.

Abbildung A.13: Link (Aufgabe 13)

Links um!

Du hast einen Spielzeugroboter, der folgende Befehle auf Zuruf ausführen kann:

<i>Befehl</i>	<i>Bedeutung</i>
Vor!	Der Roboter fährt 10 cm nach vorne.
Rechts!	Der Roboter dreht sich nach rechts (um 90 Grad, also einen Viertelkreis).

Du möchtest nun den Roboter so bewegen, dass er am Ende um 90 Grad (einen Viertelkreis) nach links gedreht ist.

Mit welcher Befehlsfolge kannst du das erreichen?

- A) Vor! Vor!
- B) Rechts! Rechts!
- C) Rechts! Rechts! Rechts!
- D) Vor! Rechts! Vor!

Abbildung A.14: Links um! (Aufgabe 14)

Morse-Code

In ESROM-Land gibt es nur fünf Buchstaben: E, S, R, O und M. Die Buchstaben kommen in der ESROM-Sprache unterschiedlich häufig vor (in Prozent):

E :	14 %
S :	18 %
R :	25 %
O :	18 %
M :	25 %

Ein Verein von Telegraphie-Freunden möchte Nachrichten in der ESROM-Sprache mit einem Morse-Code übermitteln, der mit möglichst wenigen Signalen auskommt und natürlich alle Wörter eindeutig kodiert.

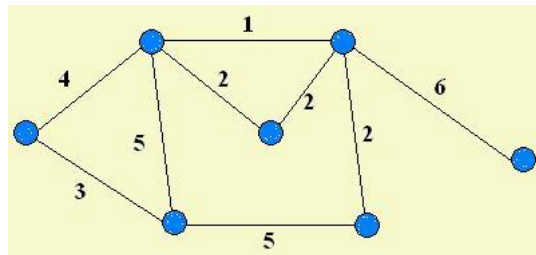
Bei welchem Morse-Code für die ESROM-Buchstaben benötigt man im Schnitt die wenigsten Signale („*“ und „-“) zur Nachrichtenübermittlung?

- | | | | | | |
|----|---------|--------|---------|--------|---------|
| A) | E = *** | S = *- | R = - | O = -* | M = * |
| B) | E = - | S = * | R = *** | O = -* | M = --- |
| C) | E = ** | S = *- | R = * | O = -* | M = - |
| D) | E = *- | S = * | R = - | O = -* | M = * |

Abbildung A.15: Morse-Code (Aufgabe 15)

Netzwerkkabel

Ein Netzwerk besteht aus sieben Computern, die durch Kabel verbunden sind. Die Kabel haben alle eine bekannte Länge (in Metern).



Einige Kabel kann man weglassen, ohne dass ein Computer komplett vom Netzwerk abgetrennt wird.

Wie viele Meter Netzwerkkabel braucht man mindestens, wenn man keinen Computer komplett abtrennen will?

- A) 18 B) 20 C) 14 D) 16

Abbildung A.16: Netzwerkkabel (Aufgabe 16)

Platzwechsel

In einer Reihe sind fünf Gesichter, zwei sind traurig und drei sind glücklich.



Nun sollen alle glücklichen Gesichter auf die linke Seite, und die traurigen Gesichter sollen auf die rechte Seite. Man darf immer zwei benachbarte Gesichter tauschen.

Wie viele Vertauschungen werden wenigstens benötigt, um die Gesichter in die gewünschte Reihenfolge zu bringen?

- A) 1 B) 5 C) 7 D) 3

Abbildung A.17: Platzwechsel (Aufgabe 17)

POP und PUSH

Wir haben vier schmale Kellerräume, um Tonnen hintereinander zu lagern. Zur Belegungsplanung dieser Keller benutzen wir die Operation:

popush (X, Y)

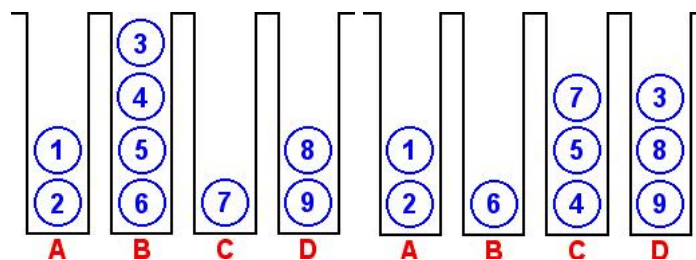
mit der Bedeutung:

Falls der Keller X nicht leer ist und

der Keller Y nicht voll ist,

dann ziehe die vorderste Tonne aus dem Keller X (pop) und schiebe sie so weit es geht in den Keller Y (push).

Wir wollen die Kellerbelegung unten links in die Kellerbelegung unten rechts ändern:



Mit welcher Folge von Operationen können wir das erreichen?

- A) popush (B,D); popush (C,D); popush (A,A); popush (B,D); popush (B,D);
- B) popush (C,D); popush (C,A); popush (B,D); popush (B,C); popush (B,C);
- C) popush (C,A); popush (B,D); popush (B,C); popush (B,C); popush (A,C);
- D) popush (B,D); popush (C,D); popush (B,C); popush (D,C); popush (B,C);

Abbildung A.18: POP und PUSH (Aufgabe 18)

Primärschlüssel

In einer Datenbank wird ein Primärschlüssel verwendet, um die Datensätze eindeutig zu identifizieren. Im Folgenden ist eine Datenbank gegeben. Jede Zeile entspricht einem Datensatz. Die Schüler ID soll ein Primärschlüssel sein.

Welche der folgenden Tabellen enthält fehlerhafte Werte für die Schüler ID?

Name	Alter	Schüler ID
Anton	18	3358
Berti	14	7194
Conni	16	5821
Det	18	3464
Edi	15	6642
Fritzchen	16	6031

A

Name	Alter	Schüler ID
Det	18	3464
Anton	18	3358
Fritzchen	16	5821
Berti	14	7194
Conni	16	5821
Edi	15	6642

B

Name	Alter	Schüler ID
Berti	14	7194
Edi	15	6642
Conni	16	5821
Fritzchen	16	6031
Anton	18	3358
Det	18	3464

C

Name	Alter	Schüler ID
Fritzchen	16	5
Anton	18	2
Det	18	6
Berti	14	1
Conni	16	4
Edi	15	3

D

Abbildung A.19: Primärschlüssel (Aufgabe 19)

Private E-Mail

Du willst eine E-Mail an neun Leute schicken, aber keiner von ihnen soll der E-Mail ansehen können, wer die anderen Empfänger sind.



The image shows a screenshot of an email client's header fields. There are three rows for 'An:', 'CC:', and 'BCC:', each with a dropdown arrow on the left and a plus sign icon on the right. Below these is a 'Betreff:' field. The fields are empty.

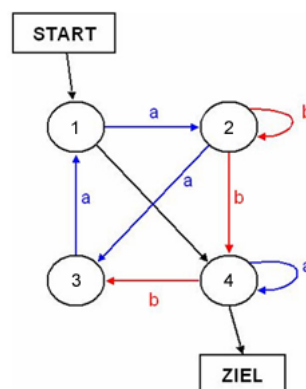
Welches ist der einfachste und schnellste Weg, dies zu erreichen?

- A) Du schickst die E-Mail neunmal los, jedes Mal mit einer einzigen Adresse im An-Feld (TO-Feld).
- B) Du schickst die E-Mail mit einer Adresse im An-Feld (TO-Feld) und acht Adressen im CC-Feld los.
- C) Du schickst die E-Mail mit drei Adressen im An-Feld (TO-Feld), drei im CC-Feld und drei im BCC-Feld los.
- D) Du schickst die E-Mail mit einem leeren An-Feld (TO-Feld) und neun Adressen im BCC-Feld (bzw. BC-Feld) los.

Abbildung A.20: Private E-Mail (Aufgabe 20)

Schnitzeljagd

Auf seinem Weg vom Start zum Ziel folgt Florian den Pfeilen, beliebig lange. Jedes Mal, wenn er einen Pfeil entlang gegangen ist, sammelt er den zugehörigen Buchstaben ein und verlängert damit eine Kette der gesammelten Buchstaben. Bei einigen Pfeilen kann er keinen Buchstaben einsammeln.



Welche der folgenden Buchstabenketten kann Florian auf seinem Weg vom Start zum Ziel nicht einsammeln?

- A) abaabba B) ba C) abaaab D) aab

Abbildung A.21: Schnitzeljagd (Aufgabe 21)

Sicheres Passwort

Du richtest dir eine neue E-mail-Adresse im Internet ein. Damit kein anderer darauf zugreifen kann, brauchst du ein sicheres Passwort. Je schwieriger ein Passwort zu erraten ist, desto sicherer ist es.

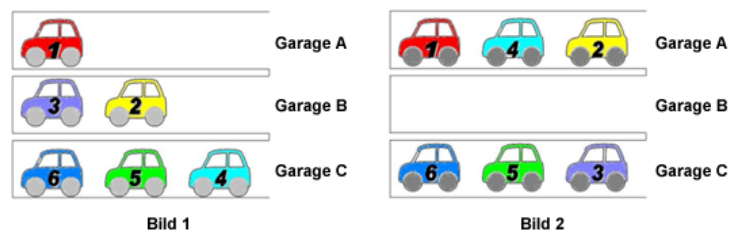
Welches der folgenden Passwörter ist am wenigsten sicher?

- A) 9 zufällig gewählte Großbuchstaben
- B) Dein Nachname, gefolgt von Deinem Geburtsjahr
- C) 5 zufällig gewählte Zeichen, also Ziffern, kleine und große Buchstaben
- D) 20 zufällig gewählte Ziffern

Abbildung A.22: Sicheres Passwort (Aufgabe 22)

Umparken

Im Hotel Biber werden die Autos der Gäste von einem Angestellten geparkt, Herrn Krause. Das ist auch gut so, weil das Hotel nur drei Garagen hat, in denen die Autos hintereinander geparkt werden müssen. In jede Garage passen maximal drei Autos. Herr Krause kennt die Fahrgewohnheiten der Gäste gut und parkt die Autos am liebsten so wie in Bild 1 gezeigt. Dann muss er nur selten umparken, wenn er einem Gast seinen Wagen bringen soll.



Am Wochenende war Herr Krause im Urlaub und Herr Brandt hat ihn vertreten. Aber als Herr Krause wiederkommt, findet er die Autos in den Garagen vor wie in Bild 2 gezeigt.

Jetzt will Herr Krause schnell die richtige Reihenfolge wiederherstellen. Leider hat er keinen weiteren Parkplatz zur Verfügung, so dass immer nur das vorderste Auto aus einer Garage X in eine andere Garage Y (so weit hinein wie möglich) umparken kann.

Diese Operation nennen wir Umparken (X, Y).

In welcher Reihenfolge muss Herr Krause die Autos umparken, um seinen bevorzugten Parkzustand (Bild 1) wieder herzustellen?

- A) Umparken (C, B), Umparken (A, C), Umparken (A, B)
- B) Umparken (C, B), Umparken (A, B), Umparken (A, C)
- C) Umparken (A, B), Umparken (C, B), Umparken (A, C)
- D) Umparken (B, C), Umparken (C, B), Umparken (A, B)

Abbildung A.23: Umparken (Aufgabe 23)

Ungeschützter Computer

Tom hat einen Computer, mit dem er auch im Internet surft. Um den Computer zu benutzen, braucht er kein Passwort. Auf dem Computer gibt es keine Firewall (also keinen Schutz gegen Kontaktversuche durch andere ans Internet angeschlossene Computer) und auch kein Programm, das gegen Viren oder andere schädliche Programme schützt.

Für welche Computer besteht durch dieses leichtsinnige Verhalten die Gefahr, durch einen Computervirus oder durch ein anderes schädliches Programm angegriffen zu werden?

- A) Für alle Computer, die mit Toms Computer im lokalen Netzwerk verbunden sind.
- B) Nur für Toms eigenen Computer.
- C) Für alle Computer auf der Welt, die mit dem Internet verbunden sind.
- D) Für alle Computer auf der Welt.

Abbildung A.24: Ungeschützter Computer (Aufgabe 24)

Verschlüsselung

Biber verschlüsselt Buchstaben mit nur zwei Ziffern 0 und 1 auf folgende Weise:

1	=	A
011	=	B
010	=	C

Mit diesem Schlüssel steht zum Beispiel „01011011“ für die Zeichenkette „CAAB“. Nun möchte Biber aber einen weiteren Buchstaben, „D“, hinzufügen. Er braucht dazu einen Schlüssel, der keine Verwechslung zulässt, so dass der Code immer eindeutig entschlüsselt werden kann. Er kann dafür z. B. nicht „11“ nehmen, weil dann „AAB“ und „DB“ mit dem selben Code „11011“ verschlüsselt würden.

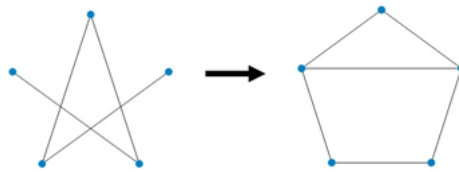
Auf welche Weise kann Biber den Buchstaben „D“ eindeutig verschlüsseln?

- A) 101
- B) 110
- C) 01110
- D) 00

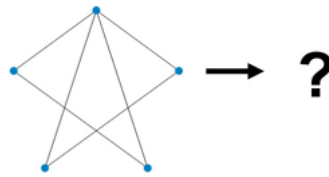
Abbildung A.25: Verschlüsselung (Aufgabe 25)

Verwandlung

Die Abbildung zeigt links eine Figur aus Punkten, die mit Linien verbunden sind. Diese Figur wird nach einer geheimen Vorschrift verwandelt; das Ergebnis ist rechts zu sehen:



Wenn du die folgende Figur nach der gleichen Vorschrift verwandelst, wie sieht dann das Ergebnis aus?



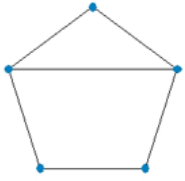
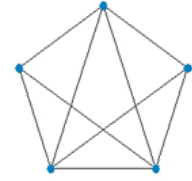
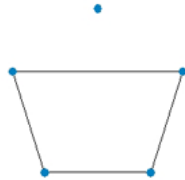
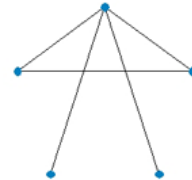
<p>A)</p> 	<p>B)</p> 
<p>C)</p> 	<p>D)</p> 

Abbildung A.26: Verwandlung (Aufgabe 26)

Wertetausch

In einer Programmiersprache haben wir zwei Variablen X und Y , die nur ganze Zahlen als Wert annehmen können. Nun möchten wir die Werte von X und Y tauschen, ohne eine weitere Variable zu benutzen.

Der Operator $:=$ weist der Variablen auf der linken Seite den Wert auf der rechten Seite zu; die Zuweisungen werden nacheinander vorgenommen.

Welche Richtige Antwort führt zum gewünschten Ergebnis?

- A) $X := Y$ B) $X := X + Y$ C) $X := X + Y$
 $Y := X$ $Y := X - Y$ $Y := X + Y$
 $X := X - Y$ $X := X - Y$
- D) Das ist ohne eine weitere Variable nicht möglich.

Abbildung A.27: Wertetausch (Aufgabe 27)

Wetter

Angenommen, das Wetter folge der Regel:

„Wenn an einem Tag die Sonne scheint, dann scheint auch am folgenden Tag die Sonne.“

Wenn heute die Sonne scheint, was kannst du daraus folgern?

- A) Die Sonne schien bisher jeden Tag und wird auch jeden weiteren Tag scheinen.
 B) Gestern schien die Sonne.
 C) Die Sonne wird nie wieder scheinen.
 D) Von heute an wird jeden Tag die Sonne scheinen.

Abbildung A.28: Wetter (Aufgabe 28)

Zahlenreihe

Nach den folgenden beiden Regeln soll Schritt für Schritt eine Reihe von Zahlen berechnet werden:

1. Ist die aktuelle Zahl ungerade, dann multipliziere sie mit drei und addiere eins.
2. Ist die aktuelle Zahl gerade, dann dividiere sie durch zwei.

Überraschend ist: Egal mit welcher Zahl man die Reihe beginnt, man kommt immer bei eins an.

Variante 1:

Du beginnst mit der Zahl 12.

Wie viele Zahlen enthält die Reihe, wenn du bei eins angekommen bist (12 und 1 werden mitgezählt)?

- A) 8 B) 11 C) 13 D) 10

Variante 2:

Du beginnst mit der Zahl 6.

Wie viele Zahlen enthält die Reihe, wenn du bei eins angekommen bist (6 und 1 werden mitgezählt)?

- A) 6 B) 8 C) 9 D) 12

Variante 3:

Du beginnst mit der Zahl 40.

Wie viele Zahlen enthält die Reihe, wenn du bei eins angekommen bist (40 und 1 werden mitgezählt)?

- A) 8 B) 9 C) 10 D) 20

Abbildung A.29: Zahlenreihe (Aufgabe 29)

Anhang B

Bildungsstandards für die Informatik

Abschlussbezogene Bildungsstandards wie sie die Kultusministerkonferenz im Fach Mathematik für den Primarbereich, den Hauptschulabschluss (Jahrgangsstufe 9), den Mittleren Schulabschluss (Jahrgangsstufe 10) und die Allgemeine Hochschulreife beschlossen hat, gibt es für das Fach Informatik noch nicht. Diese Untersuchung orientiert sich an den „Einheitlichen Prüfungsanforderungen in der Abiturprüfung Informatik“, den sogenannten EPA Informatik (KMK 2004b), und an der Empfehlung der Gesellschaft für Informatik e. V. (GI) für „Grundsätze und Standards für die Informatik in der Schule“ (GI 2008).

B.1 Anforderungsbereiche

Die Einheitlichen Prüfungsanforderungen in der Abiturprüfung Informatik, kurz EPA, wurden 1989 in der Kultusministerkonferenz verabschiedet und liegen heute in der Fassung von 2004 vor (KMK 2004b). Sie betreffen die Allgemeine Hochschulreife als Abschluss der Sekundarstufe II. Neben fachlichen und methodischen Qualifikationen und fachlichen Inhalten werden die Anforderungsbereiche Wiedergabe, Anwendung und Problemlösung beschrieben, die das Leistungsniveau der Schülerinnen und Schüler in der Abiturprüfung charakterisieren.

Anforderungsbereich I (Wiedergabe)

- Wiedergabe von bekannten Sachverhalten aus einem abgegrenzten Gebiet im gelernten Zusammenhang
- Beschreibung und Darstellung bekannter Verfahren, Methoden und Prinzipien der Informatik
- Beschreibung und Verwendung gelernter und geübter Arbeitstechniken und Verfahrensweisen in einem begrenzten Gebiet und in einem wiederholenden Zusammenhang

Anforderungsbereich II (Anwendung)

- selbstständige Verwendung (Auswählen, Anordnen, Verarbeiten und Darstellen) bekannter Sachverhalte zur Bearbeitung neuer Frage- oder Problemstellungen unter vorgegebenen Gesichtspunkten in einem durch Übung bekannten Zusammenhang
- selbstständige Übertragung des Gelernten auf vergleichbare neue Situationen, wobei es entweder um veränderte Fragestellungen oder um veränderte Sachzusammenhänge oder um abgewandelte Verfahrensweisen gehen kann
- Anwendung bekannter Verfahren, Methoden und Prinzipien der Informatik zur Lösung eines neuen Problems aus einem bekannten Problemkreis

Anforderungsbereich III (Problemlösung)

- planmäßige Verarbeitung komplexer Gegebenheiten mit dem Ziel, zu selbstständigen Gestaltungen bzw. Deutungen, Folgerungen, Begründungen, Wertungen zu gelangen
- bewusste und selbstständige Auswahl und Anpassung geeigneter gelernter Methoden und Verfahren in neuartigen Situationen. Dabei werden aus gelernten Denkmethoden bzw. Lösungsverfahren die zur Bewältigung der Aufgabe geeigneten selbstständig ausgewählt und einer neuen Problemstellung angepasst

B.2 Inhalts- und Prozessbereiche

Die Empfehlung der Gesellschaft für Informatik (GI) für Grundsätze und Standards für die Informatik in der Schule betrifft die Sekundarstufe I, die zum Mittleren Schulabschluss führt. Die „Kompetenzen über alle Jahrgangsstufen“ sind ein Diskussionsergebnis des GI-Arbeitskreises Bildungsstandards des Fachausschusses Informatische Bildung in Schulen und der Fachgruppe Didaktik der Informatik (GI 2008, S. 12ff). Eine Übersicht der „Kompetenzen über alle Jahrgangsstufen“ wird auf den beiden folgenden Seiten abgedruckt. Die fünf Inhaltsbereiche tragen die Bezeichnungen Information und Daten, Algorithmen, Sprachen und Automaten, Informatiksysteme sowie Informatik, Mensch und Gesellschaft (S. 199). Die fünf Prozessbereiche werden mit den Inhaltsbereichen verschränkt. Sie tragen die Bezeichnungen Modellieren und Implementieren, Begründen und Bewerten, Strukturieren und Vernetzen, Kommunizieren und Kooperieren, sowie Darstellen und Interpretieren (S. 200). Diese Kompetenzen werden nach Jahrgangsstufen 5 bis 7 und 8 bis 10 differenziert (GI 2008, S. 14ff) und mit Beispielaufgaben konkretisiert und erläutert (GI 2008, S. 23ff).

B.2.1 Inhaltsbereiche

Information und Daten

Schülerinnen und Schüler aller Jahrgangsstufen

- ▷ verstehen den Zusammenhang von Information und Daten sowie verschiedene Darstellungsformen für Daten,
- ▷ verstehen Operationen auf Daten und interpretieren diese in Bezug auf die dargestellte Information,
- ▷ führen Operationen auf Daten sachgerecht durch.

Algorithmen

Schülerinnen und Schüler aller Jahrgangsstufen

- ▷ kennen Algorithmen zum Lösen von Aufgaben und Problemen aus verschiedenen Anwendungsgebieten und lesen und interpretieren gegebene Algorithmen,
- ▷ entwerfen und realisieren Algorithmen mit den algorithmischen Grundbausteinen und stellen diese geeignet dar.

Sprachen und Automaten

Schülerinnen und Schüler aller Jahrgangsstufen

- ▷ nutzen formale Sprachen zur Interaktion mit Informatiksystemen und zum Problemlösen,
- ▷ analysieren und modellieren Automaten.

Informatiksysteme

Schülerinnen und Schüler aller Jahrgangsstufen

- ▷ verstehen die Grundlagen des Aufbaus von Informatiksystemen und deren Funktionsweise,
- ▷ wenden Informatiksysteme zielgerichtet an,
- ▷ erschließen sich weitere Informatiksysteme.

Informatik, Mensch und Gesellschaft

Schülerinnen und Schüler aller Jahrgangsstufen

- ▷ benennen Wechselwirkungen zwischen Informatiksystemen und ihrer gesellschaftlichen Einbettung,
- ▷ nehmen Entscheidungsfreiheiten im Umgang mit Informatiksystemen wahr und handeln in Übereinstimmung mit gesellschaftlichen Normen,
- ▷ reagieren angemessen auf Risiken bei der Nutzung von Informatiksystemen (GI 2008, S. 12f, Hervorhebungen geändert).

B.2.2 Prozessbereiche

Modellieren und Implementieren

Schülerinnen und Schüler aller Jahrgangsstufen

- ▷ erstellen informatische Modelle zu gegebenen Sachverhalten,
- ▷ implementieren Modelle mit geeigneten Werkzeugen,
- ▷ reflektieren Modelle und deren Implementierung.

Begründen und Bewerten

Schülerinnen und Schüler aller Jahrgangsstufen

- ▷ stellen Fragen und äußern Vermutungen über informatische Sachverhalte,
- ▷ begründen Entscheidungen bei der Nutzung von Informatiksystemen,
- ▷ wenden Kriterien zur Bewertung informatischer Sachverhalte an.

Strukturieren und Vernetzen

Schülerinnen und Schüler aller Jahrgangsstufen

- ▷ strukturieren Sachverhalte durch zweckdienliches Zerlegen und Anordnen,
- ▷ erkennen und nutzen Verbindungen innerhalb und außerhalb der Informatik.

Kommunizieren und Kooperieren

Schülerinnen und Schüler aller Jahrgangsstufen

- ▷ kommunizieren fachgerecht über informatische Sachverhalte,
- ▷ kooperieren bei der Lösung informatischer Probleme,
- ▷ nutzen geeignete Werkzeuge zur Kommunikation und Kooperation.

Darstellen und Interpretieren

Schülerinnen und Schüler aller Jahrgangsstufen

- ▷ interpretieren unterschiedliche Darstellungen von Sachverhalten,
- ▷ veranschaulichen informatische Sachverhalte,
- ▷ wählen geeignete Darstellungsformen aus (GI 2008, S. 13f, Hervorhebungen geändert).

Anhang C

Expertenbefragung

C.1 Versendete Unterlagen

Auf den folgenden Seiten werden die Unterlagen abgedruckt, die den Experten zugesendet wurden. Diese umfassten ein Hinweisblatt zur Aufgabenbewertung, zur Orientierung drei ausgefüllte Aufgabenprofile, die in einer Vorstudie innerhalb der Arbeitsgruppe Didaktik um Torsten Brinda an der Universität Erlangen-Nürnberg abgestimmt wurde, und fünf Aufgabenprofile mit der Möglichkeit, zutreffende Merkmalswerte anzukreuzen. Hier wird nur ein Aufgabenprofil abgedruckt. Alle Aufgaben befinden sich in Anhang A (S. 173ff).

Hinweise zur Einordnung der Aufgaben

Unter der **Erfahrungsweltnähe (1)** einer Aufgabe soll verstanden werden, wie nah sie dem persönlichen Erleben der Schüler ist. Dabei ist der *Aufgabenkontext* zu betrachten, nicht der Fachinhalt. Zum Beispiel könnte der Fachinhalt „Scheduling“ in unterschiedlich lebensweltnahe Kontexte eingebettet werden.

- Der Aufgabenkontext *Stundenplan* ist dem Schüleralltag zuzuordnen, das heißt **nah** zur Erfahrungswelt.
- Der Aufgabenkontext *Fahrplan* ist nicht unbedingt täglich erlebt, aber jederzeit erlebbar durch die Schüler, das heißt, er weist **mittlere** Entfernung zur Erfahrungswelt auf.
- Der Aufgabenkontext *Produktionsplanung* ist im Allgemeinen nicht erlebbar durch die Schüler, jedoch vorstellbar, das heißt **entfernt** von der Erfahrungswelt.

Unter dem **Abstraktionsgrad (2)** einer Aufgabe soll der Grad der Loslösung vom konkreten Beispiel verstanden werden.

- Lautet die Aufgabe, einen *Weg in einem gegebenen Netzwerk* zu suchen, so handelt es sich um einen konkreten Fall und die Aufgabe ist vom Abstraktionsgrad **konkret**.
- Lautet die Aufgabe, *mögliche Knotenausfälle in einem Netzwerk* durchzuspielen, so handelt es sich um die Betrachtung gedachter Fälle. Die Aufgabe ist von **mittlerem** Abstraktionsgrad.
- Gilt es, einen *Algorithmus zur Wegsuche in Netzwerken* zu finden, so handelt es sich um ein allgemeines Verfahren, losgelöst vom Einzelfall. Die Aufgabe ist **abstrakt**.

Die **Komplexität (3)** ergibt sich zum Beispiel aus der Anzahl der Fälle, Regeln, Vergleiche, Umformungen, die die Größe des Lösungsraums bestimmt. Zum Beispiel könnte eine Aufgabe im Kontext Schachspiel unterschiedlich komplex gestaltet werden.

- Ist nur *ein einzelner Spielzug einer Figur* zu betrachten, gilt die Aufgabe als **einfach**.
- Sind die *Spielzüge mehrerer Figuren* zu betrachten, ist die Aufgabe von **mittlerer** Komplexität.
- Ist *eine Folge von Spielzügen mehrerer Figuren* zu betrachten, gilt die Aufgabe als **komplex**.

Der **Formalisierungsgrad (4)** ist ein Merkmal der Informationsdarstellung in einer Aufgabe.

- Die Darstellung einer Information als Text (*Von drei Personen ist A die größte, B die kleinste und C die mittlere.*) oder als Bild (s. Abb. 1) ist **informell**.
- Die symbolische Repräsentation (s. Abb. 2) weist einen **mittleren** Formalisierungsgrad auf.
- Die modellhafte Darstellung, etwa als Baum oder Struktogramm (s. Abb. 3) gilt als **formal**.

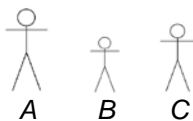


Abbildung 1:
Bildhafte Darstellung

$B < C < A$

Abbildung 2:
Symbolische Darstellung

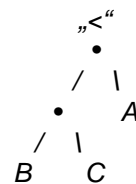


Abbildung 3:
Modellhafte Darstellung

Die **Redundanz (5)** ist ebenfalls ein Merkmal der Informationsdarstellung. Wird die gleiche Information mehrfach dargestellt, etwa als Text (*Von drei Personen ist A die größte, B die kleinste und C die mittlere.*) und zusätzlich als Bild (s. Abb. 1), so ist die Informationsdarstellung in der Aufgabe **redundant**, sonst nicht.

Der **Anforderungsbereich (6)** bezieht sich auf die geforderte Aktion. Drei Anforderungsbereiche werden in den Einheitlichen Prüfungsanforderungen in der Abiturprüfung Informatik [KK04] unterschieden.

- Die Aufgabe, *das Dualsystem zu beschreiben*, ist dem Anforderungsbereich **Wiedergabe** zuzuordnen.
- Die Aufgabe, *eine Zahl aus dem Dezimalsystem in das Dualsystem zu übertragen*, ist dem Anforderungsbereich **Anwendung** zuzuordnen.
- *Ein Verfahren zur Addition dualer Zahlen zu entwickeln*, ist dem Anforderungsbereich **Problemlösung** zuzuordnen.

Auch der **Prozessbereich (7)** bezieht sich auf die geforderte Aktion. Fünf Prozessbereiche werden in den Bildungsstandards Informatik für die Sekundarstufe I [GI08] unterschieden.

- **Modellieren und Implementieren**
 - > informatische Modelle zu gegebenen Sachverhalten erstellen
 - > Modelle mit geeigneten Werkzeugen implementieren
 - > Modelle und deren Implementierung reflektieren
- **Begründen und Bewerten**
 - > Fragen stellen und Vermutungen äußern über informatische Sachverhalte
 - > Entscheidungen bei der Nutzung von Informatiksystemen begründen
 - > Kriterien zur Bewertung informatischer Sachverhalte anwenden
- **Strukturieren und Vernetzen**
 - > Sachverhalte durch zweckdienliches Zerlegen und Anordnen strukturieren
 - > Verbindungen innerhalb und außerhalb der Informatik erkennen und nutzen
- **Kommunizieren und Kooperieren**
 - > fachgerecht über informatische Sachverhalte kommunizieren
 - > bei der Lösung informatischer Probleme kooperieren
 - > geeignete Werkzeuge zur Kommunikation und Kooperation nutzen
- **Darstellen und Interpretieren**
 - > unterschiedliche Darstellungen von Sachverhalten interpretieren
 - > informatische Sachverhalte veranschaulichen
 - > geeignete Darstellungsformen auswählen

Nach der Taxonomie von Anderson et al. (erste Dimension, kognitive Prozesse [AK01]) werden sechs **kognitive Lernzielstufen (8)** unterschieden.

- **Erinnern:** Remember – Retrieve relevant knowledge from long-term memory
- **Verstehen:** Understand – Construct meaning from instructional messages, including oral, written, and graphic communication
- **Anwenden:** Apply – Carry out or use a procedure in a given situation
- **Analysieren:** Analyze – Break material into its constituent parts and determine how the parts relate to one another and to an overall structure or purpose
- **Bewerten:** Evaluate – Make judgements based on criteria and standards
- **Erschaffen:** Create – Put elements together to form a coherent or functional whole; reorganize elements into a pattern or structure

Weiter werden nach der Taxonomie von Anderson et al. (zweite Dimension, Wissen [AK01]) vier **Arten des Wissens (9)** unterschieden.

- **Fakten:** Factual Knowledge – The basic elements students must know to be acquainted with a discipline or solve problems in it
- **Konzepte:** Conceptual Knowledge – The interrelationships among the basic elements within a larger structure that enable them to function together
- **Prozeduren:** Procedural Knowledge – How to do something, methods of inquiry, and criteria for using skills, algorithms, techniques, and methods
- **Metakognition:** Metacognitive Knowledge – Knowledge of cognition in general as well as awareness and knowledge of one's own cognition

Quellen

- [AK01] Anderson, L.W.; Krathwohl, D.R. (Hrsg.) et al.: A Taxonomy for Learning, Teaching, and Assessing. A Revision of Bloom's Taxonomy of Educational Objectives. Addison Wesley Longman, New York, 2001.
- [GI08] Gesellschaft für Informatik (GI) e.V.: Grundsätze und Standards für die Informatik in der Schule. Bildungsstandards Informatik für die Sekundarstufe I. In: LOG IN 28 (2008) 150/151, Beilage.
- [Ku04] Kultusministerkonferenz (Hrsg.): Einheitliche Prüfungsanforderungen in der Abiturprüfung Informatik. Luchterhand, München, 2004.
- [Sch08] Schlüter, K.: Je schwieriger die Aufgabe, desto klüger der Kopf? In (Brinda, T.; Fothe, M.; Hubwieser, P.; Schlüter, K. Hrsg.): Didaktik der Informatik – Aktuelle Forschungsergebnisse. 5. Workshop der GI-Fachgruppe „Didaktik der Informatik“, 24.-25. September in Erlangen. Köllen, Bonn, 2008.

Hinweise zu den Aufgabenprofilen

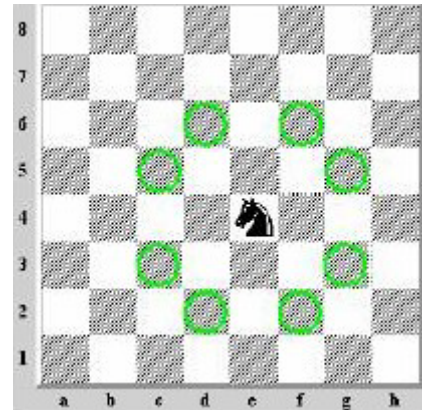
Angefügt finden Sie drei Beispielprofile zu Ihrer Orientierung. Die Aufgaben entstammen dem Vorgängerwettbewerb zum deutschen Informatik-Biber, dem EI:SPIEL blitz! 2006. Die dargestellte Einordnung entspricht der subjektiven Einschätzung durch Mitglieder unserer Arbeitsgruppe und muss nicht zwangsläufig mit Ihrer eigenen Einschätzung übereinstimmen.

Bitte füllen Sie die fünf Aufgabenprofile auf den weiteren Bögen aus. Im ersten Merkmalsblock setzen Sie je Zeile nur ein Kreuz. Im zweiten Merkmalsblock ist die Möglichkeit vorgesehen, mehr als ein Kreuz zu setzen. Unterhalb jeder Aufgabe stehen Ihnen für Kommentare zu den Merkmalen oder zur Einordnung der Aufgabe einige Zeilen zur Verfügung.

Weiteren Raum für Ihre willkommenen Anmerkungen finden Sie auf dem Bogen mit den freiwilligen statistischen Angaben. Konnten Sie die Aufgaben schlüssig einordnen? Erschienen Ihnen die Merkmale und die Merkmalskategorien angemessen? Fehlten aus Ihrer Sicht wichtige Merkmale der Aufgabenschwierigkeit?

Schachzug

Auf einem Schachbrett kann der Springer in die Felder springen, die in der Abbildung markiert sind. Jeder Zug kann beschrieben werden als eine Folge der Buchstaben U (up - ein Feld hoch), D (down - ein Feld runter), R (ein Feld nach rechts) und L (ein Feld nach links).



Welche der folgenden Buchstabenfolgen beschreibt keinen Springerzug?

- A) UUL B) DLL **C) DRU** D) URR

Schätzen Sie die Aufgabe ein. Setzen Sie dazu je Zeile ein Kreuz.

Erfahrungsweltnähe (1) Nah ~~Mittel~~ Entfernt

Abstraktionsgrad (2) ~~Konkret~~ Mittel Abstrakt

Komplexität (3) ~~Einfach~~ Mittel Komplex

Formalisierungsgrad (4) Informell ~~Mittel~~ Formal

Redundanz (5) Ja ~~Nein~~

Bei den folgenden Merkmalen sind Mehrfachantworten möglich.

Anforderungsbereich (6) Wiedergabe ~~Anwendung~~ Problemlösung

Prozessbereich (7)

Modellieren/ Implementieren	Begründen/ Bewerten	Strukturieren/ Vernetzen	Kommunizieren/ Kooperieren	Darstellen/ Interpretieren
--	------------------------	-----------------------------	-------------------------------	---

Kognitive Lernzielstufe (8)

Erinnern (remember)	Verstehen (understand)	Anwenden (apply)	Analysieren (analyze)	Bewerten (evaluate)	Erschaffen (create)
------------------------	---------------------------	---------------------------------	--------------------------	------------------------	------------------------

Art des Wissens (9)

Fakten (factual)	Konzepte (conceptual)	Prozeduren (procedural)	Metakognition (metacognitive)
---------------------	--------------------------------------	----------------------------	----------------------------------

Bemerkungen: _____

Stammbaum

Sabine hat zwei Kinder. Ihr Vater Johannes hat neben Sabine noch ein weiteres Kind, den Sohn Fabian, der wiederum selbst drei Kinder hat.

Im Allgemeinen sagen wir, dass x ein Nachfolger von y ist, entweder wenn x ein Kind von y ist, oder wenn es eine Person z gibt, so dass x ein Kind von z ist und z ein Nachfolger von y.

Wie viele Nachfolger hat Johannes?

- A) 2 B) 8 **C) 7** D) 5

Schätzen Sie die Aufgabe ein. Setzen Sie dazu je Zeile ein Kreuz.

Erfahrungsweltnähe (1)	<input checked="" type="checkbox"/> Nah	<input type="checkbox"/> Mittel	<input type="checkbox"/> Entfernt
Abstraktionsgrad (2)	<input type="checkbox"/> Konkret	<input checked="" type="checkbox"/> Mittel	<input type="checkbox"/> Abstrakt
Komplexität (3)	<input type="checkbox"/> Einfach	<input checked="" type="checkbox"/> Mittel	<input type="checkbox"/> Komplex
Formalisierungsgrad (4)	<input type="checkbox"/> Informell	<input checked="" type="checkbox"/> Mittel	<input type="checkbox"/> Formal
Redundanz (5)	<input type="checkbox"/> Ja	<input checked="" type="checkbox"/> Nein	

Bei den folgenden Merkmalen sind Mehrfachantworten möglich.

Anforderungsbereich (6)	<input type="checkbox"/> Wiedergabe	<input checked="" type="checkbox"/> Anwendung	<input type="checkbox"/> Problemlösung
--------------------------------	-------------------------------------	---	--

Prozessbereich (7)

Modellieren/ Implementieren	Begründen/ Bewerten	Strukturieren/ Vernetzen	Kommunizieren/ Kooperieren	Darstellen/ Interpretieren
--------------------------------	------------------------	-----------------------------	-------------------------------	-------------------------------

Kognitive Lernzielstufe (8)

Erinnern (remember)	Verstehen (understand)	Anwenden (apply)	Analysieren (analyze)	Bewerten (evaluate)	Erschaffen (create)
------------------------	---------------------------	---------------------	--------------------------	------------------------	------------------------

Art des Wissens (9)

Fakten (factual)	Konzepte (conceptual)	Prozeduren (procedural)	Metakognition (metacognitive)
---------------------	--------------------------	----------------------------	----------------------------------

Bemerkungen: _____

Noamesisch

Auf der kleinen Insel Noam sprechen die Eingeborenen eine ganz besondere Sprache. In langjähriger Arbeit konnte eine Sprachforscherin folgendes feststellen:

1. Es gibt die Wörter ba, di und du.
2. Durch Verdopplung eines Worts entsteht ein neues Wort; Beispiel: baba.
3. Ein neues Wort entsteht auch, indem ein Wort vorne und hinten an ein anderes Wort angefügt wird; Beispiel: baduba

Als sie versuchte, mit den Noamesen zu reden, funktionierte das ganz gut, doch bei einem der folgenden Worte zuckten die Noamesen nur verständnislos mit den Achseln.

Bei welchem?

- A) dudubabadudubabadudu
- B) didudubadududi
- C) dudubadibadibadu
- D) dididudidibadibadididididi

Schätzen Sie die Aufgabe ein. Setzen Sie dazu je Zeile ein Kreuz.

Erfahrungsweltnähe	Nah	Mittel	Entfernt <input checked="" type="checkbox"/>
Abstraktionsgrad	Konkret <input checked="" type="checkbox"/>	Mittel	Abstrakt
Komplexität	Einfach	Mittel <input checked="" type="checkbox"/>	Komplex
Formalisierungsgrad	Informell <input checked="" type="checkbox"/>	Mittel	Formal
Redundanz	Ja	Nein <input checked="" type="checkbox"/>	

Bei den folgenden Merkmalen sind Mehrfachantworten möglich.

Anforderungsbereich	Wiedergabe	Anwendung <input checked="" type="checkbox"/>	Problemlösung
----------------------------	------------	---	---------------

Prozessbereich

Modellieren/ Implementieren	Begründen/ Bewerten	Strukturieren/ Vernetzen <input checked="" type="checkbox"/>	Kommunizieren/ Kooperieren	Darstellen/ Interpretieren
--------------------------------	------------------------	---	-------------------------------	-------------------------------

Kognitive Lernzielstufe

Erinnern (remember)	Verstehen (understand)	Anwenden (apply)	Analysieren (analyze) <input checked="" type="checkbox"/>	Bewerten (evaluate)	Erschaffen (create)
------------------------	---------------------------	---------------------	--	------------------------	------------------------

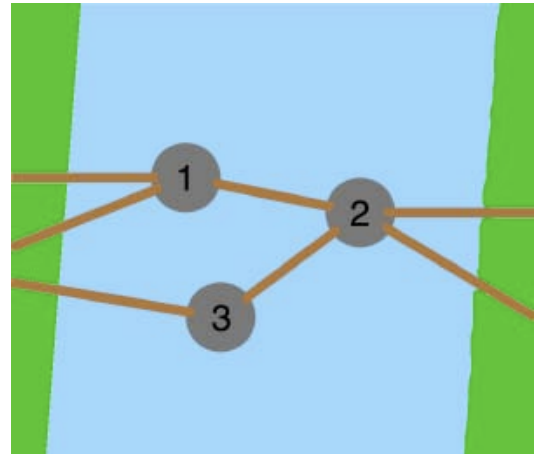
Art des Wissens

Fakten (factual)	Konzepte (conceptual)	Prozeduren (procedural) <input checked="" type="checkbox"/>	Metakognition (metacognitive)
---------------------	--------------------------	--	----------------------------------

Bemerkungen: _____

Biber am Fluss

Eine Biberfamilie lebt am Fluss. Der Fluss ist zu breit, um einen Baumstamm darüber legen zu können. Glücklicherweise ragen ein paar große Geröllsteine aus dem Wasser. So können die Biber mit mehreren Baumstämmen eine Flussüberquerung bauen:



Die Steine rollen auf sandigem Flussboden immer wieder weg. Deswegen ist ein Stein besonders wichtig: Wenn der weg rollt, gibt es keinen Weg mehr über den Fluss.

Welcher Stein ist das?

- A) Stein 1 **C) Stein 2**
 B) Stein 3 D) Alle Steine sind gleich wichtig.

Schätzen Sie die Aufgabe ein. Setzen Sie dazu je Zeile ein Kreuz.

Erfahrungsweltnähe (1)

Nah	Mittel	Entfernt
-----	--------	----------

Abstraktionsgrad (2)

Konkret	Mittel	Abstrakt
---------	--------	----------

Komplexität (3)

Einfach	Mittel	Komplex
---------	--------	---------

Formalisierungsgrad (4)

Informell	Mittel	Formal
-----------	--------	--------

Redundanz (5)

Ja	Nein
----	------

Bei den folgenden Merkmalen sind Mehrfachantworten möglich.

Anforderungsbereich (6)

Wiedergabe	Anwendung	Problemlösung
------------	-----------	---------------

Prozessbereich (7)

Modellieren/ Implementieren	Begründen/ Bewerten	Strukturieren/ Vernetzen	Kommunizieren/ Kooperieren	Darstellen/ Interpretieren
--------------------------------	------------------------	-----------------------------	-------------------------------	-------------------------------

Kognitive Lernzielstufe (8)

Erinnern (remember)	Verstehen (understand)	Anwenden (apply)	Analysieren (analyze)	Bewerten (evaluate)	Erschaffen (create)
------------------------	---------------------------	---------------------	--------------------------	------------------------	------------------------

Art des Wissens (9)

Fakten (factual)	Konzepte (conceptual)	Prozeduren (procedural)	Metakognition (metacognitive)
---------------------	--------------------------	----------------------------	----------------------------------

Bemerkungen: _____

C.2 Bewertungsverteilung, Merkmalskorrelation

Die folgenden Tabellen C.1 bis C.4 (S. 210ff) enthalten die Häufigkeitsverteilungen der Expertenbewertung aller Merkmale für alle Aufgaben. In den Tabellen C.5 bis C.7 (S. 214ff) werden die Korrelationstabellen der Merkmale in den Teilaufgabensätzen für die Jahrgangsstufen fünf bis sieben (5-7), acht bis zehn (8-10) und elf und darüber (11+) aufgeführt, analog zu Tabelle 8.1 (S. 110), die die Merkmalskorrelationen im gesamten Aufgabensatz enthält.

Tabelle C.1: Bewertungsverteilung, Teil 1: Erfahrungsweltnähe (EG; 1 = nah, 2 = mittel, 3 = entfernt), Abstraktionsgrad (AG; 1 = konkret, 2 = mittel, 3 = abstrakt) und Formalisierungsgrad (FG; 1 = informell, 2 = mittel, 3 = formal). n ist jeweils die Anzahl der Experten, die eine Bewertung abgaben.

Aufgabe	EG				AG				FG			
	1	2	3	n	1	2	3	n	1	2	3	n
Biber am Fluss	1	3	1	5	2	3	0	5	3	1	1	5
Biber und Bisons	0	3	6	9	3	5	1	9	7	2	0	9
Bibers Geheimcode	1	6	3	10	7	4	0	11	8	3	0	11
Biberzahlen	0	3	7	10	4	4	2	10	2	6	2	10
Binärbaum	0	0	3	3	2	0	1	3	0	0	3	3
Computervirus	12	5	0	17	15	2	0	17	16	1	0	17
Dateisuche	1	4	0	5	2	1	2	5	2	1	2	5
Dino-Ordnung	2	3	1	6	4	1	1	6	6	0	0	6
Endlosschleife	0	1	11	12	3	8	1	12	1	6	5	12
Falschgeld	2	7	2	11	6	5	0	11	2	5	4	11
Fenster schließen	4	1	0	5	4	1	0	5	4	1	0	5
Labyrinth	0	2	1	3	1	1	1	3	2	1	0	3
Link	4	1	0	5	5	0	0	5	5	0	0	5
Links um!	0	4	1	5	4	1	0	5	5	0	0	5
Morse-Code	0	1	4	5	3	2	0	5	2	3	0	5
Netzwerkkabel	1	3	8	12	5	7	0	12	3	7	2	12
Platzwechsel	1	1	2	4	1	2	1	4	2	2	0	4
POP und PUSH	1	3	1	5	4	1	0	5	4	1	0	5
Primärschlüssel	1	2	6	9	8	1	0	9	4	5	0	9
Private E-Mail	3	1	0	4	4	0	0	4	4	0	0	4
Schnitzeljagd	0	1	14	15	8	6	1	15	1	6	8	15
Sicheres Passwort	7	2	0	9	9	0	0	9	8	1	0	9
Umparken	0	8	2	10	8	0	2	10	5	5	0	10
Ungeschützter PC	9	2	0	11	9	2	0	11	11	0	0	11
Verschlüsselung	0	1	7	8	1	6	1	8	3	5	0	8
Verwandlung	0	0	5	5	2	2	1	5	4	0	1	5
Wertetausch	1	3	1	5	1	2	2	5	1	1	3	5
Wetter	3	2	0	5	4	1	0	5	4	0	1	5
Zahlenreihe	2	0	8	10	10	0	0	10	7	3	0	10
<i>Summe:</i>	56	73	94	223	139	68	17	224	126	66	32	224

Tabelle C.2: Bewertungsverteilung, Teil 2: Komplexitätsgrad (KG; 1 = einfach, 2 = mittel, 3 = komplex), Redundanz (RG; 1 = redundant, 2 = nicht redundant), Anforderungsbereich (AB; 1 = Wiedergabe, 2 = Anwendung, 3 = Problemlösung). n ist jeweils die Anzahl der Experten, die eine Bewertung abgaben.

Aufgabe	KG				RG			AB			
	1	2	3	n	1	2	n	1	2	3	n
Biber am Fluss	3	2	0	5	1	4	5	0	3	2	5
Biber und Bisons	3	5	1	9	0	9	9	0	2	7	9
Bibers Geheimcode	5	5	0	10	1	10	11	0	10	1	11
Biberzahlen	2	8	0	10	5	5	10	0	9	2	10
Binärbaum	1	2	0	3	1	2	3	0	3	1	3
Computervirus	16	0	0	16	2	14	16	11	4	2	15
Dateisuche	2	2	1	5	0	5	5	1	4	0	5
Dino-Ordnung	3	3	0	6	0	6	6	0	4	2	6
Endlosschleife	4	7	1	12	0	12	12	0	10	2	11
Falschgeld	4	7	0	11	5	6	11	0	9	2	11
Fenster schließen	5	0	0	5	0	5	5	5	0	0	5
Labyrinth	0	2	1	3	2	1	3	0	1	1	2
Link	5	0	0	5	0	5	5	5	0	0	5
Links um!	4	1	0	5	1	4	5	0	4	1	5
Morse-Code	0	3	2	5	0	5	5	0	4	2	4
Netzwerkkabel	2	8	2	12	1	11	12	0	5	9	12
Platzwechsel	1	3	0	4	2	2	4	0	1	3	4
POP und PUSH	1	3	1	5	0	5	5	0	4	1	5
Primärschlüssel	6	3	0	9	1	8	9	1	9	0	9
Private E-Mail	4	0	0	4	2	2	4	4	1	0	4
Schnitzeljagd	4	7	4	15	0	15	15	0	11	5	15
Sicheres Passwort	7	2	0	9	0	9	9	2	6	1	9
Umparken	6	3	1	10	3	7	10	0	8	2	10
Ungeschützter PC	7	4	0	11	0	11	11	6	5	1	11
Verschlüsselung	0	7	1	8	3	5	8	0	6	3	8
Verwandlung	0	4	1	5	0	5	5	0	1	4	5
Wertetausch	1	3	1	5	0	5	5	0	4	2	5
Wetter	4	1	0	5	0	5	5	0	5	0	5
Zahlenreihe	5	4	1	10	0	10	10	1	9	0	10
<i>Summe:</i>	105	99	18	222	30	193	223	36	142	56	219

Tabelle C.3: Bewertungsverteilung, Teil 3: Prozessbereich (PB; 1 = Modellieren und Implementieren, 2 = Begründen und Bewerten, 3 = Strukturieren und Vernetzen, 4 = Kommunizieren und Kooperieren, 5 = Darstellen und Interpretieren). n ist die Anzahl der Experten, die eine Bewertung abgaben.

Aufgabe	PB					n
	1	2	3	4	5	
Biber am Fluss	2	2	2	0	3	5
Biber und Bisons	3	3	4	0	1	9
Bibers Geheimcode	2	0	7	0	3	11
Biberzahlen	3	1	3	0	8	10
Binärbaum	0	0	0	0	2	2
Computervirus	0	14	2	3	3	16
Dateisuche	2	2	3	0	1	5
Dino-Ordnung	0	4	3	0	2	6
Endlosschleife	8	5	5	0	6	12
Falschgeld	5	4	0	0	6	11
Fenster schließen	1	0	0	0	2	3
Labyrinth	1	1	2	0	2	3
Link	0	1	0	4	0	4
Links um!	4	1	0	0	2	5
Morse-Code	2	2	3	0	2	5
Netzwerkkabel	5	2	7	0	9	12
Platzwechsel	2	0	1	0	1	4
POP und PUSH	2	0	1	0	3	4
Primärschlüssel	3	8	1	0	1	9
Private E-Mail	0	1	0	1	0	2
Schnitzeljagd	7	3	4	0	7	15
Sicheres Passwort	1	7	1	1	1	8
Umparken	7	1	2	0	5	9
Ungeschützter PC	1	8	4	0	1	11
Verschlüsselung	3	3	3	0	4	7
Verwandlung	0	2	0	0	3	4
Wertetausch	2	2	2	0	2	5
Wetter	0	3	3	0	1	5
Zahlenreihe	4	0	3	0	2	8
<i>Summe:</i>	70	80	66	9	83	210

Tabelle C.4: Bewertungsverteilung, Teil 4: Lernzielstufe (LZ; 1 = Erinnern, 2 = Verstehen, 3 = Anwenden, 4 = Analysieren, 5 = Bewerten, 6 = Erschaffen) und Art des Wissens (AW; 1 = Fakten, 2 = Konzepte, 3 = Prozeduren, 4 = Metakognition. n ist jeweils die Anzahl der Experten, die eine Bewertung abgaben.

Aufgabe	LZ							AW				
	1	2	3	4	5	6	n	1	2	3	4	n
Biber am Fluss	0	2	2	3	1	0	5	0	3	3	0	5
Biber und Bisons	0	5	1	8	1	1	9	2	6	4	3	9
Bibers Geheimcode	0	5	10	3	0	0	11	1	5	5	0	11
Biberzahlen	1	2	7	4	0	1	10	1	4	7	0	10
Binärbaum	0	1	2	2	0	0	3	0	3	0	0	3
Computervirus	5	9	2	2	6	0	16	15	5	1	1	17
Dateisuche	0	3	5	2	1	0	5	0	2	3	0	5
Dino-Ordnung	0	4	1	4	2	0	6	0	4	1	1	6
Endlosschleife	0	3	10	6	1	0	12	1	8	7	0	12
Falschgeld	1	4	8	3	1	0	11	2	5	6	0	11
Fenster schließen	5	0	1	0	0	0	5	4	1	1	0	4
Labyrinth	0	0	1	2	0	0	2	0	2	2	0	3
Link	4	2	0	0	0	0	5	5	1	0	0	5
Links um!	0	2	5	0	0	0	5	1	2	2	0	5
Morse-Code	0	1	0	4	0	0	4	1	2	4	0	5
Netzwerkkabel	0	5	4	11	1	1	12	1	10	4	1	12
Platzwechsel	0	1	2	2	0	0	4	1	1	2	0	4
POP und PUSH	0	1	4	2	0	0	5	0	3	3	0	5
Primärschlüssel	2	6	1	2	0	0	9	5	5	1	0	9
Private E-Mail	4	0	0	0	0	0	4	4	1	0	0	4
Schnitzeljagd	0	5	8	8	0	0	15	0	7	9	1	15
Sicheres Passwort	2	3	1	1	5	0	9	6	4	0	0	9
Umparken	0	3	5	3	0	1	10	0	5	5	0	10
Ungeschützter PC	5	7	1	2	4	0	11	8	5	1	1	11
Verschlüsselung	0	3	6	5	1	0	8	0	6	2	1	8
Verwandlung	0	0	2	5	0	0	5	0	0	4	1	5
Wertetausch	1	2	4	2	1	0	5	1	3	3	0	5
Wetter	0	3	3	0	1	0	5	0	3	2	0	5
Zahlenreihe	1	2	9	1	0	0	10	1	1	8	1	10
<i>Summe:</i>	31	84	105	87	26	4	221	60	107	90	11	223

Tabelle C.5: Merkmalskorrelation in Stufe 5-7. Phi-Koeffizient (Φ). Merkmale Abstraktionsgrad (AG), Formalisierungsgrad (FG), Komplexitätsgrad (KG), Anforderungsbereiche Wiedergabe (W), Anwendung (A), Problemlösung (P), Prozessbereiche Modellieren und Implementieren (MI), Begründen und Bewerten (BB), Strukturieren und Vernetzen (SV), Darstellen und Interpretieren (DI), Lernzielstufen Verstehen (V), Anwenden (Aw), Analysieren (An), Wissensarten Fakten (F), Konzepte (K), Prozeduren (P). Positive Korrelationen auf dem Signifikanzniveau 0,01 sind hervorgehoben.

	Gestaltung			Anforderung			Prozessbereich			Lernzielstufe			Wissensart			
	AG	FG	KG	W	A	P	MI	BB	SV	DI	V	Aw	An	F	K	P
AG	1	,28	,19	-,26	-,12	,68	,18	-,34	-,23	,27	-,26	,00	,65	-,38	,16	,53
FG	,28	1	,44	-,33	,03	,53	,64	-,34	-,23	,27	-,33	,52	,44	-,49	,00	,68
KG	,19	,44	1	-,43	,20	,38	-,05	-,14	,65	-,33	,10	,38	,53	-,53	-,08	,39
W	-,26	-,33	-,43	1	-,85	-,16	-,40	,19	-,40	,06	,21	-,65	-,36	,84	-,44	-,62
A	-,12	,03	,20	-,85	1	-,38	,14	-,07	,47	,03	-,10	,47	,11	-,68	,53	,42
P	,68	,53	,38	-,16	-,38	1	,44	-,21	-,18	-,15	-,19	,25	,44	-,23	-,19	,31
MI	,18	,64	-,05	-,40	,14	,44	1	-,47	-,40	,06	-,47	,63	,06	-,49	,00	,68
BB	-,34	-,34	-,14	,19	-,07	-,21	-,47	1	,19	-,39	,69	-,45	-,03	,31	,12	-,71
SV	-,23	-,23	,65	-,40	,47	-,18	-,40	,19	1	-,33	,52	,32	,06	-,49	,41	-,10
DI	,27	,27	-,33	,06	,03	-,15	,06	-,39	-,33	1	-,39	-,17	,15	-,10	,41	,29
V	-,26	-,33	,10	,21	-,10	-,19	-,47	,69	,52	-,39	1	-,19	-,11	,03	,16	-,42
Aw	,00	,52	,38	-,65	,47	,25	,63	-,45	,32	-,17	-,19	1	-,04	-,73	,05	,69
An	,65	,44	,53	-,36	,11	,44	,06	-,03	-,06	,15	-,11	-,04	1	-,53	,28	,39
F	-,38	-,49	-,53	,84	-,68	-,23	-,49	,31	-,49	-,10	,03	-,73	-,53	1	-,53	-,73
K	,16	,00	-,08	-,44	,53	-,19	,00	,12	,41	,41	,16	,05	,28	-,53	1	,05
P	,53	,68	,39	-,62	,42	,31	,68	-,71	-,10	,29	-,42	,69	,39	-,73	,05	1

Tabelle C.6: Merkmalskorrelation in Stufe 8-10. Phi-Koeffizient (Φ). Merkmale Abstraktionsgrad (AG), Formalisierungsgrad (FG), Komplexitätsgrad (KG), Anforderungsbereiche Wiedergabe (W), Anwendung (A), Problemlösung (P), Prozessbereiche Modellieren und Implementieren (MI), Begründen und Bewerten (BB), Strukturieren und Vernetzen (SV), Darstellen und Interpretieren (DI), Lernzielstufen Verstehen (V), Anwenden (Aw), Analysieren (An), Wissensarten Fakten (F), Konzepte (K), Prozeduren (P). Positive Korrelationen auf dem Signifikanzniveau 0,01 sind hervorgehoben.

	Gestaltung			Anforderung			Prozessbereich			Lernzielstufe			Wissensart			
	AG	FG	KG	W	A	P	MI	BB	SV	DI	V	Aw	An	F	K	P
AG	1	,29	,58	-,28	-,21	,56	,28	-,18	,18	,62	,00	,00	,58	-,47	,45	-,04
FG	,29	1	,44	-,48	,43	-,08	,18	-,35	-,10	,85	-,27	,44	,39	-,52	,45	,34
KG	,58	,44	1	-,48	,12	,32	-,05	-,54	,40	,69	-,27	,44	,67	-,73	,29	,17
W	-,28	-,48	-,48	1	-,65	-,15	-,28	,54	-,18	-,46	,29	-,48	-,32	,65	-,41	-,35
A	-,21	,43	,12	-,65	1	-,65	,37	-,32	-,27	,23	-,45	,74	-,12	-,30	,00	,55
P	,56	-,08	,32	-,15	-,65	1	-,19	-,23	,68	,27	,29	-,48	,48	-,26	,41	-,35
MI	,28	,18	-,05	-,28	,37	-,19	1	-,18	-,28	,28	-,28	,53	,28	-,50	,35	,48
BB	-,18	-,35	-,54	,54	-,32	-,23	-,18	1	-,34	-,54	,54	-,68	-,18	,84	-,17	-,71
SV	,18	-,10	,40	-,18	-,27	,68	-,28	-,34	1	-,03	-,18	-,10	,18	-,21	,30	-,26
DI	,62	,85	,69	-,46	,23	,27	,28	-,54	-,03	1	-,46	,54	,62	-,84	,51	,37
V	,00	-,27	-,27	,29	-,45	,29	-,28	,54	-,18	-,46	1	-,61	-,07	,44	,17	-,45
Aw	,00	,44	,44	-,48	,74	-,48	,53	-,68	-,10	,54	-,61	1	,11	-,73	,00	,75
An	,58	,39	,67	-,32	-,12	,48	,28	-,18	,18	,62	-,07	,11	1	-,55	,58	-,17
F	-,47	-,52	-,73	,65	-,30	-,26	-,50	,84	-,21	-,84	,44	-,73	-,55	1	-,32	-,55
K	,45	,45	,29	-,41	,00	,41	,35	-,17	,30	,51	,17	,00	,58	-,32	1	-,29
P	-,04	,34	,17	-,35	,55	-,35	,48	-,71	-,26	,37	-,45	,75	-,17	-,55	-,29	1

Tabelle C.7: Merkmalskorrelation in Stufe 11+. Phi-Koeffizient (Φ). Merkmale Abstraktionsgrad (AG), Formalisierungsgrad (FG), Komplexitätsgrad (KG), Anforderungsbereiche Wiedergabe (W), Anwendung (A), Problemlösung (P), Prozessbereiche Modellieren und Implementieren (MI), Begründen und Bewerten (BB), Strukturieren und Vernetzen (SV), Darstellen und Interpretieren (DI), Lernzielstufen Verstehen (V), Anwenden (Aw), Analysieren (An), Wissensarten Fakten (F), Konzepte (K), Prozeduren (P). Positive Korrelationen auf dem Signifikanzniveau 0,01 sind hervorgehoben.

	Gestaltung			Anforderung			Prozessbereich			Lernzielstufe			Wissensart			
	AG	FG	KG	W	A	P	MI	BB	SV	DI	V	Aw	An	F	K	P
AG	1	,15	,52	-,41	-,15	,32	,26	-,10	,04	,69	-,32	,14	,43	-,52	,43	,00
FG	,15	1	,27	-,48	,58	-,29	,10	-,31	,10	,12	-,43	,33	,06	-,27	,06	,17
KG	,52	,27	1	-,78	,35	,35	,33	-,68	,33	,82	-,83	,54	,54	-1,0	,20	,61
W	-,41	-,48	-,78	1	-,56	-,28	-,26	,53	-,26	-,63	,65	-,42	-,42	,78	-,42	-,48
A	-,15	,58	,35	-,56	1	-,40	,41	-,48	,00	,25	-,53	,76	-,09	-,35	,19	,58
P	,32	-,29	,35	-,28	-,40	1	-,41	-,24	,82	,13	-,11	-,47	,66	-,35	,09	,00
MI	,26	,10	,33	-,26	,41	-,41	1	-,10	-,33	,41	-,33	,58	,19	-,33	,58	,10
BB	-,10	-,31	-,68	,53	-,48	-,24	-,10	1	-,49	-,48	,68	-,51	-,17	,68	-,17	-,66
SV	,04	,10	,33	-,26	,00	,82	-,33	-,49	1	,00	-,33	-,19	,58	-,33	,19	,10
DI	,69	,12	,82	-,63	,25	,13	,41	-,48	,00	1	-,82	,71	,35	-,82	,35	,48
V	-,32	-,43	-,83	,65	-,53	-,11	-,33	,68	-,33	-,82	1	-,65	-,34	,83	-,04	-,74
Aw	,14	,33	,54	-,42	,76	-,47	,58	-,51	-,19	,71	-,65	1	-,07	-,54	,20	,60
An	,43	,06	,54	-,42	-,09	,66	,19	-,17	,58	,35	-,34	-,07	1	-,54	,20	,06
F	-,52	-,27	-1,0	,78	-,35	-,35	-,33	,68	-,33	-,82	,83	-,54	-,54	1	-,20	-,61
K	,43	,06	,20	-,42	,19	,09	,58	-,17	,19	,35	-,04	,20	,20	-,20	1	-,22
P	,00	,17	,61	-,48	,58	,00	,10	-,66	,10	,48	-,74	,60	,06	-,61	-,22	1

Literaturverzeichnis

- [Anderson 2005] ANDERSON, L. W.: Objectives, Evaluation, and the Improvement of Education. *Studies in Educational Evaluation* 31 (2005) 2/3, S. 102–113
- [Anderson und Krathwohl et al. 2001] ANDERSON, L. W.; KRATHWOHL, D. R. (Hrsg.); AIRASIAN, P. W.; CRUIKSHANK, K. A.; MAYER, R. E.; PINTRICH, P. R.; RATHS, J.; WITTRICK, M. C.: *A Taxonomy for Learning, Teaching, and Assessing. A Revision of Bloom's Taxonomy of Educational Objectives*. Addison Wesley Longman, New York, 2001
- [Backhaus et al. 2008] BACKHAUS, K.; ERICHSON, B.; PLINKE, W.; WEIBER, R.: *Multivariate Analysemethoden. Eine anwendungsorientierte Einführung*. Springer, Berlin, zwölfte, vollständig überarbeitete Auflage, 2008
- [Barnett 1979] BARNETT, J.: The Study of Syntax Variables. In: Goldin, G. A.; McClintock, C. E. (Hrsg.): *TASK VARIABLES IN MATHEMATICAL PROBLEM SOLVING*. Mathematics Education Reports, ERIC Clearinghouse for Science, Mathematics, and Environmental Education, Ohio State University, Columbus, Ohio, 1979, S. 23–68
- [Bartholomew et al. 2008] BARTHOLOMEW, D. J.; STEELE, F.; MOUSTAKI, I.; GALBRAITH, J.: *Analysis of Multivariate Social Science Data*. Chapman & Hall/CRC Press, Boca Raton, Florida, Second Edition, 2008
- [Baumert et al. 2003] BAUMERT, J.; ARTELT, C.; KLIEME, E.; NEUBRAND, M.; PRENZEL, M.; SCHIEFELE, U.; SCHNEIDER, W.; SCHÜMER, G.; STANAT, P.; WEISS, M. (Hrsg.): *PISA 2000 – Ein differenzierter Blick auf die Länder der Bundesrepublik Deutschland. Zusammenfassung zentraler Befunde*. Max-Planck-Institut für Bildungsforschung, Berlin, 2003
- [Beck 1987] BECK, K.: *Die empirischen Grundlagen der Unterrichtsforschung. Eine kritische Analyse der deskriptiven Leistungsfähigkeit von Beobachtungsmethoden*. Hogrefe, Göttingen, 1987

- [Biggs und Collis 1982] BIGGS, J. B.; COLLIS, K. F.: *Evaluating the quality of learning: The SOLO taxonomy (Structure of the Observed Learning Outcome)*. Academic Press, New York, 1982
- [Bloom et al. 1972] BLOOM, B. S. (Hrsg.); ENGELHART, M. D.; FURST, E. J.; HILL, W. H.; KRATHWOHL, D. R.: *Taxonomie von Lernzielen im kognitiven Bereich*. Beltz, Weinheim, 1972 (Originaltitel: Taxonomy of Educational Objectives. Handbook 1: Cognitive Domain. 1956)
- [Bortz 2005] BORTZ, J.: *Statistik für Human- und Sozialwissenschaftler*. Springer Medizin, Heidelberg, sechste, vollständig überarbeitete und aktualisierte Auflage, 2005
- [Bortz und Döring 2006] BORTZ, J.; DÖRING, N.: *Forschungsmethoden und Evaluation*. Springer Medizin, Heidelberg, 4., überarbeitete Auflage, 2006
- [Bower 2008] BOWER, M.: A Taxonomy of Task Types in Computing. In: ACM (Hrsg.): *Proceedings of the 13th ITiCSE*. Madrid, Spanien, 2008, S. 281–285
- [Brinda 2004] BRINDA, T.: *Didaktisches System für objektorientiertes Modellieren im Informatikunterricht der Sekundarstufe II*. Dissertation an der Universität Siegen. Siegen, 2004
- [Bruner 1966] BRUNER, J. S.: On Cognitive Growth I; On Cognitive Growth II. In: Bruner, J. S.; Olver, R. R.; Greenfield, P. M. (Hrsg.): *Studies in Cognitive Growth*. John Wiley & Sons, New York, 1966, S. 1–29; 30–67
- [Bühner 2006] BÜHNER, M.: *Einführung in die Test- und Fragebogenkonstruktion*. Pearson Studium, München, 2006
- [Carstensen 2000] CARSTENSEN, C. H.: *Mehrdimensionale Testmodelle mit Anwendungen aus der pädagogisch-psychologischen Diagnostik*. Institut für die Pädagogik der Naturwissenschaften (IPN), Kiel, 2000
- [Carstensen und Rost 2000] CARSTENSEN, C. H.; ROST, J.: *MULTIRA ein Programmsystem zur Analyse mehrdimensionaler Rasch-Modelle. Handbuch zum Computerprogramm MULTIRA*. Institut für die Pädagogik der Naturwissenschaften (IPN), Kiel, 2000
- [Chomsky 1970] CHOMSKY, N.: *Sprache und Geist*. In: Blumenberg, H.; Habermas, J.; Henrich, D.; Taubes, J. (Hrsg.): *Theorie*. Suhrkamp, Frankfurt am Main, 1970 (Originaltitel: Language and Mind. 1968)
- [Claus und Schwill 2003] CLAUS, V.; SCHWILL, A.: *Duden Informatik. Ein Fachlexikon für Studium und Praxis*. Dudenverlag, Mannheim, 2003

- [Cohors-Fresenborg 1996] COHORS-FRESENBORG, E.: Mathematik als Werkzeug zur Wissensrepräsentation. In: Kadunz, G.; Kautschitsch, H.; Ossimitz, G.; Schneider, E. (Hrsg.): *Trends und Perspektiven*. Schriftenreihe der Mathematik, Band 23, Hölder-Pichler-Tempsky, Wien, 1996, S. 85–90
- [Cohors-Fresenborg et al. 2004] COHORS-FRESENBORG, E.; SJUTS, J.; SOMMER, N.: Komplexität von Denkvorgängen und Formalisierung von Wissen. In: Neubrand, M. (Hrsg.): *Mathematische Kompetenzen von Schülerinnen und Schülern in Deutschland. Vertiefende Analysen im Rahmen von PISA 2000*. Verlag für Sozialwissenschaften, Wiesbaden, 2004, S. 109–144
- [Dagienė und Futschek 2008] DAGIENĖ, V.; FUTSHEK, G.: Bebras International Contest on Informatics and Computer Literacy: Criteria for Good Tasks. In: Mittermeir R. T.; Syslo, M. M. (Hrsg.): *Informatics Education – Supporting Computational Thinking. ISSEP 2008, 3rd International Conference on Informatics in Secondary Schools*. Springer, Berlin, Heidelberg, 2008, S. 19–30
- [Denning 2003] DENNING, P.: Great Principles of Computing. *Communications of the ACM* 46 (2003) 11, S. 15–20
- [Denning 2004] DENNING, P.: Great Principles of Computing Curricula. In: *Proceedings of the 35th SIGCSE Technical Symposium on Computer Science Education*. ACM Press, New York, 2004, S. 336–341
- [Draxler 2006] DRAXLER D.: *Aufgabendesign und basismodellorientierter Physikunterricht*. Dissertation am Fachbereich Physik der Universität Duisburg-Essen, 2006
- [Ebel und Frisbie 1991] EBEL, R. L.; FRISBIE, D. A.: *Essentials of Educational Measurement*. Prentice Hall, Englewood Cliffs, N. J., 1991
- [Fleischer et al. 2010] FLEISCHER, J.; WIRTH, J.; RUMANN, S.; LEUTNER, D.: Strukturen fächerübergreifender und fachlicher Problemlösekompetenz. Analyse von Aufgabenprofilen. In: Klieme, E.; Leutner, D.; Kenk, M. (Hrsg.): *Kompetenzmodellierung. Zwischenbilanz des DFG-Schwerpunktprogramms und Perspektiven des Forschungsansatzes*. Zeitschrift für Pädagogik 56 (2010) Beiheft, S. 239–248
- [Freischlad 2009] FREISCHLAD, S.: *Entwicklung und Erprobung des Didaktischen Systems Internetworking im Informatikunterricht*. Dissertation an der Universität Siegen. Commentarii informaticae didacticae, Band 3, Universitätsverlag Potsdam, Potsdam, 2009
- [Frey et al. 2010] FREY, A.; HEINZE, A.; MILDNER, D.; HOCHWEBER, J.; ASSEBURG, R.: Mathematische Kompetenz von PISA 2003 bis PISA 2009. In:

- Klieme, E.; Artelt, C.; Hartig, J.; Jude, N.; Köller, O.; Prenzel, M.; Schneider, W.; Stanat, P. (Hrsg.): *PISA 2009. Bilanz nach einem Jahrzehnt*. Waxmann, Münster, 2010, S. 153–176
- [Friedrich 2003] FRIEDRICH, S.: Informatik und PISA – vom Wehe zum Wohl der Schulinformatik. In: Hubwieser, P. (Hrsg.): *Informatische Fachkonzepte im Unterricht. INFOS 2003, 10. Fachtagung Informatik und Schule der Gesellschaft für Informatik (GI) 2003 in Garching bei München*. Köllen, Bonn, 2003, S. 133–144
- [Fuller et al. 2007] FULLER, U.; JOHNSON, C. G.; AHONIEMI, T.; CUKIERMAN, D.; HERNÁN-LOSADA, I.; JACKOVA, J.; LAHTINEN, E.; LEWIS, T. L.; MCGEE THOMPSON, D.; RIEDESEL, C.; THOMPSON, E.: Developing a Computer Science-specific Learning Taxonomy. *ACM SIGCSE Bulletin* 39 (2007) 4, ACM Press, New York, S. 152–170
- [GI 2000] GESELLSCHAFT FÜR INFORMATIK E. V. (GI): Empfehlungen für ein Gesamtkonzept zur informatischen Bildung an allgemein bildenden Schulen. Beilage zu *LOG IN* 20 (2000) 2
- [GI 2004] GESELLSCHAFT FÜR INFORMATIK E. V. (GI, Hrsg.): *Digitale Spaltung verhindern – Schulinformatik stärken!* Memorandum. Ulm, 2004
- [GI 2008] GESELLSCHAFT FÜR INFORMATIK E. V. (GI): Grundsätze und Standards für die Informatik in der Schule. Bildungsstandards Informatik für die Sek. I. Beilage zu *LOG IN* 28 (2008) 150/151
- [Goldin 1979] GOLDIN, G. A.: Structure Variables in Problem Solving. In: Goldin, G. A.; McClintock, C. E. (Hrsg.): *TASK VARIABLES IN MATHEMATICAL PROBLEM SOLVING*. Mathematics Education Reports, ERIC Clearinghouse for Science, Mathematics, and Environmental Education, Ohio State University, Columbus, Ohio, 1979, S. 103–169
- [Goldin und Caldwell 1979] GOLDIN, G. A.; CALDWELL, J. H.: Syntax, Content, and Context Variables Examined in a Research Study. In: Goldin, G. A.; McClintock, C. E. (Hrsg.): *TASK VARIABLES IN MATHEMATICAL PROBLEM SOLVING*. Mathematics Education Reports, ERIC Clearinghouse for Science, Mathematics, and Environmental Education, Ohio State University, Columbus, Ohio, 1979, S. 235–276
- [Goldin und McClintock 1979] GOLDIN, G. A.; MCCLINTOCK, C. E. (Hrsg.): *TASK VARIABLES IN MATHEMATICAL PROBLEM SOLVING*. Mathematics Education Reports, ERIC Clearinghouse for Science, Mathematics, and Environmental Education, Ohio State University, Columbus, Ohio, 1979.

- [Greiff und Funke 2010] GREIFF, S.; FUNKE, J.: Systematische Erforschung komplexer Problemlösefähigkeit anhand minimal komplexer Systeme. In: Klieme, E.; Leutner, D.; Kenk, M. (Hrsg.): Kompetenzmodellierung. Zwischenbilanz des DFG-Schwerpunktprogramms und Perspektiven des Forschungsansatzes. *Zeitschrift für Pädagogik* 56 (2010) Beiheft, S. 216–227
- [Gut et al. 2010] GUT, C.; LABUDDE, P.; RAMSEIER, E.: Large-scale Experimentiertests: Ansätze zur Analyse von Itemschwierigkeiten. In: Höttecke, D. (Hrsg.): *Entwicklung naturwissenschaftlichen Denkens zwischen Phänomen und Systematik. Jahrestagung der Gesellschaft für Didaktik der Chemie und Physik in Dresden 2009*. Lit Verlag, Münster, 2010, S. 245–247
- [Hansky 2010] HANSKY, S.: *Typische Schülerfehler bei Informatikaufgaben*. Zulassungsarbeit zum Ersten Staatsexamen, Department Informatik, Universität Erlangen-Nürnberg, Erlangen, 2010
- [Hartig und Höhler 2010] HARTIG, J.; HÖHLER, J.: Modellierung von Kompetenzen mit mehrdimensionalen IRT-Modellen. In: Klieme, E.; Leutner, D.; Kenk, M. (Hrsg.): Kompetenzmodellierung. Zwischenbilanz des DFG-Schwerpunktprogramms und Perspektiven des Forschungsansatzes. *Zeitschrift für Pädagogik* 56 (2010) Beiheft, S. 189–198
- [Hartig und Jude 2007] HARTIG, J.; JUDE, N.: Empirische Erfassung von Kompetenzen und psychometrische Kompetenzmodelle. In: Hartig, J.; Klieme, E. (Hrsg.): *Möglichkeiten und Voraussetzungen technologiebasierter Kompetenzdiagnostik. Expertise im Auftrag des Bundesministeriums für Bildung und Forschung (BMBF)*. Bonn, 2007
- [Hartig und Klieme 2006] HARTIG, J.; KLIEME, E.: Kompetenz und Kompetenzdiagnostik. In: Schweizer, K. (Hrsg.): *Leistung und Leistungsdiagnostik*. Springer Medizin, Heidelberg, 2006, S. 127–143
- [Hartmann et al. 2006] HARTMANN, W.; NÄF, M.; REICHERT, R.: *Informatikunterricht planen und durchführen*. Springer, Berlin, 2006
- [Hiebert 1986] HIEBERT, J. (Hrsg.): *Conceptual and procedural knowledge: The case of mathematics*. Erlbaum, Hillsdale, NJ, 1986
- [Hubwieser 2007] HUBWIESER, P.: *Didaktik der Informatik*. Springer, Berlin, 3., überarbeitete und erweiterte Auflage, 2007
- [Kauertz 2008] KAUERTZ, A.: *Schwierigkeitserzeugende Merkmale physikalischer Leistungstestaufgaben*. Dissertation. Studien zum Physik- und Chemielernen, Band 79. Logos, Berlin, 2008

- [Klafki 1985] KLAFKI, W.: *Neue Studien zur Bildungstheorie und Didaktik. Zeitgemäße Allgemeinbildung und kritisch-konstruktive Didaktik*. Beltz, Weinheim, 1985 (6., neu ausgestattete Auflage 2007)
- [Klauer 1987] KLAUER, K. J.: *Kriteriumsorientierte Tests*. Hogrefe, Göttingen, 1987
- [Klieme et al. 2003] KLIEME, E.; AVENARIUS, H.; BLUM, W.; DÖBRICH, P.; GRUBER, H.; PRENZEL, M.; REISS, K.; RIQUARTS, K.; ROST, J.; TENORTH, H.-E.; VOLLMER, H. J.: *Zur Entwicklung nationaler Bildungsstandards. Eine Expertise*. Bundesministerium für Bildung und Forschung (BMBF, Hrsg.), Berlin, 2003
- [Klieme et al. 2010] KLIEME, E.; LEUTNER, D.; KENK, M. (Hrsg.): Kompetenzmodellierung. Zwischenbilanz des DFG-Schwerpunktprogramms und Perspektiven des Forschungsansatzes. *Zeitschrift für Pädagogik 56 (2010) Beiheft*
- [Klieme und Leutner 2006] KLIEME, E.; LEUTNER, D. (Hrsg.): Kompetenzmodelle zur Erfassung individueller Lernergebnisse und zur Bilanzierung von Bildungsprozessen. Beschreibung eines neu eingerichteten Schwerpunktprogramms der DFG. *Zeitschrift für Pädagogik 52 (2006) 6*, S. 876–903
- [Kleinknecht et al. 2011] KLEINKNECHT, M.; MAIER, U.; METZ, K.; BOHL, T.: Analyse des kognitiven Aufgabenpotenzials. Entwicklung und Erprobung eines allgemeindidaktischen Auswertungsmanuals. *Unterrichtswissenschaft 39 (2011) 4*, S. 163–173
- [KMK 2004a] SEKRETARIAT DER STÄNDIGEN KONFERENZ DER KULTUSMINISTER DER LÄNDER IN DER BUNDESREPUBLIK DEUTSCHLAND (KMK, Hrsg.): *Bildungsstandards im Fach Mathematik für den Mittleren Schulabschluss. Beschluss vom 4.12.2003*. Wolters Kluwer, München, 2004
- [KMK 2004b] SEKRETARIAT DER STÄNDIGEN KONFERENZ DER KULTUSMINISTER DER LÄNDER IN DER BUNDESREPUBLIK DEUTSCHLAND (KMK, Hrsg.): *Einheitliche Prüfungsanforderungen in der Abiturprüfung Informatik. (Beschluss der Kultusministerkonferenz vom 1.12.1989 i.d.F. vom 5.2.2004)*. Wolters Kluwer, München, 2004
- [KMK 2005] SEKRETARIAT DER STÄNDIGEN KONFERENZ DER KULTUSMINISTER DER LÄNDER IN DER BUNDESREPUBLIK DEUTSCHLAND (KMK, Hrsg.): *Bildungsstandards der Kultusministerkonferenz. Erläuterungen zur Konzeption und Entwicklung*. Wolters Kluwer, München, 2005

- [Krippendorff 2011] KRIPPENDORFF, K.: *Computing Krippendorffs Alpha-Reliability*. Annenberg School for Communication, Departmental Papers (ASC), University of Pennsylvania. 2011. Online unter http://repository.upenn.edu/asc_papers/43, Stand 25.01.2011, abgerufen am 5. Mai 2014
- [Krusch 2012] KRUSCH, P.: *Analyse des Zusammenhangs zwischen Informatikunterricht und dem Erfolg in einem Informatikwettbewerb*. Zulassungsarbeit zum Ersten Staatsexamen, Department Informatik, Universität Erlangen-Nürnberg, Erlangen, 2012
- [Kujath 2007] KUJATH, B.: Vergleichende Analysen zweier Problemlöseprozesse unter dem Aspekt des Problemlöseerfolgs. In: Schubert, S. (Hrsg.): *Didaktik der Informatik in Theorie und Praxis. 12. GI-Fachtagung „Informatik und Schule – INFOS 2007“ in Siegen*. Köllen, Bonn, 2007, S. 295–306
- [Kulm 1979] KULM, G.: The Classification of Problem-Solving Research Variables. In: Goldin, G. A.; McClintock, C. E. (Hrsg.): *TASK VARIABLES IN MATHEMATICAL PROBLEM SOLVING*. Mathematics Education Reports, ERIC Clearinghouse for Science, Mathematics, and Environmental Education, Ohio State University, Columbus, Ohio, 1979, S. 1–21
- [Kunina-Habenicht et al. 2010] KUNINA-HABENICHT, O.; WILHELM, O.; MATTHES, F.; RUPP, A.: Kognitive Diagnosemodelle: Theoretisches Potential und methodische Probleme. In: Klieme, E.; Leutner, D.; Kenk, M. (Hrsg.): Kompetenzmodellierung. Zwischenbilanz des DFG-Schwerpunktprogramms und Perspektiven des Forschungsansatzes. *Zeitschrift für Pädagogik* 56 (2010) Beiheft, S. 75–85
- [Lahtinen 2007] LAHTINEN, E.: A Categorization of Novice Programmers: A Cluster Analysis Study. In: Sajaniemi, J.; Tukiainen, M. (Hrsg.): *Proceedings of the 19th annual Workshop of the Psychology of Programming Interest Group*. Joensuu, Finland, 2007, S. 32–41
- [Lefrancois 1986] LEFRANCOIS, G. R.: *Psychologie des Lernens*. Springer-Verlag, Berlin, 1986 (Originaltitel: Psychological Theories and Human Learning: Kongors Report. 1972)
- [Linck et al. 2013] LINCK, B.; MAGENHEIM, J.; NELLES, W.; NEUGEBAUER, J.; OHRNDORF, L.; SCHAPER, N.; SCHUBERT, S.; STECHERT, P.: Competence model for informatics modelling and system comprehension. In: *Proceedings of the 4th global engineering education conference, IEEE EDUCON*. Berlin, Germany, 2013, S. 85–90

- [Lind und Knoche 2004] LIND, D.; KNOCHE, D.: Testtheoretische Modelle und Verfahren bei PISA-2000-Mathematik. In: Neubrand, M. (Hrsg.): *Mathematische Kompetenzen von Schülerinnen und Schülern in Deutschland. Vertiefende Analysen im Rahmen von PISA 2000*. Verlag für Sozialwissenschaften, Wiesbaden, 2004, S. 51–69
- [Magenheim 2005] MAGENHEIM, J.: Towards a Competence Model for Educational Standards of Informatics. In: *Proceedings of the 8th IFIP World Conference on Computers in Education*. Cape Town, 2005
- [Marzano und Kendall 2007] MARZANO, R. J.; KENDALL, J. S.: *The new taxonomy of educational objectives*. Corwin Press, Thousand Oaks, CA, 2007
- [McClintock 1979] MCCLINTOCK, C. E.: Heuristic Processes as Task Variables. In: Goldin, G. A.; McClintock, C. E. (Hrsg.): *TASK VARIABLES IN MATHEMATICAL PROBLEM SOLVING*. Mathematics Education Reports, ERIC Clearinghouse for Science, Mathematics, and Environmental Education, Ohio State University, Columbus, Ohio, 1979, S. 171–234
- [Meinl 2012] MEINL, F.: *Merkmale mädchen- bzw. jungeneigneter Informatikaufgaben im Schülerwettbewerb Informatik-Biber*. Zulassungsarbeit zum Ersten Staatsexamen, Department Informatik, Universität Erlangen-Nürnberg, Erlangen, 2012
- [Meyerhöfer 2004] MEYERHÖFER, W.: Zum Kompetenzstufenmodell von PISA. *Journal für Mathematik-Didaktik* 25 (2004) 3/4, S. 294–305. Online unter http://www.math.uni-potsdam.de/prof/o_didaktik/am/Verroe, abgerufen am 8. August 2012
- [Naumann et al. 2010] NAUMANN, J.; ARTELT, C.; SCHNEIDER, W.; STANAT, P.: Lesekompetenz von PISA 2000 bis PISA 2009. In: Klieme, E.; Artelt, C.; Hartig, J.; Jude, N.; Köller, O.; Prenzel, M.; Schneider, W.; Stanat, P.: *PISA 2009. Bilanz nach einem Jahrzehnt*. Waxmann, Münster, 2010, S. 23–71
- [Nelles et al. 2010] NELLES, W.; RHODE, T.; STECHERT, P.: Entwicklung eines Kompetenzrahmenmodells – Informatisches Modellieren und Systemverständnis. *Informatik Spektrum* 33 (2010) 1, Hauptbeitrag, S. 45–53
- [Neubrand et al. 2002] NEUBRAND, M.; KLIEME, E.; LÜDTKE, O.; NEUBRAND, J.: Kompetenzstufen und Schwierigkeitsmodelle für den PISA-Test zur mathematischen Grundbildung. *Unterrichtswissenschaft* 30 (2002) 2, S. 100–119
- [Neubrand et al. 2004] NEUBRAND, M.; BIEHLER, R.; BLUM, W.; COHORS-FRESENBORG, E.; FLADE, L.; KNOCHE, N.; LIND, D.; LÖDING, W.;

- MÖLLER, G.; WYNANDS, A.; NEUBRAND, J.: Der Prozess der Itementwicklung bei der nationalen Ergänzungsuntersuchung von PISA 2000: Vom Theoretischen Rahmen zu den konkreten Aufgaben. In: Neubrand, M. (Hrsg.): *Mathematische Kompetenzen von Schülerinnen und Schülern in Deutschland. Vertiefende Analysen im Rahmen von PISA 2000*. Verlag für Sozialwissenschaften, Wiesbaden, 2004, S. 31–49
- [Nickolaus 2011] NICKOLAUS, R.: Die Erfassung fachlicher Kompetenzen und ihrer Entwicklungen in der beruflichen Bildung – Forschungsstand und Perspektiven. In: Zlatkin-Troitschanskaia, O. (Hrsg.): *Stationen Empirischer Bildungsforschung. Traditionslinien und Perspektiven*. Verlag für Sozialwissenschaften, Wiesbaden, 2011
- [OECD 2006] OECD (Organisation for economic co-operation and development, Hrsg.): *Assessing Scientific, Reading and Mathematical Literacy. A Framework for PISA 2006*. OECDpublishing, 2006
- [Oelkers und Reusser 2008] OELKERS, J.; REUSSER, K.: *Expertise: Qualität entwickeln – Standards sichern – mit Differenzen umgehen*. Bundesministerium für Bildung und Forschung (BMBF, Hrsg.), Bonn, 2008
- [PISA 2004] PRENZEL, M.; BAUMERT, J.; BLUM, W.; LEHMANN, R.; LEUTNER, D.; NEUBRAND, M.; PEKRUN, R.; ROST, J.; SCHIEFELE, U. (PISA-Konsortium Deutschland, Hrsg.): *PISA 2003: Ergebnisse des zweiten internationalen Vergleichs. Zusammenfassung*. Online unter http://pisa.ipn.uni-kiel.de/Zusammenfassung_2003.pdf, Stand 7.12.2004, abgerufen am 18. Juni 2012
- [PISA 2005] PRENZEL, M.; BAUMERT, J.; BLUM, W.; LEHMANN, R.; LEUTNER, D.; NEUBRAND, M.; PEKRUN, R.; ROLFF, H.-G.; ROST, J.; SCHIEFELE, U. (PISA-Konsortium Deutschland, Hrsg.): *PISA 2003: Ergebnisse des zweiten Ländervergleichs. Zusammenfassung*. Online unter http://pisa.ipn.uni-kiel.de/Pisa2003_E_Zusammenfassung.pdf, Stand 3.11.2005, abgerufen am 18. Juni 2012
- [PISA 2007] PRENZEL, M.; ARTELT, C.; BAUMERT, J.; BLUM, W.; HAMMANN, M.; KLIEME, E.; PEKRUN, R. (PISA-Konsortium Deutschland, Hrsg.): *PISA 2006: Die Ergebnisse der dritten internationalen Vergleichsstudie. Zusammenfassung*. Online unter http://pisa.ipn.uni-kiel.de/zusammenfassung_pisa2006.pdf, Stand 4.12.2007, abgerufen am 18. Juni 2012
- [Pohl et al. 2009] POHL, W.; SCHLÜTER, K.; HEIN, H.-W.: Informatik-Biber: Informatik-Einstieg und mehr. In: Koerber, B. (Hrsg.): *Zukunft braucht Herkunft. 25 Jahre „INFOS – Informatik und Schule“*. 13. GI-Fachtagung Informatik und Schule 2009 in Berlin. Köllen, Bonn, 2009, S. 38–49

- [Prenzel et al. 2002] PRENZEL, M.; HÄUSSLER, P.; ROST, J.; SENKBEIL, M.: Der PISA-Naturwissenschaftstest: Lassen sich die Aufgabenschwierigkeiten vorhersagen? *Unterrichtswissenschaft 30 (2002) 2*, S. 120–135
- [Puhlmann 2003] PUHLMANN, H.: Informatische Literalität nach dem PISA-Muster. In: Hubwieser, P. (Hrsg.): *Informatische Fachkonzepte im Unterricht. INFOS 2003, 10. Fachtagung Informatik und Schule der Gesellschaft für Informatik (GI) 2003 in Garching bei München*. Köllen, Bonn, 2003, S. 145–154
- [Rönnebeck et al. 2010] RÖNNEBECK, S.; SCHÖPS, K.; PRENZEL, M.; MILDNER, D.; HOCHWEBER, J.: Naturwissenschaftliche Kompetenz von PISA 2006 bis PISA 2009. In: Klieme, E.; Artelt, C.; Hartig, J.; Jude, N.; Köller, O.; Prenzel, M.; Schneider, W.; Stanat, P. (Hrsg.): *PISA 2009. Bilanz nach einem Jahrzehnt*. Waxmann, Münster, 2010, S. 177–198
- [Rost 2004a] ROST, J.: *Lehrbuch Testtheorie – Testkonstruktion*. Verlag Hans Huber, Bern, 2004
- [Rost 2004b] ROST, J.: Psychometrische Modelle zur Überprüfung von Bildungsstandards anhand von Kompetenzmodellen. *Zeitschrift für Pädagogik 20 (2004) 5*, S. 662–678
- [Roth 1971] ROTH, H.: *Pädagogische Anthropologie. Band II, Entwicklung und Erziehung*. Herrmann Schroedel, Hannover, 1971
- [Schaper 2008] SCHAPER, N.: (Arbeits-)psychologische Kompetenzforschung. In: Fischer, M., Spöttl, G. (Hrsg.): *Forschungsperspektiven in Facharbeit und Berufsbildung. Strategien und Methoden der Berufsbildungsforschung*. Peter Lang, Frankfurt am Main, 2008, S. 91–115
- [Schecker und Parchmann 2006] SCHECKER, H.; PARCHMANN, I.: Modellierung naturwissenschaftlicher Kompetenz. *Zeitschrift für Didaktik der Naturwissenschaften 12 (2006)*, S. 45–66
- [Schlüter 2008] SCHLÜTER, K.: Je schwieriger die Aufgabe, desto klüger der Kopf? In: Brinda, T.; Fothe, M.; Hubwieser, P.; Schlüter, K. (Hrsg.): *Didaktik der Informatik – Aktuelle Forschungsergebnisse. 5. Workshop der GI-Fachgruppe „Didaktik der Informatik“ in Erlangen*. Köllen, Bonn, 2008, S. 77–86
- [Schlüter 2009] SCHLÜTER, K.: Eine Studie zu den Merkmalen der Aufgabenschwierigkeit am Beispiel eines Informatik-Schülerwettbewerbs. Erster Teil: Aufgabenklassifizierung. In: Koerber, B. (Hrsg.): *Zukunft braucht Herkunft. 25 Jahre „INFOS – Informatik und Schule“. 13. GI-Fachtagung Informatik und Schule 2009 in Berlin*. Köllen, Bonn, 2009, S. 181–192

- [Schlüter 2010] SCHLÜTER, K.: Eine Studie zu den Merkmalen der Aufgabenschwierigkeit am Beispiel eines Informatik-Schülerwettbewerbs. Zweiter Teil: Empirische Aufgabenanalyse. In: Diethelm, I.; Dörge, C.; Hildebrandt, C.; Schulte, C. (Hrsg.): *Didaktik der Informatik – Möglichkeiten empirischer Forschungsmethoden und Perspektiven der Fachdidaktik. 6. Workshop der GI-Fachgruppe „Didaktik der Informatik“ in Oldenburg*. Köllen, Bonn, 2010, S. 69–80
- [Schlüter und Brinda 2008] SCHLÜTER, K.; BRINDA, T.: From exercise characteristics to competence dimensions – exemplified by theoretical computer science in secondary education. In: IFIP (Hrsg.): *Proceedings of LYICT 2008*. Kuala Lumpur, Malaysia, 2008
- [Schmidt 2012] SCHMIDT, J.: *Merkmale altersgemäßer Informatikaufgaben am Beispiel eines Schülerwettbewerbs*. Zulassungsarbeit zum Ersten Staatsexamen, Department Informatik, Universität Erlangen-Nürnberg, Erlangen, 2012
- [Schubert 2009] SCHUBERT, S.: Aus Unterrichtsbeispielen lernen – Informatikdidaktische Partnerschaften. In: Koerber, B. (Hrsg.): *Zukunft braucht Herkunft. 25 Jahre „INFOS – Informatik und Schule“*. 13. GI-Fachtagung Informatik und Schule 2009 in Berlin. Köllen, Bonn, 2009, S. 13–27
- [Schubert und Schwill 2004] SCHUBERT, S.; SCHWILL, A.: *Didaktik der Informatik*. Spektrum Akademischer Verlag, Heidelberg, Berlin, 2004
- [Schulte und Brinda 2005] SCHULTE, C.; BRINDA, T.: Beiträge der Objektorientierung zu einem Kompetenzmodell des informatischen Modellierens. In: Friedrich, S. (Hrsg.): *Unterrichtskonzepte für informatische Bildung. INFOS 2005, 11. Fachtagung Informatik und Schule der Gesellschaft für Informatik (GI) 2005 in Dresden*. Köllen, Bonn, 2005, S. 137–148
- [Schumann und Eberle 2011] SCHUMANN, S.; EBERLE, F.: Bedeutung und Verwendung schwierigkeitsbestimmender Aufgabenmerkmale für die Erfassung ökonomischer und beruflicher Kompetenzen. In: Faßhauer, U.; Fürstenau, B.; Wuttke, E. (Hrsg.): *Grundlagenforschung zum Dualen System und Kompetenzentwicklung in der Lehrerbildung*. Verlag Barbara Budrich, Opladen, 2011, S. 77–89
- [Schweizer 2006a] SCHWEIZER, K.: Klassische Leistungskonzepte. Intelligenz. In: Schweizer, K. (Hrsg.): *Leistung und Leistungsdiagnostik*. Springer Medizin, Heidelberg, 2006, S. 1–15
- [Schweizer 2006b] SCHWEIZER, K.: Klassische Leistungskonzepte. Intelligenzdiagnostik. In: Schweizer, K. (Hrsg.): *Leistung und Leistungsdiagnostik*. Springer Medizin, Heidelberg, 2006, S. 70–83

- [Schwill 1993] SCHWILL, A.: Fundamentale Ideen der Informatik. *Zentralblatt für Didaktik der Mathematik* 25 (1993) 1, S. 20–31
- [Senkbeil et al. 2005] SENKBEIL, M.; ROST, J.; CARSTENSEN, C. H., WALTER, O.: Der nationale Naturwissenschaftstest PISA 2003. Entwicklung und empirische Überprüfung eines zweidimensionalen Facettendesigns. *Empirische Pädagogik* 19 (2005) 2, S. 166–189
- [Sheard et al. 2008] SHEARD, J.; CARBONE, A.; LISTER, R.; SIMON, B.; THOMPSON, E.; WHALLEY, J. L.: Going SOLO to Assess Novice Programmers. In: ACM (Hrsg.): *Proceedings of the 13th ITiCSE*. Madrid, Spanien, 2008, S. 209–213
- [Starr et al. 2008] STARR, C. W.; MANARIS, B.; STALVEY, R. H.: Bloom's Taxonomy Revisited: Specifying Assessable Learning Objectives in Computer Science. *ACM SIGCSE Bulletin* 40 (2008) 1, ACM Press, New York, S. 261–265
- [Starruß 2010] STARRUSS, I.: *Synopse zum Informatikunterricht in Deutschland. Analyse der informatischen Bildung an allgemein bildenden Schulen auf der Basis der im Jahr 2010 gültigen Lehrpläne und Richtlinien*. Bakkalaureatsarbeit an der Technischen Universität Dresden. Dresden, 2010
- [Stechert 2009] STECHERT, P.: *Fachdidaktische Diskussion von Informatiksystemen und der Kompetenzentwicklung im Informatikunterricht*. Dissertation an der Universität Siegen. Commentarii informaticae didacticae, Band 2, Universitätsverlag Potsdam, Potsdam, 2009
- [Steinert 2007] STEINERT, M.: Lernzielgraphen und Lernerfolgsanalyse. In: Schubert, S. (Hrsg.): *Didaktik der Informatik in Theorie und Praxis. 12. GI-Fachtagung „Informatik und Schule – INFOS 2007“ in Siegen*. Köllen, Bonn, 2007, S. 147–170
- [Sternberg 1998] STERNBERG, R. J.: Metacognition, abilities, and developing expertise: What makes an expert student? *Instructional Science* 26 (1998), S. 127–140
- [Viering et al. 2010] VIERING, T.; FISCHER, H. E.; NEUMANN, K.: Die Entwicklung physikalischer Kompetenz in der Sekundarstufe I. In: Klie-me, E.; Leutner, D.; Kenk, M. (Hrsg.): *Kompetenzmodellierung. Zwischenbilanz des DFG-Schwerpunktprogramms und Perspektiven des Forschungsansatzes*. Zeitschrift für Pädagogik 56 (2010) Beiheft, S. 92–103
- [vom Hofe 1995] VOM HOFE, R.: *Grundvorstellungen mathematischer Inhalte*. Spektrum, Heidelberg, 1995

- [Webb 1979] WEBB, N.: Content and Context Variables in Problem Tasks. In: Goldin, G. A.; McClintock, C. E. (Hrsg.): *TASK VARIABLES IN MATHEMATICAL PROBLEM SOLVING*. Mathematics Education Reports, ERIC Clearinghouse for Science, Mathematics, and Environmental Education, Ohio State University, Columbus, Ohio, 1979, S. 69–102
- [Weimer 1925] WEIMER, H.: *Psychologie der Fehler*. Julius Klinkhardt, Leipzig, 1925
- [Weinert 1999] WEINERT, F. E.: *Definition and Selection of Competencies. Concepts of Competence*. Max Planck Institute for Psychological Research, München, April 1999
- [Weinert 2001] WEINERT, F. E.: Vergleichende Leistungsmessung in Schulen – eine umstrittene Selbstverständlichkeit. In: Weinert, F. E. (Hrsg.): *Leistungsmessungen in Schulen*. Beltz, Weinheim, 2001, S. 17–31
- [Winther und Achtenhagen 2009] WINTHER, E.; ACHTENHAGEN, F.: Skalen und Stufen kaufmännischer Kompetenz. *Zeitschrift für Berufs- und Wirtschaftspädagogik 105 (2009) 4*, S. 521–556
- [Wirth und Klieme 2002] WIRTH, J.; KLIEME, E.: Computer literacy im Vergleich zwischen Nationen, Schulformen und Geschlechtern. *Unterrichtswissenschaft 30 (2002) 2*, S. 136–157
- [Zlatkin-Troitschanskaia und Seidel 2011] ZLATKIN-TROITSCHANSKAIA, O.; SEIDEL, J.: Kompetenz und ihre Erfassung – das neue „Theorie-Empirie-Problem“ der empirischen Bildungsforschung? In: Zlatkin-Troitschanskaia, O. (Hrsg.): *Stationen Empirischer Bildungsforschung. Traditionslinien und Perspektiven*. Verlag für Sozialwissenschaften, Wiesbaden, 2011

Abbildungsverzeichnis

2.1	Schema eines dreidimensionalen Strukturmodells	16
3.1	Fachdidaktisches Strukturmodell	54
4.1	Methodik	61
5.1	Informationsdarstellung	72
6.1	Verteilung der Inhaltsbereiche	85
6.2	Altersgruppe und Tätigkeitsbereich der Experten	86
6.3	Häufigkeitsverteilungen der Merkmale	87
6.4	Verteilung von Erfahrungsweltnähe und Geschlecht	93
6.5	Verteilung von Anforderungsbereich und Lernzielstufe	95
8.1	Dendrogramm zur Clusteranalyse der Wettbewerbsaufgaben	113
8.2	Anordnung der Aufgaben nach ihrer Schwierigkeit	117
9.1	Aufgabenkennzahlen Schwierigkeit und Trennschärfe	134
9.2	Streudiagramme der Schwierigkeit und Trennschärfe	135
9.3	Fähigkeitsmodell B als dreidimensionales Strukturmodell	149
A.1	Biber am Fluss	174
A.2	Biber und Bisons	175
A.3	Bibers Geheimcode	175
A.4	Biberzahlen	176
A.5	Binärbaum	177
A.6	Computervirus	178
A.7	Dateisuche	178
A.8	Dino-Ordnung	179
A.9	Endlosschleife	180
A.10	Falschgeld	181
A.11	Fenster schließen	182
A.12	Labyrinth	183
A.13	Link	184
A.14	Links um!	184

A.15 Morse-Code	185
A.16 Netzwirkabel	186
A.17 Platzwechsel	186
A.18 POP und PUSH	187
A.19 Primärschlüssel	188
A.20 Private E-Mail	189
A.21 Schnitzeljagd	190
A.22 Sicheres Passwort	191
A.23 Umparken	192
A.24 Ungeschützter PC	193
A.25 Verschlüsselung	193
A.26 Verwandlung	194
A.27 Wertetausch	195
A.28 Wetter	195
A.29 Zahlenreihe	196

Tabellenverzeichnis

4.1	Aufbau der Datentabelle der Aufgabenmerkmale	63
4.2	Aufbau der Datentabellen der Antwortmuster	63
4.3	Aufbau der Datentabellen der Merkmalsquoten	64
6.1	Teilnahmedaten des Informatik Bibers	83
6.2	Wettbewerbsaufgaben des Informatik Bibers	84
6.3	Güte der Merkmalsklassifizierung	89
6.4	Aufgabenmerkmale (Teil 1)	90
6.5	Aufgabenmerkmale (Teil 2)	91
7.1	Dichotome Aufgabenmerkmale	99
7.2	Antwortmuster (5-7)	101
7.3	Antwortmuster (8-10)	101
7.4	Antwortmuster (11+)	101
7.5	Merkmalsquoten (5-7)	103
7.6	Merkmalsquoten (8-10)	103
7.7	Merkmalsquoten (11+)	103
8.1	Korrelationstabelle der Merkmale	110
8.2	Clusterzuordnung der Aufgaben	114
8.3	Aufgabenkennzahlen Schwierigkeit und Trennschärfe	116
8.4	Faktorladung der Aufgaben (5-7)	127
8.5	Faktorzuordnung der Aufgaben (5-7)	127
8.6	Faktorladung der Aufgaben (8-10)	128
8.7	Faktorzuordnung der Aufgaben (8-10)	128
8.8	Faktorladung der Aufgaben (11+)	129
8.9	Faktorzuordnung der Aufgaben (11+)	129
8.10	Faktorladung der Merkmale (5-7)	130
8.11	Faktorzuordnung der Merkmale (5-7)	130
8.12	Faktorladung der Merkmale (8-10)	131
8.13	Faktorzuordnung der Merkmale (8-10)	131
8.14	Faktorladung der Merkmale (11+)	132
8.15	Faktorzuordnung der Merkmale (11+)	132

9.1	Merkmalsprofile der Cluster 1 und 2	140
9.2	Merkmalsprofile der Cluster 3 und 4	141
9.3	Merkmalsprofile der Faktoren (5-7)	152
9.4	Merkmalsprofile der Faktoren (8-10)	153
9.5	Merkmalsprofile der Faktoren (11+)	154
10.1	MultiRa Designmatrizen (5-7)	159
10.2	MultiRa-Designmatrizen (8-10)	160
10.3	MultiRa-Designmatrizen (11+)	161
10.4	Güte der Datenanpassung	162
C.1	Bewertungsverteilung (Teil 1)	210
C.2	Bewertungsverteilung (Teil 2)	211
C.3	Bewertungsverteilung (Teil 3)	212
C.4	Bewertungsverteilung (Teil 4)	213
C.5	Korrelationstabelle der Merkmale (5-7)	214
C.6	Korrelationstabelle der Merkmale (8-10)	215
C.7	Korrelationstabelle der Merkmale (11+)	216