



Entwicklung und Überprüfung eines Strukturmodells der Diagrammkompetenz für den Biologieunterricht

Dissertation
zur Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultät
der Christian-Albrechts-Universität
zu Kiel

vorgelegt von
Simone Lachmayer

Kiel
2008

Referent/Referentin:

Korreferent/Korreferentin:

Tag der mündlichen Prüfung:

Zum Druck genehmigt:

Kiel, den _____

Der Dekan

Teile dieser Arbeit wurden bereits veröffentlicht unter:

- Lachmayer, S., Nerdel, C. & Prechtel, H. (2007). Modellierung kognitiver Fähigkeiten beim Umgang mit Diagrammen im naturwissenschaftlichen Unterricht. *Zeitschrift für Didaktik der Naturwissenschaften*, 13, S. 145-160.
- Lachmayer, S., Nerdel, C. & Prechtel, H. (2007). Development of an instrument to assess diagram literacy in biology education. In European Science Education Research Association (ESERA) (Ed.), *ESERA 2007 International Conference. Abstracts* (p. 224). Malmö: DSDRA, Malmö University.
- Lachmayer, S., Nerdel, C. & Prechtel, H. (2007). Development of an instrument to assess diagram literacy in biology education. Synopsis of Oral Presentation at the ESERA 2007 International Conference. Verfügbar unter: [http://195.178.227.107/ESERA/FMPro?-db=presentations.fp5&-format=record%5fdetail%5fpres.htm&-lay=detail%5fsymposium&-sortfield=datum&-sortfield=sal&-sortfield=starttid&Last%5fname=lachmayer&-recid=35187&-find=\[14.05.2008\]](http://195.178.227.107/ESERA/FMPro?-db=presentations.fp5&-format=record%5fdetail%5fpres.htm&-lay=detail%5fsymposium&-sortfield=datum&-sortfield=sal&-sortfield=starttid&Last%5fname=lachmayer&-recid=35187&-find=[14.05.2008]).
- Lachmayer, S., Nerdel, C. & Prechtel, H. (2007). Diagrammkompetenz im Biologieunterricht. In H. Bayrhuber, F. X. Bogner, D. Graf, H. Gropengießer, M. Hammann, U. Harms, et al. (Hrsg.), *Ausbildung und Professionalisierung von Lehrkräften. Abstracts. Internationale Tagung der Fachgruppe Biologiedidaktik im VBIO – Verband Biologie, Biowissenschaften & Biomedizin, 16.09. bis 20.09.2007 in Essen* (S. 179-182). Kassel: Universität Kassel.
- Lachmayer, S., Nerdel, C. & Prechtel, H. (2007). Diagrammkompetenz im Biologieunterricht. In H. Bayrhuber, D. Höttecke & H. Vogt (Hrsg.), *Kompetenzen, Kompetenzmodelle, Kompetenzentwicklung – Empirische Forschung in den Fachdidaktiken. Abstracts. 3. Kongress der Gesellschaft für Fachdidaktik (GFD), 16.09. bis 20.09.2007 in Essen* (S. 40). Kassel: Universität Kassel.
- Lachmayer, S., Prechtel, H. & Nerdel, C. (2006). Diagrammkompetenz im Biologieunterricht. In D. Krüger, C. Sander & S. Waßenhoven (Hrsg.), *Achte Frühjahrsschule der Sektion Biologiedidaktik im Verband Deutscher Biologen vom 24. bis 26. März 2006 an der Freien Universität Berlin* (S. 38-39). Freie Universität Berlin.
- Lachmayer, S., Nerdel, C. & Prechtel, H. (2007). Diagrammkompetenz im Biologieunterricht. In K. Bätz, W. Homann, M. Wilde & N. Grotjohann (Hrsg.), *Neunte Frühjahrsschule der Sektion Biologiedidaktik im Verband Deutscher Biologen vom 23.-25.03.2007* (S. 40-41). Universität Bielefeld.

Inhaltsverzeichnis

ZUSAMMENFASSUNG	1
ABSTRACT.....	2
1 EINLEITUNG	3
2 EINSATZ VON DIAGRAMMEN IM BIOLOGIEUNTERRICHT AUS FACHDIDAKTISCHER PERSPEKTIVE	5
2.1 Diagramme als Medien der Fachwissenschaft Biologie.....	5
2.2 Diagramme als Lehr- und Lernmittel im Biologieunterricht.....	6
2.3 Problemstellung	7
3 THEORETISCHE GRUNDLAGEN	9
3.1 Begriffsklärungen	9
3.1.1 Diagramme, realistische Bilder und Texte	9
3.1.2 Kategorisierung der Diagrammtypen	11
3.2 Struktur von Linien- und Säulendiagrammen	13
3.2.1 Grundlegende Begriffe	13
3.2.2 Liniendiagramm versus Säulendiagramm	15
3.2.3 Diagramme im Mathematik- und Biologieunterricht	17
3.3 Exkurs: Curriculare Relevanz und Bildungsstandards.....	19
3.4 Kompetenz und Kompetenzmodellierung.....	22
3.4.1 Kompetenzbegriff.....	22
3.4.2 Kompetenzmodelle und ihre Überprüfung.....	24
3.5 Umgang mit Diagrammen	27
3.5.1 Kognitionspsychologische Grundlagen.....	27
3.5.2 Differenzierte theoretische Überlegungen und empirische Befunde.....	31
3.5.3 Diagrammtypen	40
3.6 Vorschlag eines Kompetenzstrukturmodells	41
3.6.1 Informationsentnahme	42
3.6.2 Konstruktion	44
3.6.3 Integration.....	47
3.7 Lernerhintergrund und Lernermerkmale	48
3.7.1 Lernerhintergrund.....	48
3.7.2 Lernermerkmale	49

4	FORSCHUNGSFRAGEN.....	52
4.1	Überprüfung der Modellstruktur	52
4.1.1	Komponenten	53
4.1.2	Diagrammtypen	53
4.2	Differenzierte Beschreibung der Modellstruktur.....	54
4.2.1	Relationen der Komponeten	54
4.2.2	Schwierigkeiten der Komponenten	55
4.2.3	Binnenstrukturen der einzelnen Komponenten	55
4.2.4	Vergleich der Schwierigkeiten bezogen auf die Diagrammtypen	55
4.3	Lernerhintergrund und Lernermerkmale	56
4.3.1	Einfluss des Lernerhintergrunds.....	56
4.3.2	Einfluss von Lernermerkmalen	56
4.4	Qualitative Fehleranalyse	57
5	METHODEN.....	58
5.1	Testtheoretischer Hintergrund.....	58
5.1.1	Grundlagen der probabilistischen Testtheorie.....	59
5.1.2	Probabilistische Testmodelle.....	62
5.1.3	Zusammenfassung	68
5.2	Erhebungsverfahren.....	69
5.2.1	Überblick über die Studien.....	69
5.2.2	Stichproben.....	70
5.2.3	Organisation der einzelnen Testung	71
5.2.4	Material.....	71
5.3	Auswertungsverfahren.....	81
5.3.1	Kodierung	82
5.3.2	Itemselektion	83
5.3.3	Bearbeitung der Forschungsfragen.....	90
6	ERGEBNISSE.....	106
6.1	Modellstruktur.....	106
6.1.1	Komponenten	106
6.1.2	Diagrammtypen	112
6.1.3	Zusammenfassung	114
6.2	Differenzierte Beschreibung der Modellstruktur.....	115
6.2.1	Relationen.....	115
6.2.2	Schwierigkeiten der Komponenten	125
6.2.3	Analyse der Binnenstrukturen der Komponenten Informationsentnahme und Konstruktion	127
6.2.4	Schwierigkeiten der Diagrammtypen	133

6.3	Lernerhintergrund und Lernermerkmale	134
6.3.1	Einfluss des Lernerhintergrunds	134
6.3.2	Einfluss der Lernermerkmale	137
6.4	Qualitative Fehleranalyse	141
6.4.1	Auswahl des Diagrammtyps	141
6.4.2	Achsenbelegung	142
6.4.3	Achsenbeschriftung	142
6.4.4	Skalierung	143
6.4.5	Werteintrag	145
7	DISKUSSION	146
7.1	Komponenten der Diagrammkompetenz und deren Relationen	146
7.1.1	Gegenüberstellung von Informationsentnahme und Konstruktion	147
7.1.2	Relation der Komponente Integration zu Informationsentnahme und Konstruktion	148
7.2	Binnenstrukturen	150
7.2.1	Ablesen	150
7.2.2	Konstruktion	152
7.2.3	Integration	153
7.3	Betrachtung der verschiedenen Diagrammtypen	154
7.4	Lernerhintergrund und Lernermerkmale	156
7.4.1	Lernerhintergrund	156
7.4.2	Lernermerkmale	157
7.5	Testinstrument	161
7.6	Implikationen für die Unterrichtspraxis	162
7.6.1	Komponenten der Diagrammkompetenz und deren Relationen	162
7.6.2	Binnenstrukturen	164
7.6.3	Diagrammtypen	165
8	AUSBLICK	166
9	LITERATURVERZEICHNIS	168
10	LEHRPLANVERZEICHNIS	180
11	ABKÜRZUNGSVERZEICHNIS	182
12	ANHANG	183
12.1	Tabellenverzeichnis	183
12.2	Abbildungsverzeichnis	185

12.3	Testinstrumente	188
12.3.1	Übersicht über die Items in den Testheften der Pilotierung	188
12.3.2	Items zu den Lernermerkmalen	190
12.4	Itemkennwerte	195
12.4.1	Pilotierung: Items zur Diagrammkompetenz.....	195
12.4.2	Pilotierung: Items zu den Lernermerkmalen	199
12.4.3	Haupttestung: Items zur Diagrammkompetenz	201
12.5	Ergänzungen zur Ergebnisdarstellung.....	203

Zusammenfassung

Dem Einsatz von Diagrammen kommt im Biologieunterricht eine vielfältige Bedeutung zu. Sie fungieren einerseits als Lehr- und Lernmittel zur Unterstützung textgebundener Informationsvermittlung, andererseits als direkter Lerngegenstand bei der Datenauswertung und Ergebnisdarstellung im Rahmen von Prozessen der Erkenntnisgewinnung im Unterricht. Schwierigkeiten von Schülerinnen und Schülern beim Umgang mit diesen hoch konventionalisierten Darstellungsformen weisen jedoch auf die Notwendigkeit der Schulung von Kenntnissen und Fähigkeiten der Lernenden zur Interpretation und Konstruktion von Diagrammen hin. Entsprechend wird der Erwerb dieser Fähigkeiten, die in der vorliegenden Arbeit unter dem Begriff der Diagrammkompetenz zusammengefasst werden, in Lehrplänen und Bildungsstandards als bedeutsames Ziel des Biologieunterrichts herausgestellt.

Als Grundlage sowohl der Förderung als auch der Diagnose der Diagrammkompetenz wird ein fundiertes Kompetenzmodell benötigt, das die relevanten Kompetenzstrukturen sowie deren Zusammenhänge darlegt. Hauptzielsetzung der vorliegenden Forschungsarbeit war die theoriegeleitete Entwicklung eines solchen Strukturmodells für die Diagrammkompetenz sowie dessen empirische Validierung, die sich zum einen auf zwei verschiedene Klassenstufen, zum anderen auf zwei verschiedene Diagrammtypen, Linien- und Säulendiagramm, erstreckt.

Basierend auf kognitionspsychologischen Befunden zum Verstehen von Diagrammen wurde ein theoretisches Kompetenzstrukturmodell entwickelt, welches als Hauptkomponenten die Fähigkeiten zur Informationsentnahme, zur Konstruktion und zur Integration voneinander unterscheidet. Diese Modellstruktur konnte umfassend empirisch bestätigt werden. Die drei Komponenten, sowie weitere Unterkomponenten, ließen sich für beide Klassenstufen als zu differenzierende Fähigkeiten nachweisen. Insbesondere zeigte sich, dass das formulierte Kompetenzmodell für beide Diagrammtypen Gültigkeit hat.

Zur differenzierteren Beschreibung der Struktur der Diagrammkompetenz wurden zudem die Zusammenhänge der einzelnen Komponenten untereinander weiter geklärt. Es wurde ein geringer Zusammenhang zwischen Informationsentnahme und Konstruktion, starke Bezüge hingegen zwischen Informationsentnahme und Integration festgestellt.

Als didaktische Implikation ergibt sich aus den Befunden der vorliegenden Arbeit, dass eine effektive Förderung der Diagrammkompetenz von Schülerinnen und Schülern im Biologieunterricht über die Schulung der Interpretationsfähigkeit gegebener Diagramme hinausgehen muss. Ein kompetenter Umgang mit Diagrammen erfordert eigene Fähigkeiten zur Konstruktion sowie Fähigkeiten, Informationen aus unterschiedlichen Quellen wie Text und Diagramm miteinander in Beziehung zu setzen. Diese hängen zwar mit den Fähigkeiten, Informationen aus Diagrammen zu entnehmen, zusammen, können aber nicht mit diesen gleichgesetzt werden.

Abstract

In biology education the use of graphs is of manifold relevance. On the one hand, graphs are applied as teaching material, because they are assumed to support learning from text. On the other hand, graphs are a subject of biology tuition by themselves, since they are used for data evaluation and presentation of results within classroom processes of scientific discovery. Students' difficulties regarding the handling of graphs indicate that knowledge of and abilities in the handling of graphs need to be taught in science classes. According to this, the requirement of these abilities, which are referred to as graph competence in this study, is considered as an important goal of biology education by curricula and standards of science education.

A graph competence model, that describes the relevant competence structures and their relations, is required as a basis for both diagnosing and fostering graph competence in science classrooms. The primary aim of the present study was the development of such a structural competence model and its empirical validation in two different grades and for two different graph types, line and bar graphs.

A competence model was theoretically derived referring to research on graph comprehension in cognitive psychology. It assumes that information extraction, construction and integration are the main components of graph competence. The model structure could be confirmed in empirical data, i.e. the named three components, as well as further subcomponents could be distinguished empirically as different abilities. In particular, it was found that the postulated competence model is valid for both types of graphs.

Furthermore, the relations between the different components were examined in order to give a more detailed description of the competence structure. It was found that information extraction and construction bear little relation to each other whereas there is a strong relationship between information extraction and integration.

A didactical implication that arises from the results of the present study is the necessity to promote not only students' abilities in interpreting graphs but also their abilities in constructing and relating them to other sources of information. These abilities are related to but cannot be equated with the abilities in reading graphs.

1 Einleitung

Dem Einsatz von Diagrammen kommt im Biologieunterricht eine vielfältige Bedeutung zu. Zum einen stellen Diagramme Unterrichtsmedien dar, mit denen abstrakte Zusammenhänge visualisiert und veranschaulicht werden können (Kattmann, 2006). Dabei wird erwartet, dass Diagramme als Ergänzung zu textlichen Erklärungen den Wissens- und Verständniserwerb der Lernenden erleichtern (Ainsworth, 1999). Zum anderen sind Diagramme zur Vermittlung naturwissenschaftlicher Denk- und Arbeitsweisen im Unterricht geeignet, da an ihnen sowohl die Datenauswertung als auch die Ergebnispräsentation, wie sie im Rahmen eines naturwissenschaftlichen Erkenntnisprozesses stattfindet, erlernt und geübt werden kann.

Derweil benötigen Schülerinnen und Schüler¹ Kenntnisse und Fähigkeiten zum Umgang mit Diagrammen, um von ihrem Einsatz im Biologieunterricht zu profitieren, da Diagramme hoch konventionalisierte Darstellungsformen darstellen, deren Verständnis intuitiv kaum möglich ist (Schnotz, 1994). Entsprechend der Bedeutung von Diagrammen im Biologieunterricht ist die Vermittlung dieser Fähigkeiten in den deutschen Lehrplänen sowie den Bildungsstandards für die Biologie (KMK, 2005a) verankert. So sollen Schüler am Ende der zehnten Klassenstufe Kompetenzen erworben haben, die sie befähigen, Diagramme zum Wissenserwerb sowie im Rahmen der Erkenntnisgewinnung flexibel einzusetzen, zu analysieren und zu konstruieren. Aus fachdidaktischer Sicht stellt sich dabei die Frage, über welches Wissen, welche Fähigkeiten und Fertigkeiten die Lernenden genau verfügen müssen, um diese Anforderungen bewältigen zu können. Darüber hinaus ist die Struktur der Gesamtheit dieser Fähigkeiten, die in der vorliegenden Arbeit im Begriff der Diagrammkompetenz zusammengefasst werden, noch weitestgehend unklar. Zur Erfassung, Diagnose und Förderung von Diagrammkompetenz ist jedoch eine detaillierte Beschreibung dieser Struktur nötig.

Zielsetzung der vorliegenden Arbeit ist deshalb die differenzierte Analyse der Struktur der Diagrammkompetenz von Schülern im Biologieunterricht. Diese Analyse erfolgt durch die theoriegeleitete Entwicklung und empirische Überprüfung eines Kompetenzmodells, welches einzelne Kompetenzstrukturen abzubilden vermag sowie auf Zusammenhänge dieser Strukturen hinweist.

Dazu wird zunächst in Kapitel 2 der Einsatz von Diagrammen im Unterricht aus fachdidaktischer Perspektive detaillierter beleuchtet. Im anschließenden theoretischen Teil der Arbeit (Kapitel 3) wird die Entwicklung eines Strukturmodells der Diagrammkompetenz nachgezeichnet, beginnend mit einer Klärung der verwendeten Begrifflichkeiten in Kapitel 3.1. Dabei werden zum einen Diagramme Texten und

¹Sofern in den folgenden Ausführungen nicht explizit zwischen Schülerinnen und Schülern unterschieden wird, schließt der Begriff Schüler die weibliche Form „Schülerinnen“ mit ein.

realistischen Bildern als unterschiedlichen Repräsentationsformen gegenüber gestellt. Zum anderen wird anhand einer Kategorisierung der gängigen Diagrammtypen herausgearbeitet, warum sich die vorliegende Arbeit auf Linien- und Säulendiagramme eingrenzt. In Kapitel 3.2 wird die Struktur der genannten Diagrammtypen sowie ihre unterschiedliche Anwendung im Mathematik- und Biologieunterricht beschrieben, bevor in Kapitel 3.3 die Relevanz der Diagrammkompetenz im naturwissenschaftlichen Unterricht anhand der gängigen Curricula sowie der Bildungsstandards erläutert wird. In Kapitel 3.4 werden der in der vorliegenden Arbeit verwendete Kompetenzbegriff erörtert und einige Grundlagen der Kompetenzmodellierung vorgestellt. Daraufhin werden in Kapitel 3.5 die kognitionspsychologischen Grundlagen sowie die empirischen Befunde ausführlich dargestellt, die zur Modellierung der Diagrammkompetenz herangezogen wurden. Das entwickelte Kompetenzstrukturmodell wird in Kapitel 3.6 beschrieben. In Kapitel 3.7 sind mögliche Einflussfaktoren auf die Diagrammkompetenz, die in der vorliegenden Untersuchung berücksichtigt wurden, aufgeführt.

Im empirischen Teil der Arbeit wird die Überprüfung des Kompetenzstrukturmodells beschrieben. Er beginnt mit der Vorstellung der Forschungsfragen und Hypothesen in Kapitel 4. In Kapitel 5 werden nach einer ausführlichen Darstellung des testtheoretischen Hintergrunds die in der vorliegenden Arbeit verwendeten Erhebungs- und Auswertungsmethoden vorgestellt. Die Ergebnisse sind in Kapitel 6 ausgeführt. Ihre Diskussion sowie die Ableitung zentraler Implikationen für die Unterrichtspraxis erfolgt in Kapitel 7. In Kapitel 8 wird in einem Ausblick auf mögliche sich der vorliegenden Arbeit anschließende Studien zur Diagrammkompetenz eingegangen.

2 Einsatz von Diagrammen im Biologieunterricht aus fachdidaktischer Perspektive

2.1 Diagramme als Medien der Fachwissenschaft Biologie

Mathematische Methoden werden in der Biologie bei der Herleitung von Gesetzmäßigkeiten, bei der Auswertung von Versuchsdaten und beim Vergleich von Forschungsergebnissen eingesetzt (Meyer, 1988). Mithilfe von Winkelfunktionen werden beispielsweise Räuber-Beute-Interaktionen modelliert und anhand von Differential- und Integralgleichungen Fragestellungen aus der Enzymkinetik bearbeitet (Bohl, 2006)². Mit den mathematischen Methoden kommt auch mathematischen Darstellungen wie Diagrammen insbesondere bei empirischen Untersuchungen eine hohe Bedeutung zu. Diese erstreckt sich sowohl auf die Datenauswertung als auch auf die Darstellung und Präsentation von Ergebnissen. Bei der Datenauswertung dient das Auftragen von Messdaten in ein Diagramm zum einen der Veranschaulichung und Visualisierung der Daten und ist zum anderen selbst Bestandteil der Analyse (Dytham, 2005, S. 9; Meyer, 1988). So ist beispielsweise vielfach erst nach dem Auftragen gemessener Daten in ein Diagramm ein bestimmter Trend ablesbar (Krohn, 1991). Die Ergebnisdarstellung erfolgt insbesondere in wissenschaftlichen Publikationen häufig mithilfe von Diagrammen, dabei werden beispielsweise Streu-, Linien- und Säulendiagramme genutzt (Roth, Bowen & McGinn, 1999).

Mit den beschriebenen Funktionen stellen Diagramme ein zentrales Medium für die Fachwissenschaft Biologie dar. Darüber hinaus wird der Umgang mit Graphiken, zu denen auch Diagramme zählen, ihre Erstellung zum Zwecke der Datenanalyse sowie zur Präsentation von Untersuchungsergebnissen, als Bestandteil einer naturwissenschaftlichen Kompetenz angesehen (Deutsches PISA-Konsortium, 2000; Bowen & Roth, 2005; vgl. auch Padilla, McKenzie & Shaw, 1986). Dabei wird von vielen Autoren die Notwendigkeit, die Vermittlung dieser Kompetenz und damit der Fähigkeiten zum Gebrauch von Diagrammen explizit in naturwissenschaftliche Curricula aufzunehmen, bekräftigt (z. B. Roth et al., 1999; Dreyfus & Eisenberg, 1990; Padilla et al., 1986). In Deutschland spiegelt sich bezogen auf die Biologie diese Notwendigkeit jüngst in den nationalen Bildungsstandards für den mittleren Schulabschluss (KMK, 2005a, vgl. 3.3) wider, in denen explizit Standards formuliert wurden, die sich auf den Umgang mit Diagrammen beziehen³. Auch in zahlreichen deutschen Lehrplänen für die Biologie wird der Umgang mit Diagrammen im Unterricht thematisiert (vgl. 3.3). Es wird deutlich, dass

² Einige Lehrbücher befassen sich ausschließlich mit der Erklärung der mathematischen Methoden, die in der Biologie eingesetzt werden (Hütt, 2001; für die Ökologie z. B. Leyer, 2007).

³ Auch in den Ausführungen für die beiden anderen naturwissenschaftlichen Fächer finden sich entsprechende Standards (KMK, 2005b,c).

die Verwendung von Diagrammen in naturwissenschaftlichen Kontexten als fachwissenschaftliche Methode im Biologieunterricht selbst geschult werden muss und nicht etwa dem Mathematikunterricht überlassen werden darf. Die Vermittlung der entsprechenden Fähigkeiten stellt damit eine direkte Aufgabe des Biologieunterrichts dar.

2.2 Diagramme als Lehr- und Lernmittel im Biologieunterricht

Wie beschrieben sind Diagramme als fachspezifische Medien direkter (Lern-)Gegenstand des Biologieunterrichts. Darüber hinaus besitzen sie ein hohes Potential für die Wissens- und Verständnisvermittlung, so dass sie generell im naturwissenschaftlichen Unterricht vielfach als Lernmittel eingesetzt werden (Moore, 1993; Roth et al., 1999). Diagramme können, wie andere Abbildungen, organisierende oder erklärende Funktion besitzen (vgl. Carney & Levin, 2002). Sie erlauben im Vergleich zur linearen Darstellungsweise in Texten eine simultane Darbietung größerer Informationsmengen. Zudem erleichtern sie das Schlussfolgern (Larkin & Simon, 1987) und stellen auf diese Weise eine wertvolle Unterstützung der textgebundenen Informationsvermittlung dar (Ainsworth, 1999). Insbesondere wird dabei der eigenständigen, aktiven Konstruktion von Diagrammen durch die Lernenden eine lernförderliche Wirkung zugeschrieben (Cox, 1999; Stern, Aprea & Ebner, 2003). In diesem Fall fungieren Diagramme als Lernmittel, anhand derer sich Lernende durch die aktive Auseinandersetzung Wissen aneignen.

Sowohl bei der Präsentation vorgefertigter Diagramme als auch bei der eigenständigen Diagrammkonstruktion durch die Lernenden werden Diagramme im Biologieunterricht insbesondere in Verbindung mit erklärenden (Lehrbuch-)Texten und anderen textlichen Informationsquellen eingesetzt. Allerdings verweisen die Ergebnisse verschiedenster Studien auf spezifische Probleme von Lernenden bei der Nutzung von Texten und Abbildungen für den Wissens- und Verständniserwerb. Insbesondere die Forschung zum Multimedialernen hat aufgezeigt, dass der Einsatz von bildhaftem Material wie illustrierenden Abbildungen und Diagrammen das Lernen nicht in jedem Falle unterstützt. Vielfach zeigen Lernende Schwierigkeiten, sich mithilfe von Text-Bild-Kombinationen Wissen über komplexe Prozesse anzueignen. Beispielsweise fällt es ihnen schwer, die relevanten Informationen aus bildhaftem Material zu entnehmen (für Diagramme speziell Peeck, 1993; Guthrie, Weber & Kimmerly, 1993, Carpenter & Shah, 1998, S. 97) oder Informationen aus Text und Bild aufeinander zu beziehen und zu verknüpfen (Salomon, 1984; Weidenmann, 1988; Chandler & Sweller, 1991; Lewalter, 1997; Ainsworth, 1999; Bodemer, Plötzner, Feuerlein & Spada, 2004; Nerdel, Prechtel & Bayrhuber, 2003; Kramer, Prechtel & Bayrhuber, 2005). Die Folge der genannten Probleme ist vor allem ein geringer Verständniszuwachs.

Eine besondere Anforderung beim Lernen mit zugleich textlichem und bildhaftem Material stellt somit die (mentale) Zusammenführung der Informationen aus den unterschiedlichen Repräsentationsformen dar (Seufert, 2003; Ainsworth, Bibby, & Wood, 2002). In vielen Fällen zeigt sich, dass eine höhere Vertrautheit mit der bildhaften Darstellungsform zu einem besseren Lernergebnis (z. B. Peeck, 1993, S. 232) und insbesondere zu einer besseren Verknüpfungsleistung führt. Ist die bildliche Darstellung unvertraut, beziehen die Lernenden beispielsweise die verschiedenen Repräsentationsformen lediglich über deren Oberflächenmerkmale aufeinander, während Lernende mit höherer Vertrautheit über die dahinterliegenden Konzepte Bezüge herstellen können (Kozma & Russell, 1997). Auch bezogen auf die eigenständige Konstruktion von Diagrammen gibt es Hinweise dafür, dass eine höhere Vertrautheit mit der Darstellungsform Diagramm hilfreich und sogar notwendig dazu ist, dass die o.g. Erwartungen an lernförderliche Effekte erfüllt werden (Stern, Aprea & Ebner, 2003). Diese Effekte stellen sich keineswegs automatisch ein (vgl. Aprea & Ebner, 2003).

Zusammenfassend sind somit Kenntnisse über die Darstellungsform und eine höhere Vertrautheit bei deren Anwendung erforderlich, damit das Potential, das der Einsatz von Diagrammen als Lehr- und Lernmittel im Unterricht birgt, ausgeschöpft werden kann.

2.3 Problemstellung

In den obigen Ausführungen ist ausgehend von der Bedeutung, die Diagramme in der Fachwissenschaft Biologie besitzen, der Umgang mit Diagrammen als direktes Lernziel für den Biologieunterricht herausgestellt worden. Darüber hinaus wurde das Potential von Diagrammen als Lehr- und Lernmittel im Biologieunterricht und die spezifischen Hindernisse zu deren lernförderlichem Einsatz beleuchtet. Um vom Einsatz von Diagrammen im Unterricht profitieren können, benötigen Lernende Wissen und Fähigkeiten bezüglich der Handhabung dieser Darstellungsformen. Da die lernförderliche Wirkung beim Einsatz von Diagrammen häufig ausbleibt, stellt sich zum einen die Frage, welche Kenntnisse über die Darstellungsform und welche Fähigkeiten zu deren Interpretation für einen erfolgreichen Umgang mit Diagrammen zu schulen sind, und zum zweiten, in welchem Maße dazu auch Fähigkeiten zur eigenständigen Konstruktion vorhanden sein sollten. Damit verbunden ist die Frage nach dem Zusammenhang von eigenständigem Konstruieren und inhaltlichem Verstehen (Shah & Hoeffner, 2002).

Die genannten Fragestellungen sind von besonderem Interesse für die Fachdidaktiken der naturwissenschaftlichen Fächer, in denen Diagramme häufig im Unterricht eingesetzt werden. Insbesondere für die Lehrenden dieser Fächer ist es wichtig, die Lernenden zum fachgemäßen Umgang und insbesondere zu einer lernförderlichen Nutzung von Diagrammen zu befähigen. Eine besondere Herausforderung stellt dabei die Diagnose

dieser Fähigkeiten dar, auf deren Basis gezielte Förderungsmaßnahmen vorgenommen werden können. An dieser Stelle setzt die vorliegende Arbeit an. Es wird ein Modell der Diagrammkompetenz vorgeschlagen und empirisch überprüft. Damit wird zum einen der Zusammenhang zwischen Diagrammkonstruktion und Diagrammverstehen beleuchtet. Zum anderen wird damit die Grundlage für eine differenzierte Diagnose der Fähigkeiten zum Umgang mit Diagrammen geschaffen, durch die gezielte Förderungsmaßnahmen für den Biologieunterricht abgeleitet werden können.

3 Theoretische Grundlagen

3.1 Begriffsklärungen

Der Begriff Diagramm ist in unterschiedlichen Kontexten auf vielfältige Weise besetzt. So werden in der Biologie beispielsweise schematische, grundrissartige Darstellungen von Blüten, wie in Abbildung 3.1 dargestellt, als „Blütendiagramme“ (Sauermost, 2000) und baumartig verzweigte Übersichten über Erbschemata (für ein Beispiel siehe Abbildung 3.1) als Baumdiagramme bezeichnet. Beide Beispiele weichen von der Auffassung, Diagramme seien „bildliche Formen mathematischer Abbildungen“ (Kattmann, 2006, S. 340), ab.

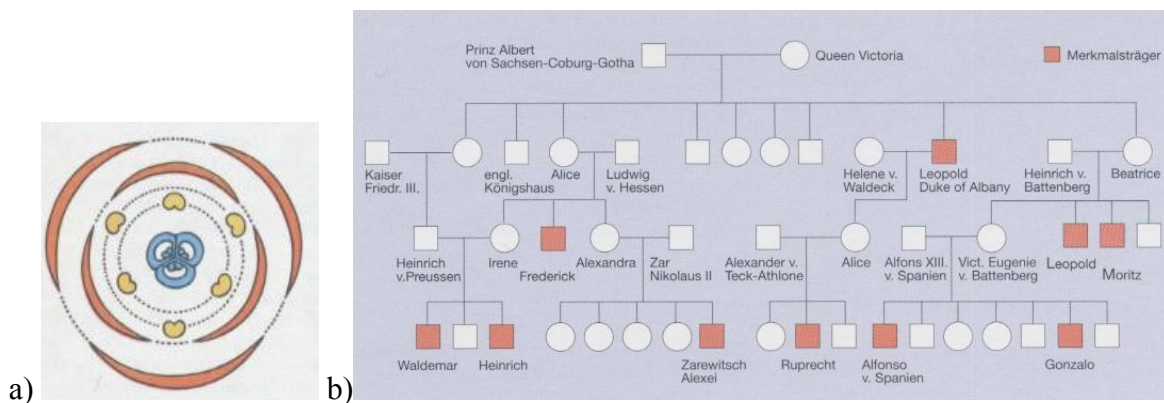


Abbildung 3.1: Beispiele für Blütendiagramm (a) und Baumdiagramm (b). (a aus Vogel & Angermann, 1992, S. 558; b aus Bickel et al., 1995, S. 125).

Aufgrund der Heterogenität der Begriffsverwendung ist es notwendig, den Begriff des Diagramms in der vorliegenden Arbeit zu präzisieren. Zu diesem Zweck werden Diagramme im Folgenden zunächst sowohl von anderen bildlichen Darstellungen als auch von Texten abgegrenzt, indem die Klassifikation von deskriptionalen und depiktionalen Repräsentationen (Schnotz, 2001) herangezogen wird (3.1.1)⁴. Daraufhin wird vor dem Hintergrund der Vielfalt an unterschiedlichen Typen von Diagrammen spezifiziert, auf welche dieser Typen sich die vorliegende Arbeit bezieht (3.1.2).

3.1.1 Diagramme, realistische Bilder und Texte

Diagramme, Bilder und Texte sind Repräsentationen – sie stellen reale Objekte oder reale Sachverhalte dar und geben sie wieder. Schnotz (2001) schlägt eine Klassifikation der Repräsentationsformen basierend auf dem Grad der Ähnlichkeit zwischen dem realen Objekt und seiner externen Darstellung, seiner Repräsentation vor. Diese Klassifikation wird im Folgenden erläutert.

⁴ Diese Klassifikation wird an späterer Stelle (s. 3.5.1) erneut aufgegriffen, da sie für die Theorie der Diagrammverarbeitung bedeutsam ist.

Texte sind verschieden von dem, was sie beschreiben: sie bestehen aus willkürlich festgelegten Zeichen, Buchstaben, Wörtern und Zahlen, die selbst keinerlei Ähnlichkeit mit dem realen Objekt haben, sondern mit ihm nur durch eine Konvention verknüpft sind. Das Wort „Katze“ hat beispielsweise keine Ähnlichkeit mit dem realen Tier. Darstellungen dieser Art bezeichnet Schnotz (2001) als *deskriptionale Repräsentationen*. Unter **Texten** werden in der vorliegenden Arbeit Sätze aus deskriptionalen Repräsentationen nicht festgelegter Anzahl verstanden.

Im Unterschied zu Texten zählen bildliche Darstellungen zu den *depiktionalen Repräsentationen*, die dadurch gekennzeichnet sind, dass sie eine Ähnlichkeit zwischen dem realen Objekt und seiner Darstellung aufweisen (Schnotz, 2001). Diese Ähnlichkeit kann offensichtlich sein wie im Falle der *realistischen Bilder*, zu denen Gemälde, Fotografien, Strich- und Schemazeichnungen zählen. Sie kann allerdings ebenfalls abstrakt sein, wie im Falle der **Diagramme**, die synonym als logische Bilder bezeichnet werden (Schnotz, 2001, 1994). Logische Bilder weisen keine *sichtbare* Ähnlichkeit mit dem dargestellten Sachverhalt auf, sie stimmen mit ihm allerdings in einer gewissen Hinsicht überein: die (logischen) Relationen zwischen den Merkmalen sind innerhalb des Bildes und innerhalb des abgebildeten Sachverhalts gleich, damit erhalten logische Bilder die Struktur des dargestellten Sachverhalts (Schnotz, 2001). Diese strukturelle Übereinstimmung ist, anders als bei den realistischen Bildern, durch Konvention bestimmt. So würden im realistischen Bild zur Darstellung verschieden großer Blütendurchmesser maßstabsgetreue Zeichnungen der Blüten verwendet, während diese im logischen Bild beispielsweise durch verschieden hohe Säulen angegeben werden. Diese abstrakte strukturelle Übereinstimmung beruht auf den Konventionen, dass verschiedene Säulen für verschiedene Pflanzenarten stehen und die Höhe der Säulen die Größe des Blütendurchmessers angibt. Im Hinblick auf die Konventionen, die Diagramme mit dem, was sie darstellen, verknüpfen, sind Diagramme den deskriptionalen Repräsentationen vergleichbar⁵. Dennoch werden sie aufgrund der abstrakten Ähnlichkeit und den dadurch gewährleisteten Strukturerehalt den depiktionalen Repräsentationen zugeordnet (Schnotz, 1994, 2001). Damit umfassen depiktionale Repräsentationen realistische und logische Bilder⁶.

Vor dem Hintergrund dieser Ausführungen sind die eingangs erwähnten Blütendiagramme nicht den Diagrammen im Sinne der logischen Bilder zuzurechnen, sondern den Schemazeichnungen und damit den realistischen Bildern. Blütendiagramme stellen eine hoch kontextualisierte Darstellungsform dar, mit der eine spezifische Interpretation verbunden ist, welche nicht auf „andere“ Typen von Diagrammen übertragen werden kann.

⁵ Darüber hinaus enthalten Diagramme in Form der Achsenbeschriftungen oft selbst Text und damit deskriptionale Repräsentationen.

⁶ Eine weitere Kategorie bilden nach Schnotz (2001, in Anlehnung an Issing, 1983, S. 13) die *analogen Bilder*.

Das Beispiel des Baumdiagramms aus Abbildung 3.1 wird im folgenden Abschnitt behandelt.

3.1.2 Kategorisierung der Diagrammtypen

Wie oben ausgeführt, sind Diagramme bildhafte Darstellungen, in denen die Relationen innerhalb eines realen Sachverhalts mit den Relationen innerhalb der Darstellung in abstrakter Hinsicht übereinstimmen. Dabei unterscheiden sich die verschiedenen Diagrammtypen dahingehend, wie die Relationen in der bildhaften Darstellung wiedergegeben sind (vgl. Kosslyn, 1989). Es resultiert eine Vielzahl an verschiedenen Diagrammtypen, die auf unterschiedliche Art interpretiert und konstruiert werden müssen (vgl. Novick, 2001, Moore, 1993). Im Folgenden wird in Anlehnung an Schnotz (1994) eine Kategorisierung der Diagrammtypen danach vorgeschlagen, ob die dargestellten Relationen oder Zusammenhänge qualitativer oder quantitativer Art sind.

Zur Visualisierung *qualitativer Zusammenhänge* zwischen Variablen werden Fluss-, Pfeil- und Baumdiagramme gebraucht, in denen einzelne Elemente (z. B. Worte, Symbole) durch Linien oder Pfeile miteinander verbunden sind. Diese Pfeile repräsentieren Relationen zwischen den Elementen und müssen dabei überwiegend inhaltlich interpretiert werden (Schnotz, 2001). Es ist darauf hinzuweisen, dass im englischen Sprachgebrauch für die beschriebenen Diagrammtypen die Oberbegriffe *charts* (z. B. Kosslyn, 1989) und *diagrams* (z. B. Winn, 1987) existieren, die allerdings nicht synonym verwendet werden⁷.

In Achsendiagrammen wie Linien-, Balken-, Säulen-, und Streudiagrammen ist mindestens eine der Variablen kontinuierlich (vgl. 3.2.1), d. h. diese Diagrammtypen dienen der Visualisierung von *quantitativen Zusammenhängen* zwischen Variablen (Schnotz, 1994). Auf den Skalen sind dabei die jeweiligen Ausprägungen, die eine Variable annehmen kann, angegeben. Dadurch kann ein Element, d. h. ein Datenpunkt oder eine Säule, eindeutig durch seine relative Lage zu den Achsen beschrieben werden. Somit werden beschriftete Pfeile zwischen den einzelnen Datenpunkten überflüssig, denn die Beziehungen zwischen den Elementen ergeben sich ebenfalls durch die Lage der einzelnen Elemente zu den Achsen (Schnotz, 1994). Der Terminus *Achsendiagramme* beschreibt die Diagramme zur Visualisierung quantitativer Zusammenhänge allerdings nicht erschöpfend, da auch Kreisdiagramme (nach der Festlegung, dass mindestens eine dargestellte Variable kontinuierlich ist) den letzteren zuzurechnen sind (Schnotz, 1994). Kosslyn (1989) fasst die genannten Diagrammtypen (Achsen- und Kreisdiagramme) unter dem Begriff *graphs*⁸

⁷ Während nach Kosslyn (1989) Pfeil- und Flussdiagramme als *charts* bezeichnet werden und der Begriff *diagram* sich eher auf Schemazeichnungen bezieht, versteht Winn (1987) unter *charts* Tabellen und Tafeln und schließt bei seiner Beschreibung von *diagram* sowohl Pfeil- und Flussdiagramme als auch Schemazeichnungen mit ein.

⁸ Vergleiche dazu die englischen Bezeichnungen für Linien-, Säulen- und Kreisdiagramme: *line graphs*, *bar graphs* und *pie graphs*.

zusammen, da ihnen gemein ist, dass sie die räumliche Ausdehnung von Flächen oder Linien zur Darstellung von Quantitäten nutzen.

Einen Überblick über die Kategorisierung der genannten Diagrammtypen zeigt die folgende Abbildung.

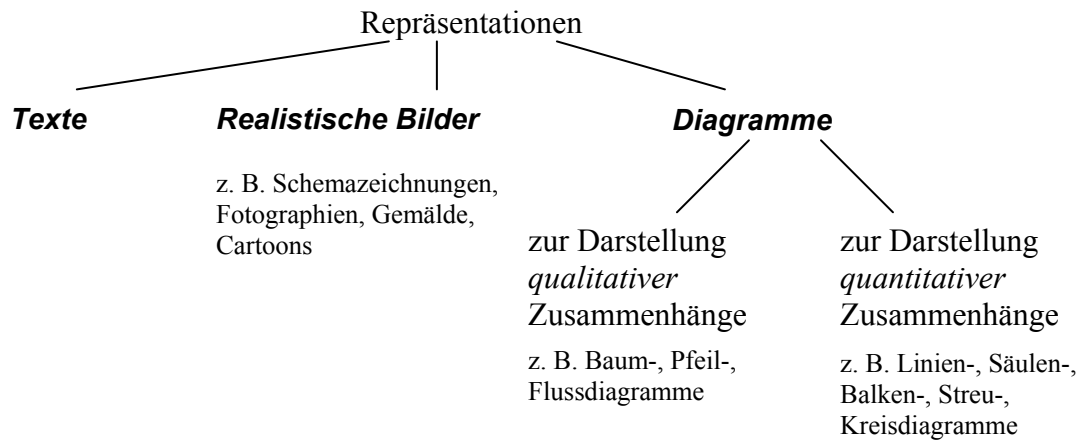


Abbildung 3.2: Kategorisierung der Diagrammtypen mit Bezug zu realistischen Bildern und Texten.

Die Diagramme zur Visualisierung qualitativer Zusammenhänge unterscheiden sich noch in weiterer Hinsicht von denen zur Visualisierung quantitativer Zusammenhänge. Für die Pfeil- und Baumdiagramme gibt es keine kontextunabhängigen Festsetzungen zur Interpretation, die von einem auf ein beliebiges anderes Diagramm übertragen werden können (vgl. Heiser & Tversky, 2006). So kann ein Pfeil in einem Ablaufdiagramm auf eine zeitliche Abfolge hindeuten, er kann jedoch wie beispielsweise im Falle eines Stammbaumes für eine verwandtschaftliche Beziehung, wie „Kind von“ oder „Bruder von“ stehen. Im Gegensatz dazu besitzen Achsen- und Kreisdiagramme eine durch Konvention festgelegte Form, so dass sie nach bestimmten Regeln interpretiert und konstruiert werden müssen. Sie sind damit in überaus vielfältigen Kontexten anwendbar (Kosslyn, 1989; Zelazny, 2001)⁹. Aus diesem Grund sind es vor allem die Achsendiagramme, auf die in der vorliegenden Arbeit Bezug genommen wird. Aufgrund deren besonders häufigen Vorkommens in Biologielehrbüchern (vgl. Moore, 1993; Roth et al., 1999) wird dabei vorrangig auf Linien- und Säulendiagramme eingegangen. Die Struktur dieser beiden Diagrammtypen wird im folgenden Abschnitt detaillierter vorgestellt.

⁹ Für den betriebswirtschaftlichen Bereich können nach Zelazny (2001) 90% der interessierenden Analysenergebnisse mithilfe der Diagrammtypen Säulen-, Balken-, Linien-, Streu- und Kreisdiagramm dargestellt werden.

3.2 Struktur von Linien- und Säulendiagrammen

Im Folgenden werden die Strukturen von Linien- und Säulendiagrammen erläutert und einander gegenübergestellt. Dazu werden die zum Verständnis der genannten Diagrammtypen relevanten grundlegenden Strukturbegriffe erklärt (3.2.1), bevor auf die Spezifika der beiden Diagrammtypen eingegangen wird (3.2.2). Abschließend wird herausgestellt, worin die Unterschiede bei der Verwendung der genannten Diagramme im Biologie- und Mathematikunterricht liegen (3.2.3).

3.2.1 Grundlegende Begriffe

Linien- und Säulendiagramme stellen, wie in Abschnitt 3.1.2 beschrieben, quantitative Zusammenhänge zwischen verschiedenen Merkmalen dar. Für diese Merkmale, die im Gegensatz zu einer Konstanten mindestens zwei Ausprägungen aufweisen, wird der Begriff der *Variable* verwendet (vgl. Bortz, 1999, S. 7). In der einfachsten Form wird in Diagrammen ein Zusammenhang zwischen *zwei* Variablen visualisiert. Dabei lassen sich unabhängige und abhängige Variable unterscheiden. Als *unabhängige Variable (UV)* wird diejenige bezeichnet, die eine (vermutete) Auswirkung auf die andere – die *abhängige Variable (AV)* – hat. In einer Untersuchung ist die UV diejenige, die systematisch variiert wird, und die AV diejenige, die gemessen wird.

Für die Darstellung wird eine durch zwei rechtwinklig aufeinanderstehende Achsen begrenzte Fläche genutzt. Die waagerechte Achse wird als *Abszisse* und die senkrechte Achse als *Ordinate* bezeichnet (Schnotz, 1994). In Anlehnung an die mathematischen Funktionsgleichungen $f(x) = y$, die x als UV und y als AV charakterisieren, werden für Abszisse und Ordinate auch die Bezeichnungen *x-Achse* und *y-Achse* verwendet. Entsprechend ist es auch in konkreten Anwendungen, in denen nicht x und y verwendet werden, Konvention, die UV („ x “) auf der x -Achse und die AV („ y “) auf der y -Achse einzutragen (Meyer, 1988; Dytham, 2005). Demzufolge kann die Belegung der Achsen mit Variablen als Indiz gewertet werden, welche Abhängigkeit im dargestellten Zusammenhang vorliegt, wobei die Abhängigkeit der beiden Variablen in den meisten Fällen auch inhaltlich oder aufgrund einer gegebenen Fragestellung (vgl. Bortz, 1999, S. 8) erschließbar ist.

Entsprechend der Achsenbelegung werden auf den Achsen die Ausprägungen der Variablen, die zur Vereinfachung im Folgenden als *Werte* bezeichnet werden, markiert und eingetragen. Indem ausgehend von diesen Markierungen zu den Achsen senkrechte Linien in die Diagrammfläche gezeichnet werden, wird dadurch die Diagrammfläche mit einem fiktiven Gitternetz versehen. Die Kreuzungspunkte der Linien auf der Diagrammfläche definieren alle möglichen Paarungen je eines Wertes der UV und eines der AV. Diese Wertepaare symbolisieren die *Daten* (i. e. mehrere einzelne Datenpunkte), die in ein

Diagramm eingetragen werden können. Bei der Konstruktion eines konkreten Diagramms werden damit bestimmte Kreuzungspunkte als Datenpunkte auf der Diagrammfläche markiert.

In Abhängigkeit davon, welche Werte eine Variable annehmen kann und in welcher Relation diese zueinander stehen, ergeben sich verschiedene Arten von *Skalierungen* für die Diagrammachsen, die im Folgenden kurz vorgestellt werden. Für detaillierte Informationen sei auf Bortz (1999) und Backhaus et al. (2003) verwiesen. Im Fall einer **nominalen Skalierung** lassen sich die einzelnen Werte lediglich unterscheiden. Man spricht auch von *kategorialer* Skalierung (vgl. Dytham, 2005, S. 33), weil die Werte Kategorien darstellen, die unterschiedlich sind, aber sich nicht etwa in eine Reihenfolge bringen lassen. Ein Beispiel ist die Variable *Pflanzenart* mit den Werten *Löwenzahn* und *Gänseblümchen*. Im Gegensatz dazu ist die Aufreihung im Fall der **ordinalen Skalierung** möglich. Ein typisches Beispiel ist die Variable *Schulnote*, deren Ausprägungen *ungenügend* bis *sehr gut* in einer Rangfolge hinsichtlich der Güte der erbrachten Leistung stehen. Zu beachten ist an dieser Stelle, dass die Unterschiede zwischen den einzelnen Noten keineswegs gleich sein müssen¹⁰. Es ist durchaus möglich, dass ein sehr gut benoteter Schüler nur geringfügig besser ist als ein gut benoteter Schüler, letzterer hingegen deutlich besser als ein Schüler mit einer befriedigenden Leistungsbewertung. Die ordinale Skalierung gibt keinen Aufschluss über die Größe des Abstandes zwischen den Werten einer Variablen. Eine **metrische Skalierung**¹¹ unterscheidet sich in dieser Hinsicht von der ordinalen Skalierung. Im Fall der metrischen Skalierung existiert ebenfalls eine Reihenfolge der Werte, darüber hinaus sind jedoch die Abstände zwischen den Werten festgelegt. Dieser Skala liegt eine Metrik und damit die Möglichkeit der Messung von Werten zugrunde. Theoretisch existieren zwischen je zwei Werten unendlich viele weitere Werte, weshalb diese Art der Skalierung auch als *kontinuierlich* bezeichnet wird (Dytham, 2005, S. 32). Zu einer metrischen Skalierung gehört darum die Angabe der **Einheit**, die die Abstände zwischen den auf der Skala markierten Werten erklärt. Beispiele für Variablen, die sich metrisch skalieren lassen und damit abkürzend als metrische Variablen bezeichnet werden¹², sind die *Temperatur*, deren Werte *verschieden große Gradzahlen* sind, oder die *Höhe eines Baumes*, deren Werte *verschiedene Zahlen an Zentimetern* sind.

Linien- und Säulendiagramm unterscheiden sich zum einen hinsichtlich der Struktur der eingetragenen Daten, welche ihrerseits von den zugrunde liegenden Skalierungen der Variablen bestimmt wird. Zum anderen bedienen sich diese beiden Diagrammtypen

¹⁰ auch wenn dies durch die Bezeichnungen 1 bis 6 für die Noten impliziert werden könnte.

¹¹ Darunter werden an dieser Stelle sowohl Intervallskalierung als auch Verhältnisskalierung, die einen definierten Nullpunkt hat, zusammengefasst.

¹² Entsprechendes gilt für nominale und ordinale Variablen.

verschiedener Arten der Markierung von Datenpunkten. Diese Unterschiede werden nachfolgend beschrieben.

3.2.2 Liniendiagramm versus Säulendiagramm

Säulendiagramme werden in der Regel zur Darstellung des Zusammenhangs zwischen einer nominal oder ordinal skalierten unabhängigen Variable (UV, vgl. 3.2.1) und einer metrisch skalierten abhängigen Variable (AV, vgl. 3.2.1) verwendet (Köhler et al., 2002, S. 18, 22). Die Datenpunkte bestehen damit aus einer Kategorie (der UV) und einem Zahlenwert (der AV). Die Visualisierung dieser Datenpunkte erfolgt in Form senkrechter Säulen¹³ zwischen der Abszisse und dem Gitternetzpunkt in Höhe des Zahlenwertes. Damit ist einzig die Höhe der Säule entscheidend für die Kennzeichnung des AV-Wertes, die flächige Markierung in Form der Säule dient in erster Linie der einfacheren Wahrnehmung. In Säulendiagrammen werden aus Gründen der Lesbarkeit Freiräume zwischen den Säulen der einzelnen Kategorien gelassen (Dytham, 2005, S. 48; Köhler, Schachtel & Voleske, 2002)¹⁴.

Grundsätzlich *kann* ein Säulendiagramm auch dann verwendet werden, wenn die UV metrisch skaliert ist (Köhler et al., 2002). Da bei metrischer Skalierung meist viele Ausprägungen der UV betrachtet werden, wird in diesen Fällen allerdings die Darstellung in Form eines Liniendiagramms empfohlen (Dytham, 2005; Andreß, 2001). Im Gegensatz zum Säulendiagramm ist ein Liniendiagramm *nur* dann zur Darstellung eines Zusammenhangs zwischen zwei Variablen geeignet, wenn beide Variablen metrisch skalierbar sind (Meyer, 1988; Dytham, 2005, S. 59). Die Datenpunkte bestehen folglich aus zwei Zahlenwerten, die auch Koordinaten genannt werden. Die Markierung der Datenpunkte erfolgt in Form von Kreuzen oder Punkten, die durch eine Linie verbunden sind. Sind die Daten aus einer Untersuchung hervorgegangen, beschreibt die Linie hypothetische Werte. Die Linie ist damit eine Interpolation, die deutlich von den real gemessenen Daten unterschieden werden muss (Kattmann, 2006; Meyer, 1988). Insbesondere sollte beim Vorliegen einer nominalen UV keine Linie zwischen Datenpunkte gezeichnet, sondern ein Säulendiagramm gewählt werden. Kattmann (2006, S. 342) sieht bei Auswahl des unangemessenen Diagrammtyps die Gefahr, dass

¹³ Balkendiagramme unterscheiden sich dadurch von Säulendiagrammen, dass in ersteren statt senkrechten Säulen waagerechte Balken verwendet werden (Schnotz, 1994). Zu beachten ist dabei, dass in Balkendiagrammen die UV auf der Ordinate eingetragen wird.

¹⁴ Im Vergleich dazu werden die Säulen in einem Histogramm in der Regel dicht aneinander gezeichnet (Dytham, 2005, S. 47). Auch in diesem Diagrammtyp liegt eine ordinale UV vor, die allerdings aus einer metrischen UV gebildet wird, indem die metrische Skala in einzelne Abschnitte eingeteilt wird. Durch die aneinanderliegenden Säulen wird dem direkten Aneinandergrenzen der Abschnitte auf der metrischen Skala entsprochen. Da allerdings die AV in Histogrammen die Häufigkeit beschreibt, mit der ein Wert der UV aufgetreten ist, (vgl. Dytham, 2005), ist nicht jedes Säulendiagramm, dessen ordinale UV auf eine metrische Skalierung zurückgeht, als Histogramm zu bezeichnen.

„Unterschiede oder Gemeinsamkeiten vorgetäuscht werden, die in der Sache gar nicht vorhanden sind“. Neben der Darstellung konkreter Daten kann ein Liniendiagramm zur Visualisierung eines generellen Trends verwendet werden. In diesem Fall ist nicht zwischen der Markierung der Datenpunkte und der Interpolationslinie zu unterscheiden.

Für beide genannten Diagrammtypen – Linien- und Säulendiagramme – ist das Eintragen mehrerer **Datenreihen** möglich. Bisher wurde der Fall geschildert, dass zu jedem Wert der UV jeweils ein Wert in Abhängigkeit betrachtet wurde, es ergab sich eine einzige Datenreihe. In Form weiterer Datenreihen wird der gleiche Zusammenhang zu bestimmten anderen Bedingungen, beispielsweise an einem anderen Objekt, an einem anderen Standort oder zu einem anderen Zeitpunkt¹⁵ dargestellt. Verschiedene Datenreihen werden durch andere Farben, Schraffierungen oder Symbole gekennzeichnet, welche in einer **Legende** erklärt werden.

Darüber hinaus können sowohl in Linien- als auch in Säulendiagrammen zwei verschiedene AVs dargestellt werden. Ein typisches Beispiel aus dem Biologieunterricht ist die Darstellung der Temperatur und des Sauerstoffgehalts des Wassers in einem See in Abhängigkeit von der Wassertiefe. Beide AVs benötigen verschiedene Skalen, die, wie in der Abbildung 3.3, häufig links und rechts der Diagrammfläche eingezeichnet werden. In diesem Fall ist neben der Spezifizierung der Datenreihen zu kennzeichnen, welche Daten sich auf welche Skala beziehen.

¹⁵ Wiederholte Messungen können demnach in Form mehrerer Datenreihen dargestellt werden. Gleichzeitig wäre die Berechnung des Mittelwerts aus den gemessenen Werten oder die Auftragung der Messwerte in Form eines Streudiagramms möglich (Dytham, 2005, S. 60). In Streudiagrammen werden vorrangig stochastische Zusammenhänge anhand einer begrenzten Zahl von Beobachtungen ermittelt (Schnotz, 1994), indem beispielsweise Regressionsgeraden errechnet werden. Die Gerade beschreibt dabei den stochastischen Trend.

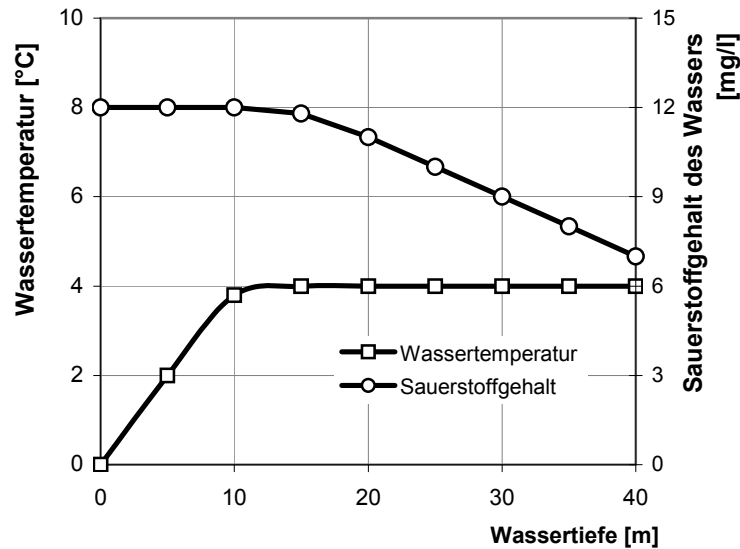


Abbildung 3.3: Liniendiagramm mit zwei abhängigen Variablen, die auf zwei verschiedenen Achsen eingetragen sind.

Über die dargestellten Grundformen von Linien- und Säulendiagrammen hinaus existieren einige Sonderformen, wie beispielsweise kombinierte Linien- und Säulendiagramme sowie gestaffelte Säulendiagramme. Da sie allerdings in der Biologie relativ selten vorkommen, werden sie aus der hier beschriebenen Untersuchung ausgeklammert.

3.2.3 Diagramme im Mathematik- und Biologieunterricht

Im Folgenden wird darauf eingegangen, worin Unterschiede bei der Verwendung der oben beschriebenen Diagrammformen im Mathematik- und im Biologieunterricht bestehen.

Ein bedeutender Unterschied ist das Abstraktionsniveau. In der Biologie beschreiben Diagramme Zusammenhänge zwischen *konkreten* Variablen. Sie werden basierend auf Messwerten zur Analyse eines Zusammenhangs oder zur verallgemeinerten Darstellung eines bereits bekannten Zusammenhangs verwendet. In der Mathematik sind Diagramme in vielen Fällen rein *abstrakt*. Sie beschreiben keine konkreten Variablen, sondern den Zusammenhang zwischen zwei Zahlenmengen. Die Art des Zusammenhangs steht im Blickpunkt, es interessieren dessen mathematischen Eigenschaften (vgl. LP SH Mathe¹⁶, S. 61, Eindeutigkeit und Umkehrbarkeit verschiedener Zuordnungen). Ein solcher Zusammenhang wird häufig in Form einer Funktion beschrieben, entsprechend werden Diagramme in der Mathematik häufig einschränkend als bildhafte Darstellungen von Funktionen charakterisiert. Eine solche bildliche Darstellung einer Funktion, der

¹⁶ Zur Vereinfachung werden Lehrpläne u. ä. hier und im Folgenden abweichend von der gebräuchlichen Zitationsweise abgekürzt. Die Abkürzungen werden in einem separaten Lehrplanverzeichnis (s. Kapitel 10) erklärt.

Funktionsgraph¹⁷, liefert beispielsweise Anhaltspunkte darüber, ob ein linearer Zusammenhang vorliegt oder nicht (LP SA Mathe, S. 55f). Anhand mehrerer Funktionsgraphen können Fragestellungen wie „Wie ändert sich der Graph der Gleichung $y = x + b$, wenn man verschiedene Werte für b einsetzt?“ bearbeitet werden. Entsprechend dem Abstraktionsniveau der Variablen unterscheiden sich Diagramme in Mathematik und Biologie damit hinsichtlich des Abstraktionsniveaus der Fragestellungen. Wenngleich auch in der Biologie die Art eines Zusammenhangs von Interesse ist, so liegt der Fokus dabei auf der inhaltlichen Interpretation (vgl. Leinhardt, Zaslavsky & Stein, 1990): Wie wirkt die eine Variable auf die andere? Welche Auswirkungen hat der dargestellte Zusammenhang auf die Variablen?

Um diese Interpretation leisten zu können, müssen die konkreten inhaltlichen Variablen mit den graphischen Symbolen, die in einem Diagramm verwendet werden, verknüpft werden (vgl. Bertin, 1974; Carpenter & Shah, 1998; Roth & Bowen, 2003). In der Literatur gibt es zahlreiche Hinweise auf spezifische Schwierigkeiten bei der inhaltlichen Interpretation (z. B. Bell & Janvier, 1981; Janvier, 1986, Padilla et al., 1986). Beispielsweise beschränkt sich die Interpretation eines Diagramms auf eine Beschreibung der graphischen Strukturen, beispielsweise eines Linienverlaufs, wenn es Schülern nicht gelingt, die graphischen Formen des Diagramms mit dem Inhalt in Beziehung zu setzen (Preece & Janvier, 1992). Insbesondere ist ein (positiver wie negativer) Einfluss des inhaltlichen Vorwissens der Schüler auf die inhaltliche Interpretation eines Diagramms zu berücksichtigen (Preece & Janvier, 1992; Scaife & Rogers, 1996; Friel, Curcio & Bright, 2001).

Neben dem Diagrammverstehen kann es auch hinsichtlich der selbstständigen Konstruktion im konkreten Anwendungsfall zu Schwierigkeiten kommen. Auf der einen Seite wird berichtet, dass Schüler Diagramme häufig als Bilder ansehen (Janvier, 1981; Bell & Janvier, 1981). Dabei interpretieren sie beispielsweise ein Liniendiagramm, das die Geschwindigkeiten eines Rennwagens auf einer Rennstrecke darstellt, als Bild des Parcours. Auf der anderen Seite haben Schüler, die in abstrakten Szenarien Wissen über Diagramme erworben haben, Probleme, ein Diagramm zur Veranschaulichung konkreter Variablen zu zeichnen (Leinhardt et al., 1990). Im Speziellen ergeben sich – im Vergleich zur abstrakten Verwendung – beim Vorliegen konkreter Variablen Anforderungen, die beispielsweise die Belegung der Achsen mit unabhängiger und abhängiger Variable betreffen. So ist die Achsenbelegung beim Zeichnen eines Funktionsgraphen nach der Konvention vorgegeben: Die x -Werte werden auf der Abzisse und die y -Werte auf der Ordinate eingetragen. Die Konvention, die unabhängige Variable auf der x -Achse einzutragen, wird dabei nicht expliziert. Soll der Zusammenhang zwischen zwei konkreten

¹⁷ Der Graph ist dabei als eine Form des Liniendiagramms zu empfehlen (vgl. 3.2.2), da die genannten Zahlenmengen, zwischen denen ein Zusammenhang beschrieben wird, metrische Variablen sind.

Variablen in einem Diagramm dargestellt werden, muss die unabhängige Variable identifiziert werden, um sie auf der (nach Konvention) richtigen Achse einzutragen. Ebenso muss im konkreten Fall beachtet werden, welches Skalenniveau vorliegt und in Folge dessen welcher Diagrammtyp zur Darstellung des interessierenden Zusammenhangs überhaupt geeignet ist (vgl. 3.2.2)¹⁸.

Auch in den Lehrplänen für Mathematik wird die Anwendung von Diagrammen in konkreten Anwendungssituationen thematisiert. Ziel dessen ist allerdings häufig die Vertiefung des mathematischen Verständnisses für die Funktion oder eine Steigerung der Motivation durch einen Alltagsbezug und nicht die inhaltliche Interpretation der Diagramme (Leinhardt et al., 1990). Es erscheint somit aufgrund der vorrangig abstrakten Verwendung in der Mathematik und der beschriebenen besonderen Anforderungen bei der Anwendung in konkreten Situationen notwendig, den Umgang mit den genannten Diagrammtypen in den naturwissenschaftlichen Unterrichtsfächern selbst zu schulen. Dies zeigt sich insbesondere daran, dass, wie im folgenden Exkurs ausgeführt, diese Schulung explizit in den deutschen Curricula verankert ist.

3.3 Exkurs: Curriculare Relevanz und Bildungsstandards

Der Umgang mit Diagrammen ist traditioneller Weise Thema der Mathematik. Entsprechend finden sich in den deutschen Lehrplänen¹⁹ dieses Unterrichtsfaches detaillierte Angaben zum Einsatz verschiedener Diagrammtypen und den jeweils an ihnen behandelten Fragestellungen (z. B. Kreisdiagramme zur Visualisierung von Brüchen und Winkeln (LP SH Mathe, S. 21, S. 28²⁰) oder Funktionsgraphen zur Veranschaulichung bestimmter Eigenschaften linearer Funktionen (LP SA Mathe, S. 55f)). Der Umgang mit Diagrammen ist jedoch ebenfalls explizit in deutschen Lehrplänen für die naturwissenschaftlichen Fächer verankert. Grundlegend wird für alle drei Fächer das selbstständige Konstruieren und Auswerten graphischer Darstellungen von Messdaten (insbesondere zum Auffinden von Zusammenhängen) als Lernziel festgelegt (LP SH Phy, S. 18; LP SA Bio, S. 99; LP HH, S. 91; LP Saar ChemGr, S. 6). Zudem sollen Schüler in der Lage sein, Diagramme zu interpretieren (LP Ni, S.14; LP Sach Chem, S. 6), bekannte Zusammenhänge in Diagrammen wiederzuerkennen sowie in graphische Darstellungen zu übertragen (LP Bb Bio, S. 24). Die genannten Fähigkeiten zur Verwendung von Diagrammen werden in einigen Lehrplänen als Bestandteil einer zu vermittelnden Methodenkompetenz (z. B. LP SH Bio, S. 17; LP Saar Bio9, S. 51) oder als

¹⁸ Leinhardt und ihre Kollegen (1990) vermuten aufgrund der Gewöhnung an zwei gleich-skalierte metrische Skalen durch Einführung der Diagramme im Mathematikunterricht Schwierigkeiten beim Vorkommen anderer Skalierungen im naturwissenschaftlichen Unterricht.

¹⁹ Unter diesem Begriff werden im Folgenden Rahmenpläne, Bildungspläne, Kerncurricula und alle weiteren Formen curriculärer Vorgaben/Richtlinien der Bundesländer zusammengefasst.

²⁰ Siehe gesondertes Lehrplanverzeichnis (Kapitel 10).

Ziel der allgemeinen Bildung (z. B. LP Bb Bio, S. 24) angesehen²¹. Dabei wird der Umgang mit Diagrammen einerseits als fachtypische Arbeitsweise betrachtet (LP Saar Bio9, S. 51), andererseits dessen fächerübergreifender und fächerverbindender Charakter hervorgehoben (LP Saar Bio9, S. 51; LP Bb Bio, S. 62).

Über diese grundlegenden Setzungen hinaus finden sich themengebundene Anlässe für die Vermittlung von Fähigkeiten zum Umgang mit Diagrammen, von denen im Folgenden einige exemplarisch für jedes der drei naturwissenschaftlichen Fächer wiedergegeben werden. Genannt wird beispielsweise die Darstellung gewonnener Messdaten zum Basiskonzept Stoff-Teilchen (LP Ni, S. 52) und die Interpretation von Energiediagrammen in Chemie (LP Saar Chem8, S. 16; LP RP, S. 219, S. 243; LP Th Chem, S. 107) sowie die Verwendung von (Fluss-)Diagrammen über den Energiebedarf des Menschen zum Thema Ernährung (LP SA Bio, S. 66, S. 140f; LP MV Bio, S. 26; LP RP, S. 42) in Biologie. Ferner sollen in Biologie Diagramme zu Räuber-Beute-Beziehungen, insbesondere den Volterraschen Regeln, ausgewertet (LP SA Bio, S. 140f; LP Saar Bio9, S. 2; LP Th Bio, S. 51) und die Darstellung von Ergebnissen in Form von Diagrammen anhand der Analyse abiotischer Faktoren im Ökosystem gelernt werden (LP SA Bio, S. 143f; LP HH, S. 48; vgl. auch LP Saar Bio9, S. 2). In der Physik wird die Verwendung von Diagrammen so eng an Unterrichtsinhalte geknüpft, dass mehrere spezielle Diagramme aufgeführt werden. Diese sind beispielsweise $N(t)$ - bzw. $m(t)$ -Diagramme zum radioaktiven Zerfall²² (LP MV Phy, S. 26) und $s(t)$ -Diagramme sowie $v(t)$ -Diagramme²³ einer geradlinig gleichmäßig beschleunigten Bewegung beim Thema Kinematik und Dynamik (LP MV Phy S. 31; LP SH Phy, S. 97, LP HH, S. 97). Wenngleich in den genannten Diagrammen der funktionale Zusammenhang der beiden Variablen benannt wird, ist eine „behutsame Mathematisierung“ vorgesehen (LP Ni, S. 15; LP SH Phy, S. 97). Ziel ist weniger die mathematische Beschreibung von Diagrammen, sondern deren konkretes inhaltliches Verstehen.

Die genannten Beispiele verdeutlichen, dass die in der vorliegenden Arbeit fokussierten Fähigkeiten zum Umgang und Gebrauch von Diagrammen für alle drei naturwissenschaftlichen Fächer direkt lehrplanrelevant sind. Diagramme sollen in naturwissenschaftlichen Kontexten erstellt und hinsichtlich naturwissenschaftlicher Fragestellungen interpretiert werden. Es geht dabei vordergründig um die Erfassung und Beschreibung eines Zusammenhangs und weniger um die Identifizierung des zugehörigen mathematischen Funktionstyps. Durch die Einbettung in eine Methodenkompetenz werden

²¹ Insbesondere werden die Erfolge von Schülern beim Erklären und Auswerten von Diagrammen sowohl im schriftlichen als auch im mündlichen Bereich zur Leistungsmessung herangezogen (LP SA Bio, S. 25).

²² $N(t)$ bezeichnet die Anzahl der radioaktiven Teilchen zum Zeitpunkt t , $m(t)$ beschreibt die Masse des noch nicht zerfallenen radioaktiven Stoffes zum Zeitpunkt t .

²³ $s(t)$ beschreibt die zurückgelegte Strecke zum Zeitpunkt t , $v(t)$ bezeichnet die Geschwindigkeit zum Zeitpunkt t .

fachübergreifende Aspekte, sowie durch die inhaltsgebundene Vermittlung fachspezifische Aspekte des Umgangs mit Diagrammen aufgeführt.

Neben der grundlegenden Einbettung des Umgangs mit Diagrammen in die Methodenkompetenz und der Verankerung an definierten Inhaltsgebieten werden die Fähigkeiten zum Diagrammgebrauch in mehreren Lehrplänen (z. B. LP Ni; LP Berlin; LP BW; LP HH) explizit bestimmten *Kompetenzbereichen* zugeordnet. Dies lässt eine direkte Bezugnahme zu den nationalen Bildungsstandards erkennen, die im Folgenden kurz skizziert werden. Mit der Einführung der Bildungsstandards für den mittleren Schulabschluss (KMK, 2005a, 2005b, 2005c) erfährt der Unterricht an allgemeinbildenden Schulen eine Neuorientierung an Kompetenzen, die die Schüler bis zum Ende der Mittelstufe erworben haben sollen. Hierzu wurden von den Expertenkommissionen für die jeweiligen Fächer Kompetenzen in unterschiedlichen Bereichen beschrieben und Standards formuliert, die dazu dienen, den *Outcome* zu konkretisieren, seine Erreichung einer Überprüfung zugänglich zu machen (KMK, 2005d) sowie Ansätze zur gezielten Förderung im Unterricht zu liefern. Einen gemeinsamen strukturellen Rahmen der drei naturwissenschaftlichen Fächer Biologie, Chemie und Physik bilden die vier Kompetenzbereiche Fachwissen, Erkenntnisgewinnung, Kommunikation und Bewertung, für die jeweils eigene Standards formuliert wurden.

Standards zum Umgang mit Diagrammen sind dabei vor allem dem Kompetenzbereich Kommunikation zugeordnet²⁴. Schüler sollen in der Lage sein, sich Wissen aus unterschiedlichen Quellen zu erschließen und anzueignen (KMK, 2005a, S. 15; KMK, 2005b, S. 12; KMK, 2005c, S. 12.). Dazu gehören insbesondere die Fähigkeiten, Aussagen unterschiedlicher Informationsquellen wie Texten, Bildern, Diagrammen und Tabellen zu erfassen und aufeinander zu beziehen (KMK, 2005a, S. 11). Diese Fähigkeiten werden als „wesentlicher Bestandteil einer erweiterten Lese- bzw. Verstehenskompetenz“ (KMK, 2005a, S.11) aufgefasst. Darüber hinaus sollen die Schüler über Fähigkeiten verfügen, gemessene Daten mit mathematischen und bildlichen Gestaltungsmitteln darzustellen (KMK, 2005a, S. 15) und auszuwerten (KMK, 2005c, S. 11).

In Anlehnung an die konkreten Standards sind in den entsprechenden Lehrplänen Erwartungen an Schülerfähigkeiten formuliert. So heißt es beispielsweise „Schülerinnen und Schüler stellen vorgegebene oder selbst ermittelte Messdaten eigenständig in Diagrammen dar und wählen dazu eine geeignete Diagrammform“ (LP Ni, S. 77) und „Schülerinnen und Schüler können Sachverhalte und Zusammenhänge in Diagramme übertragen“ (LP HH, S. 33). Darüber hinaus werden die Standards zum Teil ebenfalls an bestimmten Inhalten, insbesondere an den in den vorangegangenen Beispielen genannten,

²⁴ Die Schwerpunktsetzung unterscheidet sich dabei zwischen den Fächern. In der Chemie (KMK, 2005b) und der Physik (KMK, 2005c) fällt die Darstellung von erhobenen Daten sowohl in die Kommunikation als auch in die Erkenntnisgewinnung.

konkretisiert. Schüler sollen bezogen auf die Mechanik beispielsweise Weg-Zeit-Diagramme lesen und interpretieren (LP HH, S. 104; LP Berlin, S. 16), zur Veranschaulichung der Immunreaktion Diagramme erstellen (LP Berlin, S. 40) oder Werte aus einem Diagramm zur Optik entnehmen (LP HH, S. 104).

Aufgrund des gemeinsamen strukturellen Rahmens durch die Bildungsstandards ähnelt sich der Wortlaut der expliziten Angaben zum Umgang mit Diagrammen in allen drei naturwissenschaftlichen Fächern. Die verwendeten Diagramme sind allerdings zum Teil stark inhaltspezifisch, wie oben anhand der thematischen Bezüge dargelegt wurde. Aus diesem Grund ist es sinnvoll, die Fähigkeiten zum Umgang mit Diagrammen vorerst bezogen auf die einzelnen naturwissenschaftlichen Fächer zu beleuchten, so wie in der vorliegenden Arbeit für die Biologie durchgeführt. Bezugnehmend auf die Kompetenzbereiche in den Bildungsstandards ist dabei zu präzisieren, in welcher Weise die Fähigkeiten für die Verwendung von Diagrammen eine Kompetenz oder Teile einer Kompetenz darstellen. Im Folgenden wird daher der Begriff der Kompetenz näher beleuchtet.

3.4 Kompetenz und Kompetenzmodellierung

Wie oben ausgeführt, sind Fähigkeiten zum Umgang mit Diagrammen ein direktes Ziel schulischer (und insbesondere naturwissenschaftlicher) Bildung. Zur Beschreibung der Bildungsziele wird in der aktuellen Diskussion in den Erziehungswissenschaften der Begriff der Kompetenz herangezogen und zur Diagnose von Kompetenzen werden Kompetenzmodelle diskutiert (Klieme & Hartig, 2008). Im Folgenden wird die Verwendung des Begriffs der Diagrammkompetenz in der vorliegenden Arbeit dargelegt (3.4.1). Im Zusammenhang damit werden Grundlagen der Kompetenzmodellierung und der Überprüfung von Kompetenzmodellen vorgestellt (3.4.2).

3.4.1 Kompetenzbegriff

In der aktuellen fachdidaktischen Forschung wird der Kompetenzbegriff²⁵ nach Franz Weinert sehr häufig verwendet. Weinert (2001a) beschreibt Kompetenzen als „die bei Individuen verfügbaren oder von ihnen erlernbaren kognitiven Fähigkeiten und Fertigkeiten, bestimmte Probleme zu lösen, sowie die damit verbundenen motivationalen, volitionalen²⁶ und sozialen Bereitschaften und Fähigkeiten, die Problemlösungen in variablen Situationen erfolgreich und verantwortungsvoll nutzen zu können“. Wenngleich

²⁵ In der aktuellen nationalen sowie internationalen Bildungsforschung herrscht ein intensiver Diskurs über den Begriff der Kompetenz (vgl. z. B. Weinert, 2001a; Rychen & Salganik, 2001), der im Rahmen der vorliegenden Arbeit nicht wiedergegeben werden kann. Es sei auf Klieme und Hartig (2008) verwiesen, die diesen Diskurs aus erziehungswissenschaftlicher Perspektive nachzeichnen.

²⁶ Klieme (2003, S. 21) erklärt dazu „Volition = willentliche Steuerung von Handlungen und Handlungsabsichten.“

Weinert damit die Verbindung zwischen kognitiven und motivationalen Anteilen formuliert, spricht er sich selbst an anderer Stelle (2001b) für eine Einschränkung des Kompetenzbegriffs auf die kognitiven Faktoren aus. Hartig und Klieme (2006) leiten daraus für ihre Ausführungen den Kompetenzbegriff unter Ausschluss der motivationalen und affektiven Faktoren ab und fassen Kompetenzen als „kontextspezifische kognitive Leistungsdispositionen, die sich funktional auf bestimmte Klassen von Situationen und Anforderungen beziehen“, auf (Hartig & Klieme, 2006, S. 128). Beide Begriffsverwendungen stehen in unmittelbarem Bezug zueinander, sie unterscheiden sich lediglich darin, ob motivationale Faktoren der Kompetenz zugerechnet werden (Weinert) oder nicht (Hartig & Klieme). Um Missverständnissen vorzubeugen, wird in der vorliegenden Arbeit der Kompetenzbegriff, der auf den kognitiven Bereich einschränkt, als *Kompetenzbegriff nach Hartig und Klieme* bezeichnet.

In beiden Kompetenzbegriffen werden Kompetenzen nicht als abstrakte psychologische Konstrukte aufgefasst, sondern als kontextualisierte Dispositionen, Probleme in konkreten Anforderungssituationen lösen zu können (Klieme, 2003, S. 72). Diese Dispositionen lassen sich charakterisieren als Kenntnisse, Fertigkeiten und Routinen (Hartig & Klieme, 2006), sie stellen dadurch eine Verbindung von Wissen und Können (Klieme, 2003) dar. Dies bedeutet insbesondere, dass davon ausgegangen wird, dass Kompetenzen durch Erfahrung und Lernen in relevanten Anforderungssituationen erworben sowie durch äußere Interventionen beeinflusst werden können (Baumert, Stanat & Demmrich, 2001).

Die Kontextspezifität und die Erlernbarkeit von Kompetenzen stellen nach beiden Definitionen zentrale Charakteristika von Kompetenz dar (Klieme & Hartig, 2008). Durch diese Charakteristika lässt sich das Konzept der Kompetenz zum einen vom Konzept der Intelligenz abgrenzen. Intelligenz wird erstens durch generalisierte, auf nicht spezifizierte Situationen verallgemeinerbare Leistungsdispositionen charakterisiert, zweitens als relativ stabiles und damit als kaum erlernbares Persönlichkeitsmerkmal angesehen (Hartig & Klieme, 2006). Zum anderen unterscheidet insbesondere die Kontextspezifität die hier vorgestellte Auffassung von Kompetenz von der Auffassung, Kompetenzen seien allgemeine, bereichsübergreifende Qualifikationen. Solch ein weitfassender Kompetenzbegriff ist grundlegender Bestandteil in der Kompetenzkonzeption Heinrich Roths (1971, zitiert in Klieme & Hartig, 2008), in welcher die nicht kontextbezogenen Bereiche Selbst-, Sach- und Sozialkompetenz unterschieden werden. Diese Unterscheidung spielt eine bedeutsame Rolle in der Diskussion um Schlüsselqualifikationen oder Schlüsselkompetenzen (Klieme & Hartig, 2008) und wird insbesondere in manchen Lehrplänen (z. B. LP SH Bio; LP MV Bio; LP Th Bio) aufgegriffen²⁷. Als mögliche Begründung für die aktuelle Akzeptanz des Kontextbezugs, wie er im Kompetenzbegriff nach Weinert und auch nach Hartig und Klieme enthalten ist (im Vergleich zu der

²⁷ Dort wird die Sachkompetenz in Fach- und Methodenkompetenz ausdifferenziert.

Auffassung der bereichsübergreifenden Schlüsselqualifikationen), führen Schecker und Parchmann (2006) Ergebnisse aus der Expertiseforschung (z. B. Friege & Lind, 2003) an, nach denen Fähigkeiten, die als fachübergreifend angesehen wurden, sehr stark vom betrachteten Gegenstandsbereich abhängen.

Wie bereits erwähnt, unterscheiden sich die beiden hier vorgestellten Kompetenzbegriffe durch die Einbeziehung bzw. den Ausschluss motivationaler Faktoren. In der derzeitigen Kompetenzforschung in Deutschland besteht indes Konsens darüber, dass kognitive und motivationale Faktoren in Bezug auf Kompetenz in empirischen Untersuchungen getrennt erfasst werden sollten, nicht zuletzt um ihre Beziehungen aufklären zu können (KMK, 2005a, S.11). So bezieht sich beispielsweise das von der deutschen Forschungsgesellschaft geförderte Schwerpunktprogramm „Kompetenzmodelle zur Erfassung individueller Lernergebnisse und zur Bilanzierung von Bildungsprozessen“ (Klieme & Leutner, 2006) ebenfalls auf kognitive Leistungsdispositionen. Ebenso werden in den PISA-Studien motivationale Aspekte getrennt von den Aufgaben zur Messung der Kompetenz in den drei Bereichen Lesen, Mathematik und Naturwissenschaften erfasst (z. B. Baumert et al., 2001, S. 57ff).

Vor diesem Hintergrund wird in der vorliegenden Arbeit der Kompetenzbegriff nach Hartig und Klieme herangezogen. Danach wird unter der **Diagrammkompetenz** ein Gefüge kognitiver Leistungsdispositionen verstanden, das Schüler dazu befähigt, mit Diagrammen anforderungsgemäß umzugehen. Die Anforderungen, die beim Einsatz von Diagrammen im Biologieunterricht als Anwendungsfall für den naturwissenschaftlichen Unterricht entstehen, bilden den Kontext, der sich beispielsweise vom Gebrauch von Diagrammen im Mathematikunterricht unterscheidet (vgl. 3.2.3). Diagrammkompetenz wird als nicht von Natur aus gegeben angesehen, sondern muss erlernt werden (Dreyfus & Eisenberg, 1990).

Aufgrund des dargelegten engen Bezugs der Kompetenzbegriffe nach Hartig und Klieme sowie nach Weinert lässt sich Diagrammkompetenz im hier beschriebenen Sinne als der kognitive Anteil einer Kompetenz nach Weinert bezeichnen. Damit ist eine direkte Anknüpfung an das Verständnis von Kompetenz in den nationalen Bildungsstandards für den mittleren Schulabschluss (KMK, 2005a,b,c) möglich, in denen der Kompetenzbegriff nach Weinert (ebenfalls mehr oder weniger eingeschränkt auf die kognitiven Faktoren) verwendet wird. In dieser Hinsicht kann Diagrammkompetenz aufgefasst werden als Anteil der Kommunikationskompetenz (vgl. 3.3).

3.4.2 Kompetenzmodelle und ihre Überprüfung

Ein zentrales Element bei der Erfassung und Erforschung von Kompetenzen sind Kompetenzmodelle. Die Operationalisierung einer Kompetenz sollte sich nach Klieme (2003, S. 73) auf konkrete Anforderungssituationen beziehen; Kompetenzmodelle

vermitteln zwischen dem Verhalten von Personen in diesen konkreten Anforderungssituationen und den Kompetenzen als abstrakten Bildungszielen (Klieme, 2003, S. 71). Damit stellen sie das Bindeglied zwischen dem pädagogischen Konstrukt „Kompetenz“ und realen Aufgabensammlungen dar, indem in Kompetenzmodellen konkrete bereichsspezifische Leistungserwartungen formuliert werden und dadurch präzisiert wird, welches Verhalten bei der Bearbeitung von Testaufgaben als Hinweis auf die Kompetenz einer Person zu werten ist (Klieme, 2008, S. 24).

Dabei bedürfen Kompetenzmodelle der empirischen Überprüfung, bei der beleuchtet wird, ob sich die im Modell formulierten Aspekte der Kompetenz tatsächlich bei den Lernenden nachweisen lassen (Klieme, 2003, S. 82). Schecker und Parchmann (2006) unterscheiden diesbezüglich zwischen normativen und deskriptiven Kompetenzmodellen. In normativen Modellen werden theoretisch begründet Strukturen der Kompetenz formuliert, über die Lernende verfügen sollen, während in deskriptiven Modellen bereits empirisch bestätigte Strukturen und typische Muster der kognitiven Dispositionen wiedergegeben werden.

Theoretisch begründete, normative Kompetenzmodelle dienen einerseits der Konstruktion von Tests, mithilfe derer spezifische Stärken und Schwächen einzelner Schüler aufgedeckt werden können, und damit der Individualdiagnostik und Förderung einzelner Schüler im schulischen Alltag (Klieme, 2003, S. 83). Andererseits lassen sich die Modelle selbst mithilfe der Tests hinsichtlich der Angemessenheit ihrer Strukturen evaluieren. Dabei wird darauf hingewiesen, dass es für diesen Zweck sinnvoller erscheint, einen kleineren Kompetenzbereich detaillierter zu erfassen, als ein großes Spektrum mit wenigen Aufgaben abzudecken (Klieme, 2003, S. 83).

Modelle, die ein Gefüge von Teilaspekten einer Kompetenz beschreiben, werden häufig als *Kompetenzstrukturmodelle* bezeichnet (z. B. Schecker & Parchmann, 2006). Im Gegensatz dazu werden in *Kompetenzstufenmodellen*²⁸ spezifische Fähigkeiten bei unterschiedlichen Ausprägungen einer Kompetenz beschrieben (Hartig & Klieme, 2006). In vielen Fällen werden die Kompetenzstufen dabei erst auf der Basis von Messergebnissen konstruiert. Dazu werden die Schwierigkeiten der eingesetzten Testaufgaben auf einer Kompetenzskala verortet, welche nachträglich in Abschnitte eingeteilt wird. Die Beschreibung der Kompetenzstufen erfolgt dann anhand der spezifischen Eigenschaften der zur Messung verwendeten Aufgaben (vgl. Hartig, 2007). Die Kompetenzstufenmodelle, die in den letzten Jahren aus den internationalen Schulleistungsstudien hervorgegangen sind, sind größtenteils auf diese Weise entstanden (Hartig, 2007) und sind demnach als deskriptive Kompetenzmodelle zu bezeichnen.

Kompetenzstrukturmodelle sind im Gegensatz dazu oft normativer Natur, durch ihre empirische Überprüfung entstehen deskriptive Modelle. Die bei der Überprüfung von

²⁸ die auch als Kompetenzentwicklungsmodelle (Hammann, 2004; Schecker & Parchmann, 2006) bezeichnet werden.

Kompetenzstrukturmodellen zentrale Fragestellung ist demnach, welche Teilaspekte der Kompetenz differenziert erfasst werden können (Hartig & Klieme, 2006, S. 132). Dazu eignen sich mehrdimensionale probabilistische Testmodelle (Klieme & Hartig, 2008, S. 25), in denen jeweils unterschiedliche Teilaspekte getrennt oder zusammengefasst als Dimensionen abgebildet werden können. Ist das Modell, das die gemessenen Daten am besten beschreibt, dasjenige, welches die theoretisch postulierten Kompetenzstrukturen abbildet, werden diese Strukturen als validiert angesehen.

Klieme und Hartig (2006, S. 131) heben hervor, dass sich die Strukturierung von Kompetenzen an den zu bewältigenden Anforderungen orientiert. Vor diesem Hintergrund wird in der vorliegenden Arbeit ein *Kompetenzstrukturmodell* der Diagrammkompetenz theoretisch abgeleitet, in dem die Anforderungen beim Umgang mit Diagrammen beschrieben werden. Basierend auf diesem Modell werden Testaufgaben entwickelt, mithilfe derer das Kompetenzmodell empirisch überprüft wird. Dadurch wird die Grundlage gebildet, Unterschiede in den Fähigkeiten von Schülern beim Diagrammgebrauch im Biologieunterricht zu erfassen und so aufzudecken, an welcher Stelle eine differenzierte Förderung anzusetzen hat, um die Schüler zu einem lernförderlichen Gebrauch von Diagrammen zu befähigen.

Zur Überprüfung des genannten Kompetenzstrukturmodells kommen in der vorliegenden Arbeit *mehrdimensionale* Rasch-Modelle zum Einsatz (vgl. 5.1.2.4). Insbesondere die Verwendung des Dimensionsbegriffs bei der Modellierung von Kompetenzen mit mehrdimensionalen Rasch-Modellen kann allerdings zu Verwirrung führen, wenn die Teilaspekte von Kompetenzen als *Kompetenzdimensionen* bezeichnet werden. Zudem wird der Begriff der Kompetenzdimension in der aktuellen Forschung keineswegs einheitlich verwendet. So wird beispielsweise im nationalen Naturwissenschaftstest von PISA 2003 die naturwissenschaftlichen Kompetenz als bestehend aus zwei *Facetten* angenommen, deren unterschiedliche Ausprägungen als *Dimensionen* bezeichnet werden (Senkbeil et al., 2005). In anderen Beschreibungen gliedert sich eine Kompetenz in mehrere *Dimensionen* mit verschiedenen *Komponenten* (Schecker & Parchmann, 2006).

Aufgrund der möglicherweise irreführenden Konnotation wird der Begriff der Dimension in der vorliegenden Arbeit lediglich für die Beschreibung der Rasch-Modelle herangezogen. In Anlehnung an Klieme (2003, z. B. S. 74), der Kompetenzstrukturmodelle auch als *Komponentenmodelle* bezeichnet, wird für die unterschiedlichen Anteile der Diagrammkompetenz durchweg der Begriff *Komponente* verwendet. Eine Komponente kann dabei in weitere *Unterkomponenten* gegliedert sein und eine Unterkomponente kann aus mehreren *Aspekten* bestehen. Damit werden auf der theoretischen Ebene *Komponenten der Diagrammkompetenz* formuliert, die auf der methodischen Ebene durch *Dimensionen des Rasch-Modells* abgebildet werden (vgl. 5.1.2.4).

3.5 Umgang mit Diagrammen

Die in Kapitel 2 beschriebenen Anforderungen beim Umgang mit Diagrammen im Biologieunterricht können in zwei Bereiche eingeteilt werden: Erstens müssen Lernende Diagramme zur Visualisierung gegebener oder selbst erhobener Daten eigenständig konstruieren (vgl. 2.1). Die eigenständige Konstruktion ist insbesondere gefordert, wenn sich Lernende durch die aktive Auseinandersetzung mit einem Inhalt Wissen aneignen sollen (vgl. 2.2, Lernmittel). Zweitens müssen Diagramme sowohl in ihrer Funktion als Lehrmittel (vgl. 2.2) als auch bei der Datenauswertung (vgl. 2.1) von den Lernenden gelesen und interpretiert werden. In beiden Bereichen ist es erforderlich, Diagramme in Bezug zu textlichen Darstellungen zu setzen. Letztere bezeichnen zum einen in Textform bereitgestellte Informationen, auf deren Grundlage ein Diagramm konstruiert oder auf die es bezogen wird. Zum anderen werden darunter die Texte, die als schriftliche, sprachliche Äußerungen beim Lesen eines Diagramms von den Schülern produziert werden, verstanden.

Vor diesem Hintergrund scheint es in einem Modell der Diagrammkompetenz angebracht, die kognitive Verarbeitung von Text zu einem gewissen Maße zu berücksichtigen. Kognitionspsychologische Ansätze dazu liefert das integrative Modell der Text- und Bild- bzw. Diagrammverarbeitung²⁹, das von der Arbeitsgruppe um Wolfgang Schnotz entwickelt wurde (2001; Schnotz & Bannert, 2003). Dieses enthält Annahmen darüber, wie aus Text und Diagramm entnommene Informationen zueinander in Beziehung gesetzt werden, und liefert damit die kognitionspsychologischen Grundlagen für die Modellierung der Diagrammkompetenz, die im Folgenden vorgestellt werden.

3.5.1 Kognitionspsychologische Grundlagen

Das Modell der Text- und Diagrammverarbeitung nach Schnotz und Mitarbeitern (Schnotz, 2001; Schnotz & Bannert, 2003) greift die bereits in Abschnitt 3.1.1 vorgestellte Unterscheidung von deskriptionalen Repräsentationen, zu denen Texte zählen, und depiktionalen Repräsentationen, zu denen Bilder und Diagramme gehören, auf. Grundlegende Annahme ist, dass die verschiedenen externen, d. h. real vorliegenden, Repräsentationen mit verschiedenen mentalen Repräsentationsformen der Speicherung im Arbeitsgedächtnis korrespondieren. Dabei handelt es sich um eine propositionale Repräsentation im deskriptionalen Fall sowie um ein mentales Modell im depiktionalen Fall (Schnotz, 2001). Propositionale Repräsentationen bestehen aus einzelnen Propositionen. Sie werden als mentale Gedächtniseinheiten aufgefasst, die in ihrer Struktur Sätzen der gesprochenen Sprache ähneln und einen Gegenstand mit Hilfe mentaler Symbole *beschreiben* (Schnotz, 2001). Im Gegensatz dazu stellen mentale Modelle

²⁹ Unter Bildern versteht Schnotz an dieser Stelle realistische Bilder. Da das Modell im Folgenden nur in Bezug auf Diagramme betrachtet wird, wird es zur Vereinfachung als Modell der Text- und Diagrammverarbeitung bezeichnet.

Objekte dar, in denen die realen oder logischen Strukturen eines Gegenstandes respektive eines Sachverhalts analog *abgebildet* sind (Schnotz & Bannert, 2003).

Schnotz (2001) geht davon aus, dass Text und Diagramm als unterschiedliche Repräsentationsformen auf unterschiedliche Weise verarbeitet werden. So wird bei der Verarbeitung eines Textes eine propositionale Repräsentation gebildet (deskriptionale Repräsentationen), bei der Verarbeitung eines Diagramms ein mentales Modell (depiktionale Repräsentationen)³⁰. Mentales Modell und propositionale Repräsentation stehen in Wechselwirkung und sind ineinander überführbar. So können anhand eines mentalen Modells Informationen in Form von Propositionen generiert werden, was Schnotz (2001) als *Modellinspektion* bezeichnet. Bei dem umgekehrten Prozess, der *Modellkonstruktion*, wird basierend auf der propositionalen Repräsentation ein mentales Modell aufgebaut oder ein bereits bestehendes Modell verändert (Schnotz, 2001).

Bei der gleichzeitigen Verarbeitung von Diagramm und Text, die aufeinander bezogen sind, dienen die Prozesse Modellinspektion und Modellkonstruktion dem Abgleich und der Zusammenführung von Informationen aus den beiden Quellen. Propositionale Repräsentationen, die aus der Modellinspektion des mentalen Modells eines Diagramms resultieren, werden zu denen in Beziehung gesetzt, die bei der Verarbeitung eines Textes generiert wurden. Gleichfalls erfolgt auch ein Abgleich der mentalen Modelle: Über die Modellkonstruktion können propositionale Informationen, die bei der Verarbeitung eines Textes gebildet wurden, in das mentale Modell des zugehörigen Diagramms integriert werden (Schnotz & Bannert, 2003). Wird beispielsweise zu einem Liniendiagramm, das die Individuenzahlen einer Schildlauspopulation über mehrere Monate hinweg zeigt (vgl. Abbildung 3.4), ein Text gelesen, der die Individuenzahlen einer Marienkäferpopulation in Relation zur den Schildläusen beschreibt (Räuber-Beute-Verhältnis), so ist anzunehmen, dass der Verlauf einer zweiten Linie in das mentale Modell, das bei der Verarbeitung des Diagramms gebildet wurde, eingefügt wird.

Propositionale Repräsentation und mentales Modell stehen als mentale Repräsentationen zueinander im gleichen Verhältnis wie Text und Diagramm als externe Repräsentationen (Schnotz, 2001). Zur Erläuterung der beschriebenen Wechselwirkung zwischen propositionaler Repräsentation und mentalem Modell (und damit zwischen Text und Diagramm) wird nachfolgend detaillierter auf die dem Modell zugrundeliegende Theorie zur Diagrammverarbeitung³¹ eingegangen.

³⁰ Vgl. dazu auch Webber und Feeney (2004), die empirische Hinweise darauf liefern, dass Diagramme in Form mentaler Modelle verarbeitet werden und diese insbesondere dazu genutzt werden, Schlussfolgerungen aus zwei aufeinander bezogenen Diagrammen abzuleiten.

³¹ Bei der Textverarbeitung beruft sich Schnotz (2001) u.a. auf Theorien von van Dijk & Kintsch (1983). Sie wird an dieser Stelle nicht näher ausgeführt, da sie für das Verständnis der genannten Wechselwirkung in Bezug auf Diagramme nachrangig ist.

Nach welchen *genauen* Mechanismen die Verarbeitung eines Diagramms abläuft, ist nach Schnotz (2002) noch weitestgehend unklar. Eine Theorie dazu beschreibt Schnotz (2001; Schnotz & Bannert, 2003) in seinem Modell, in dem er das Diagrammverstehen als aus zwei Schritten bestehend darstellt. Der erste Schritt ist die Wahrnehmung, bei der die graphischen Komponenten des Diagramms, wie Linien, Punkte und Flächen, unterschieden und identifiziert werden. Als zweiter Schritt erfolgt davon ausgehend die inhaltliche Analyse dieser wahrgenommenen graphischen Strukturen. Dabei werden die räumlichen Relationen des Diagramms semantisch interpretiert und auf analoge konzeptuelle Relationen des entstehenden mentalen Modells übertragen (Schnotz & Bannert, 2003)³². Daneben können bei der semantischen Interpretation auch auf *direktem* Wege Propositionen gebildet und so eine propositionale Repräsentation des Diagramminhalts aufgebaut werden (Schnotz, 2001). In diesem Fall führt die Verarbeitung der im Diagramm enthaltenen Informationen nicht unbedingt und zuerst zu einem mentalen Modell, wie es zunächst als Regelfall (s. o.) beschrieben wurde. Über den Prozess der Modellkonstruktion kann jedoch letztlich über den Umweg der propositionalen Repräsentation wiederum ein mentales Modell aufgebaut werden.

Unabhängig davon, zu welcher mentalen Repräsentation die semantische Analyse zunächst führt, nimmt Schnotz in Anlehnung an Pinker (1990) an, dass sie von bestimmten, im Langzeitgedächtnis gespeicherten kognitiven Schemata (nach Pinker (1990, S. 94) „graph schema“, im Folgenden als **Diagrammschema** bezeichnet³³) angeleitet wird. Die Annahme der Existenz solcher Diagrammschemata findet sich auch in anderen Theorien des Diagrammverstehens (Lohse, 1993; Carpenter & Shah, 1998; Freedman & Shah, 2002)³⁴. Diagrammschemata enthalten Wissen darüber, wie Informationen in Diagrammen gespeichert sind und wie sie ihnen entnommen werden (Pinker, 1990; Lohse, 1993; Schnotz, 2001). Beispielsweise ist damit gemeint, an welchen Stellen auf eine gegebene Frage hin Strukturen im Diagramm zu suchen sind, und wie eine im Diagramm wahrgenommene Information mit Inhalt gefüllt wird.

Wie oben erwähnt, führt die semantische Analyse eines Diagramms (unter Anleitung der Diagrammschemata) zur Speicherung der in ihm dargestellten Informationen direkt in Form von Propositionen oder in Form eines mentalen Modells, dem durch Modellinspektion Propositionen entnommen werden. Propositionen wiederum können nach dem Modell der integrativen Text- und Diagrammverarbeitung bei Bedarf in

³² Schnotz und Bannert (2003) bezeichnen diesen Prozess als „analogical structure mapping“ (S. 146).

³³ Der Begriff *graph* wird an dieser Stelle nicht mit *Graph* übersetzt, da der Graph im Deutschen eingeschränkt auf die Linie in einem Liniendiagramm verwendet wird. Stattdessen wird das graph schema als Diagrammschema übersetzt, da sich graph im Englischen, wie in 3.1.2 dargestellt, umfassend auf mehrere (Achsen-)Diagrammtypen bezieht.

³⁴ Da es nicht Ziel der vorliegenden Arbeit ist, die genannten Theorien gegeneinander abzuwägen, werden sie an diesem Punkt nicht eingehender behandelt. Vielmehr werden sie im Folgenden an geeigneten Stellen herangezogen, um Prozesse der Entnahme von Informationen aus Diagrammen zu beleuchten.

sprachliche Äußerungen umgesetzt und damit externalisiert werden (Schnotz, 2001). In beiden Fällen bildet damit diese Versprachlichung den Abschluss der Entnahme von Informationen aus dem Diagramm, die dabei insbesondere eine Übersetzung von einer depiktionalen zu einer deskriptionalen Repräsentation darstellt. Die bei der Diagrammverarbeitung aktivierten Diagrammschemata können somit als Grundlage dieses Übersetzungsprozesses von Diagramm zu Text angesehen werden. Offen ist jedoch, welche kognitiven Mechanismen bei der Konstruktion von Diagrammen wirksam sind. Analog zu der von Schnotz explizit aufgeführten Externalisierung der Propositionen auf der Seite der deskriptionalen Repräsentationen kann auf der Seite der depiktionalen Repräsentationen angenommen werden, dass die Diagrammschemata, die die semantische Analyse anleiten, auch in der umgekehrten Richtung wirken. Konstruktionsleistungen können demzufolge durch die Bildung einer propositionalen Repräsentation, dem darauf folgenden Aufbau eines mentalen Modells (Modellkonstruktion) sowie dessen Externalisierung in Form eines Diagramms erklärt werden. Entsprechend kann Konstruktion als Übersetzung von Text zu Diagramm aufgefasst und als der Informationsentnahme entgegengesetzt verlaufend gegenübergestellt werden.

Darüber hinaus erklärt das Modell von Schnotz (2001), welche Übersetzungs- und Umwandlungsprozesse beim Vorliegen der zwei externen Informationsquellen Text und Diagramm zur Integration der Informationen aus beiden Quellen in eine gemeinsame Repräsentation ablaufen. Über Modellinspektions- und Modellkonstruktionsprozesse werden die beiden unterschiedlichen mentalen Repräsentationen, die bei der Verarbeitung von Text und Diagramm generiert werden, zueinander in Beziehung gesetzt und in eine gemeinsame mentale Repräsentation integriert (Schnotz, 2001; Schnotz & Bannert, 2003). Je nach Art dieser gemeinsamen Repräsentation (propositionales Format oder mentales Modell) erfolgt deren Externalisierung auf die oben angesprochene Weise, also entweder in Form eines Texts oder eines Diagramms.

Wie zu Beginn dieses Kapitels erläutert, wird in der vorliegenden Arbeit unter *Text* sowohl geschriebener (Lehr- und Lernmittel-)Text als auch die schriftlichen Äußerungen der Schüler beim Interpretieren eines Diagramms aufgefasst. Bezug nehmend auf das vorgestellte Modell der integrativen Text- und Diagrammverarbeitung (Schnotz, 2001) können damit die Informationsentnahme aus Diagrammen respektive die Diagrammkonstruktion als Übersetzungsprozesse von Diagramm zu Text respektive von Text zu Diagramm aufgefasst werden. Darüber hinaus weist das Modell darauf hin, dass die beim Vorliegen von Text und Diagramm nötigen Verarbeitungsprozesse in Form der Integration anteilig Prozesse sowohl der Informationsentnahme als auch der Konstruktion umfassen.

3.5.2 Differenzierte theoretische Überlegungen und empirische Befunde

Um Strukturen für ein Modell der Diagrammkompetenz ableiten zu können, werden im Folgenden empirische Befunde vorgestellt, die auf Schwierigkeiten der Schüler und damit auf Unterschiede in ihrer Diagrammkompetenz respektive verschiedene Kompetenzstrukturen hinweisen können. An gegebener Stelle werden sie durch theoretische Überlegungen zu bestimmten Strukturen ergänzt. Gemäß den obigen Ausführungen zum integrativen Modell des Text- und Diagrammverstehens (Schnotz, 2001) können grundlegend drei Bereiche zum Umgang mit Diagrammen unterschieden werden, die Informationsentnahme (3.5.2.1), die Konstruktion (3.5.2.2) und die Integration (3.5.2.3).

3.5.2.1 Informationsentnahme

Bereits in Abschnitt 3.5.1 wurden differenziertere Modelle des Diagrammverstehens angesprochen. Diese Modelle stimmen darin überein, dass zu mindestens einem Zeitpunkt bei der Verarbeitung eines Diagramms Informationen, die in der graphischen Struktur des Diagramms enthalten sind (z. B. „Linie steigt an“) mit den im Diagramm dargestellten Variablen verknüpft werden müssen (Shah & Hoeffner, 2002), um zu einer inhaltlichen Aussage zu gelangen. Sie unterscheiden sich darin, *wann genau* im Verarbeitungsprozess diese Verknüpfung stattfindet. Carpenter und Shah (1998) beispielsweise gehen wie Schnotz (2001) davon aus, dass das Diagrammverstehen mit der Wahrnehmung einer bestimmten graphischen Struktur (z. B. einer ansteigenden Linie) im Diagramm beginnt. Die Diagrammschemata liefern das Wissen, dass die visuelle Information „ansteigende Linie“ semantisch als „je mehr, desto mehr“ zu interpretieren ist³⁵. Nach dieser Interpretation erfolgt die Verknüpfung mit den dargestellten Variablen, die dann beispielsweise zur Formulierung der Information „je höher die Umgebungstemperatur, desto höher der Wasserverlust der Maus“ führt. Nach dem Modell von Carpenter und Shah (1998) laufen beim Verstehen des gesamten Diagramms diese Schritte iterativ ab, bis alle graphischen Komponenten verarbeitet sind.

Im Vergleich dazu nimmt Bertin (1974) an, dass die *Identifizierung* der im Diagramm dargestellten Variablen die Voraussetzung für das Verstehen eines Diagramms ist. Er grenzt diese Identifizierung von den eigentlichen *Ableseprozessen*, die für ihn die Betrachtung der graphischen Strukturen beinhalten, ab. In Anlehnung an diese begriffliche Differenzierung werden die nachfolgenden Ausführungen in *Identifizierung* und *Ablesen* unterteilt.

³⁵ wenn die entsprechende Skalierung gegeben ist

Identifizierung

Nach Bertin (1974) ist die *Identifizierung* als eine Betrachtung des Rahmens anzusehen, durch die sich ein Diagrammleser einen Überblick darüber verschafft, worum es in dem Diagramm geht. Er vergegenwärtigt sich die aufgeführten Variablen, ordnet sie den graphischen Strukturen zu (Bertin, 1974), erkennt, welche Variable auf welcher Achse eingetragen ist, und erhält Hinweise darauf, welches die unabhängige und welches die abhängige Variable ist. In Liniendiagrammen kann bereits die Achsenbelegung einen Hinweis auf die Abhängigkeit geben, da es Konvention ist, die abhängige Variable auf der Ordinate einzutragen (vgl. 3.2.1). Darüber hinaus muss der Diagrammleser beim Vorliegen mehrerer abhängiger Variablen durch die Betrachtung der Beschriftung an den einzelnen Linien, Säulen etc. oder der Legende zuordnen, welches Symbol, welche Farbe oder welche Linie welche Variable repräsentiert.

Durch die genannten Identifizierungsprozesse ist es dem Betrachter möglich, ohne Beachtung der Datenpunkte bereits eine Aussage über den im Diagramm dargestellten Sachverhalt zu treffen. So kann für das Diagramm in Abbildung 3.4 ausgedrückt werden: „Das vorliegende Diagramm zeigt die Individuenzahlen von Schildläusen und Marienkäfern über einen bestimmten Zeitraum hinweg.“

Wenn auch von Bertin (1974) nicht aufgeführt, so kann die Beachtung der Skalierung der Identifizierung zugerechnet werden. Bezieht man die Skalenreichweite der unabhängigen Variable mit ein, wird die oben genannte Aussage präziser: „Das vorliegende Diagramm zeigt die Individuenzahlen von Schildläusen und Marienkäfern über einen Zeitraum von 100 Monaten hinweg.“ Neben der Präzisierung der Informationsentnahme kann das Einbeziehen der Skalierung insbesondere dann von Nutzen sein, wenn verschiedene Diagramme miteinander verglichen werden müssen. Eine entsprechende Aufgabe in der Studie von Kerslake (1981), bei der Schüler erkennen mussten, dass zwei Graphen, die unterschiedliche Steigungen haben, gleich sind, bzw. dass solche, die von der Steigung der Linie her gleich aussehen, aufgrund verschiedener Skalierungen nicht gleich sind, konnte von 63,4% der Vierzehnjährigen und 68,5% der Fünfzehnjährigen richtig beantwortet werden.

Hinweise auf dem Diagrammlesen vorausgehende Identifizierungsprozesse liefert die Studie von Maichle (1994). Sie stellte Schülern die Aufgabe, gegebene Aussagen zu einem Diagramm anhand des Diagramms auf ihre Richtigkeit hin zu überprüfen. Ohne explizite Aufforderung dazu vertieften sich einige Probanden zuerst in das Diagramm, benannten beispielsweise die Achsen, identifizierten die Einheiten und die Reichweite der Skalen, bevor sie sich der eigentlichen Überprüfung der Aussagen zuwandten. Die Personen, die ohne vorherige Identifizierungsprozesse sofort mit der Überprüfung der Aussagen begannen, benötigten für die dazu nötigen Ableseprozesse deutlich mehr Zeit.

Wenngleich später nicht ausführlich in das integrative Modell des Text- und Diagrammverstehens aufgenommen, greift Schnotz in einem früheren Beitrag, der sich speziell auf den Wissenserwerb mit Diagrammen bezieht (Schnotz, 1994), die Identifizierung in seiner Analyse des Diagrammlesens implizit auf. Er beschreibt vom Diagrammleser benötigte Vorkenntnisse über die Darstellungsform, beispielsweise das Wissen über die Konventionen zum Aufbau eines Liniendiagramms aus Abszisse, Ordinate, Skalen und Graph, also die Wissensstruktur, die als notwendig für die genannten Identifizierungsprozesse angesehen werden kann. In Anlehnung an Pinker (1990) verwendet Schnotz (1994) für diese Wissensstruktur den Begriff der globalen Graphikschemas (vgl. Diagrammschema in 3.5.1).

Im Vergleich zu den globalen Schemata nimmt Schnotz (1994) an, dass die eigentlichen Ableseprozesse, die sich auf Darstellungsdetails wie Kurvenverläufe, Minima und Maxima beziehen, von *lokalen* Graphikschemas geleitet werden. Das Ablesen wird im folgenden Absatz detaillierter beleuchtet.

Ablesen

Das Ablesen bezieht sich auf die Datenpunkte, die in einem Diagramm als Punkte einer Linie oder in Form einer Säule eingetragen sind. Es wird von Bertin (1974) in drei Stufen eingeteilt, welche durch die Art der Fragestellung charakterisiert werden.

Die erste „elementare“ Stufe des Ablesens bezieht sich auf ein *einzelnes Element* einer Variable, ist also z. B. die Antwort auf die Frage „Wie viele Schildläuse wurden an *einem bestimmten Tag* gezählt?“ (s. Abbildung 3.4). Bertin (1974) beschreibt als zweite „mittlere“ Stufe des Ablesens das Beantworten von Fragen, die sich aus *einer Gruppe von Elementen* ergeben, beispielsweise der Frage „Wie veränderte sich die Anzahl der Schildläuse *in den ersten fünf Monaten*?“ Die obere Stufe des Ablesens ist nach Bertin die Antwort auf die Frage, die sich auf *alle Elemente einer Variable* bezieht (z. B. „Wie änderte sich die Anzahl der Schildläuse *über den gesamten Zeitraum hinweg*?“). Diese Stufung Bertins wird von Wainer (1992) aufgegriffen und durch eine Beschreibung der auf jeder Ablesensstufe vorzunehmenden Handlung präzisiert (z. B. durch das Erkennen eines Trends in den Daten auf der zweiten Stufe und dem Vergleichen verschiedener Trends bzw. dem Auffinden von Gruppierungen auf der oberen Stufe). Er deutet damit an, dass das Ablesen von Informationen mit zunehmender Stufung komplexer wird.

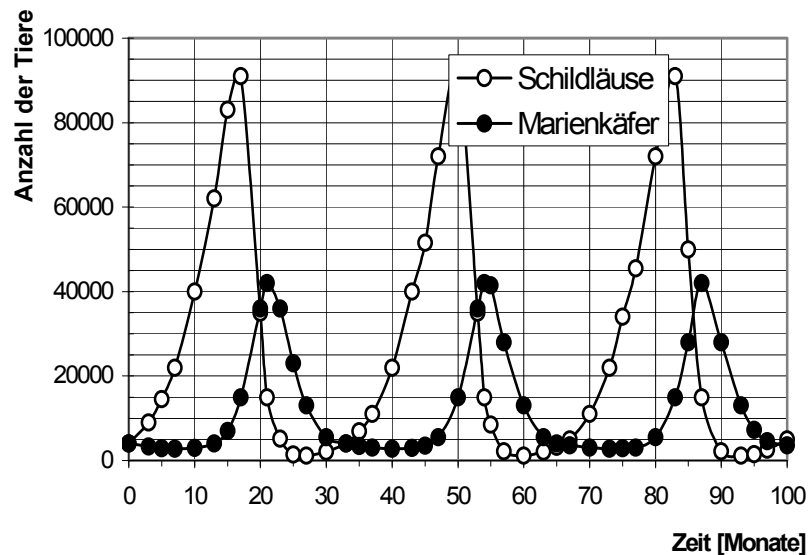


Abbildung 3.4: Liniendiagramm zur Populationsentwicklung von Schildläusen und Marienkäfern in einem Räuber-Beute-System.

Schnotz (1994) lehnt sich in seiner Beschreibung der drei Ablesestufen an die Kategorisierung Bertins und Wainers an und erweitert sie um eine detailliertere Betrachtung der Komplexität der Informationsentnahme auf der jeweiligen Stufe. Unter einem *Ablese* **erster Ordnung** versteht Schnotz (1994) ein Ablese, bei dem Attribute lediglich einer einzelnen graphischen Komponente bestimmt werden müssen. Im Falle eines Liniendiagramms bedeutet es, zu einem vorgegebenen Wert der unabhängigen Variablen die passende Koordinate der abhängigen Variablen, den „Funktionswert“, abzulesen („Wie viele Marienkäfer wurden im 25. Monat gezählt?“, vgl. Abbildung 3.4). Im Falle des Säulendiagramms bezieht es sich darauf, die Höhe einer Säule einer bestimmten Kategorie zu benennen.

Aufgaben zum Ablese erster Ordnung können von den Schülern recht erfolgreich gelöst werden, wie folgende Befunde belegen: Eine Aufgabe aus der TIMS-Studie, bei der die Schüler aus vier Alternativen die richtigen Koordinaten zu einem in ein Koordinatenkreuz gezeichneten Punkt auswählen sollten, wurde von den Achtklässlern zu 59% richtig gelöst (Baumert et al., 1998a). In einer Studie von Kerslake (1981) gelang Fünfzehnjährigen das Ablese der Koordinaten eines Punktes zu 73,6% bzw. 95,2%. Auch in einer Studie von Padilla, McKenzie und Shaw (1986) wurden die Koordinaten eines Punktes zu einem hohen Anteil von 84 % korrekt bestimmt.

Mit dem *Ablese* **zweiter Ordnung** bezeichnet Schnotz ebenso wie Bertin (1974) den Vergleich einzelner Werte. Ein solcher Vergleich kann dabei im Längsschnitt und im Querschnitt vorgenommen werden. Ein Längsschnittvergleich charakterisiert den Vergleich von Werten einer Datenreihe (z. B. „Die Individuenzahl der Schildläuse stieg in den ersten zehn Monaten an.“, vgl. Abbildung 3.4). Im Gegensatz dazu ist ein Vergleich der Populationsgrößen von Marienkäfern und Schildläusen (also ein Vergleich zwischen

den Datenreihen) zu einem bestimmten Zeitpunkt ein Querschnittsvergleich. Im Säulendiagramm bezeichnet ein Ablesen zweiter Ordnung im Längsschnitt das Vergleichen einzelner Säulenhöhen über die gegebenen Kategorien hinweg. Wenn mehrere Datenreihen vorhanden sind, können analog zum Querschnitt im Liniendiagramm die Säulenhöhen der verschiedenen Reihen in einer bestimmten Kategorie verglichen werden.

In Liniendiagrammen kann zusätzlich zwischen dem *Vergleich zweier Werte* und dem *Ablesen eines Trends* unterschieden werden (Maichle, 1994). Zur Benennung eines Trends müssen mindestens drei Werte einbezogen werden. So beinhaltet „Die Anzahl der Schildläuse stieg vom ersten bis zum vierten Monat an.“ eine relationale Aussage über die Werte im zweiten und dritten Monat, während durch den Vergleich zweier Einzelwerte („Im vierten Monat gab es mehr Schildläuse als im ersten.“) keine solche Aussage getroffen wird. Bezogen auf zwei Punktwerte differenziert Maichle zudem nach qualitativen Vergleichen („Die Anzahl der Schildläuse lag höher.“) und quantitativen Vergleichen („Die Anzahl lag um 5000 höher.“). Dabei werden für einen quantitativen Vergleich explizit Rechenfähigkeiten benötigt. Pereira-Mendoza und Mellor (1991) berichten davon, dass Schwierigkeiten von Schülern beim Vergleich zweier Werte häufig auf Probleme beim Vollziehen der notwendigen Rechenoperationen zurückgeführt werden können.

Unter einem *Ablesen dritter Ordnung* versteht Schnotz das Vergleichen mehrerer Trends bzw. das Vergleichen der Unterschiede zwischen jeweils zwei Werten (wie insbesondere im Säulendiagramm). Wie beim Ablesen zweiter Ordnung entspricht der Trendvergleich im Längsschnitt dem Vergleich innerhalb einer Datenreihe („Die Anzahl der Marienkäfer nahm erst ab und stieg dann wieder an.“) und im Querschnitt dem Vergleich zwischen zwei Datenreihen („Nachdem die Anzahl der Schildläuse zugenommen hat, beginnt die Anzahl der Marienkäfer zu steigen.“).

Auch in der Terminologie Curcios (1987) finden sich drei Ablesestufen. Im Unterschied zu Bertin und Schnotz beschreibt Curcio die Identifizierung nicht als einen dem Ablesen vorangehenden Prozess, sondern integriert zumindest Teile der Identifizierungsprozesse, beispielsweise das Ablesen der Achsenbeschriftungen, in ihre erste Ablesensstufe, dem *Lesen der Daten*. Abgesehen von diesem Identifizierungsanteil entspricht Curcios erste Stufe dem Ablesen erster Ordnung nach Schnotz. Curcios zweite Stufe, das *Lesen zwischen den Daten*, umfasst das Erkennen von Trends sowie deren Vergleiche und differenziert damit nicht zwischen dem Ablesen zweiter und dritter Ordnung. Demnach kann die Kategorisierung nach Schnotz als feinere Graduierung von Curcios erster und zweiter Stufe des Ablesens aufgefasst werden: Teile der Identifizierung und Ablesen erster Ordnung werden bei Curcio als Lesen der Daten bezeichnet, Ablesen zweiter und dritter Ordnung als Lesen zwischen den Daten. Curcio bezieht indes das *Lesen über die Daten hinaus* als höchste Kategorie in ihre Betrachtungen ein. Die Beantwortung von Aufgaben

dieses Typs (bzgl. der Abbildung 3.4 z. B. „Wie entwickeln sich beide Populationen nach den 100 im Diagramm dargestellten Monaten weiter?“) erfordert vom Diagrammläser, Schlussfolgerungen zu ziehen oder Vorhersagen basierend auf den gegebenen Daten zu machen. Eine entsprechende Aufgabe der TIMS-Studie, bei der zum Vorhersagen eines Wertes eine Gerade verlängert werden musste, wurde von den Schülern der achten Klasse in Deutschland zu 84 % richtig beantwortet (Baumert et al., 1998b). Die Lösungshäufigkeit einer ähnlichen Aufgabe, bei der eine Kurve verlängert werden musste, lag hingegen bei 59 % (Baumert et al., 1998a). Nach Pereira-Mendoza und Mellor (1991) äußerten Probanden bezüglich solcher Extrapolationen häufig, die Aufgabe nicht lösen zu können, da die betreffenden Werte im Diagramm nicht dargestellt seien. In ihrer Studie konnte eine Aufgabe zum Lesen über die Daten hinaus nur von 18 % der teilnehmenden Sechstklässler richtig gelöst werden, während das Lesen der Daten 98 % und das Lesen zwischen den Daten 78 % der Probanden gelang.

3.5.2.2 Konstruktion

Wie in Abschnitt 3.2.3 ausgeführt, bezieht sich die Konstruktion von Diagrammen im Biologieunterricht, im Vergleich zur Konstruktion eines Funktionsgraphen zu den Variablen x und y , auf die Anwendung konkreter Variablen. Grundsätzlich zeigen sich dabei Schwierigkeiten sogar in der gymnasialen Oberstufe, wie das folgende Beispiel aus der TIMS-Studie belegt. Die Lösungswahrscheinlichkeit einer Aufgabe, bei der ein Liniendiagramm angefertigt werden sollte, das die Beziehung zwischen der Körpergröße einer Person und ihrem Alter von der Geburt bis zu einem Alter von 30 Jahren zeigt, lag bei 41 % (Baumert et al., 1999).

Da, wie in Abschnitt 3.5.1 bereits erwähnt, kein kognitionspsychologisches Modell zur Konstruktion von Diagrammen bekannt ist, werden im Folgenden empirische Befunde aufgeführt, die sich differenzierter auf bestimmte Schwierigkeiten von Schülern beim Konstruktionsprozess in konkreten Anwendungssituationen beziehen und dadurch auf Fähigkeitsstrukturen hinweisen können.

So ist die Wahl des Diagrammtyps, der gegebenen Daten bzw. einem gegebenen Sachverhalt angemessen wäre (vgl. 3.2.2), als schwierig einzustufen. Aufgaben, bei denen Schüler der achten und neunten Klasse aus vier vorgegebenen Alternativen das passende Diagramm auswählen mussten, wurden in einer Studie von Baker, Corbett und Koedinger (2001) mit einer Wahrscheinlichkeit von weniger als 25 % richtig gelöst.

Schwierigkeiten bei der Belegung der Achsen (also der Festlegung, welche Variable auf welcher Achse eingetragen wird) und dadurch auch bei der Festsetzung der unabhängigen und abhängigen Variablen haben Padilla und Kollegen (1986) aufgedeckt. Von den Items, bei denen aus vier vorgegebenen Antwortalternativen diejenige Achsenbelegung

herausgesucht werden musste, die einer gegebenen Fragestellung entsprach, wurden im Mittel nur 46 % gelöst. Ein Beispiel für eine solche Aufgabe zeigt Abbildung 3.5.

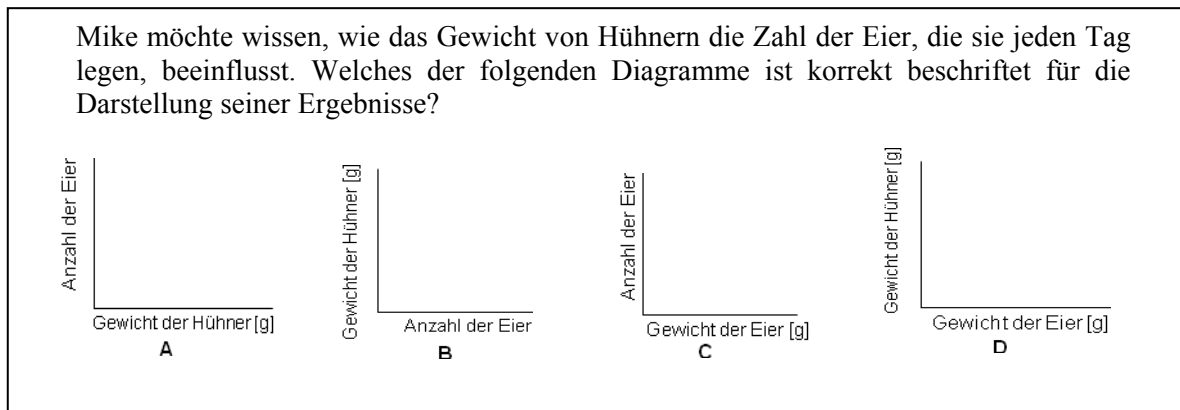


Abbildung 3.5: Beispiel für ein Item zur Achsenbelegung (adaptiert nach Padilla et al., 1986).

Zusätzlich scheint die Wahl der passenden Skalierung Schülern Schwierigkeiten zu bereiten. In der Erhebung von Padilla und Kollegen (1986) wiesen die entsprechenden Aufgaben zum Zeichnen einer Skala eine geringe mittlere Lösungshäufigkeit von 32 % auf und waren damit deutlich schwieriger als Aufgaben zu anderen Konstruktionsaspekten. Bei der Erstellung eines Liniendiagramms muss per Konvention für die unabhängige Variable eine metrische Skala gezeichnet werden. Verschiedene Autoren (Wavering, 1989, für Liniendiagramme; Baker et al., 2001, Baker et al., 2002, für Streudiagramme) berichten davon, dass Versuchspersonen vorgegebene (ungeordnete) Werte so auf der Skala anordnen, als wären sie kategorial, also z. B. 4 1 3 5, statt sie in der metrischen Reihenfolge aufzuführen. Ein weiterer typischer Fehler in der Studie von Wavering (1989) war das Eintragen der gegebenen Werte zwar in der richtigen Reihenfolge, aber in immer gleichen Abständen, unabhängig vom eigentlichen Wert, z. B. 1 3 4 5, so dass eine unregelmäßige und somit nicht-metrische Skala entstand. Auch Kerslake (1981) weist auf die Schwierigkeiten von Schülern hin, die Größe der Intervalle zwischen den beschrifteten Teilstrichen einer Achse auszuwählen. Befunde zu vergleichbaren Schwierigkeiten hinsichtlich der Skalierung einer kategorialen Variablen, beispielsweise dem Eintragen der Kategoriennamen in Säulendiagrammen, liegen nicht vor.

Im Vergleich zur Konstruktion von Teilen des Diagrammrahmens, also des Achsenkreuzes mit Skalierung, scheint den Schülern das Eintragen der konkreten Datenpunkte, das einem Ablesen erster Ordnung entspricht, leichter zu fallen. Die Studie von Kerslake (1981) zeigt, dass Dreizehn- bis Fünfzehnjährige das Eintragen von Punkten in Koordinatendarstellung in ein Koordinatensystem zu 89 % bis knapp 95 % korrekt ausführen konnten. Ähnlich hohe Lösungshäufigkeiten von 84 % wurden in der Studie von Padilla und Kollegen (1986) hinsichtlich des Eintragens von Datenpunkten ermittelt. Beiden Befunden unterliegt dabei die Verwendung ganzzahliger Werte der einzutragenden

Punkte. Beim Eintragen des Punktes (1½,4) zeigten sich hingegen geringere Lösungshäufigkeiten von knapp 77 % bis 80 % (Kerslake, 1981).

Hinsichtlich der Verwendung von Linien zur Punktverbindung liegt ein Befund von Kerslake (1981) vor. Er berichtet davon, dass Schüler die Verbindungslinie zwischen zwei Punkten nicht als Interpolation auffassen und nicht davon ausgehen, dass auf dieser Linie weitere, sogar unendlich viele Punkte liegen. 46 % der fünfzehnjährigen Probanden begründeten die Verbindung von Punkten mit „weil man die Punkte immer verbindet“ oder „weil es dann ordentlicher aussieht“ (Kerslake, 1981, S. 124).

Wie in Abschnitt 3.2.2 bereits angedeutet, können Liniendiagramme auch zur Darstellung genereller Trends genutzt werden, ohne dass konkrete Datenpunkte vorgegeben wären. Ein Beispiel wäre die Visualisierung der Aussage: „Je älter ein Baum ist, desto dicker ist sein Stamm.“ Aufgaben zur direkten Übersetzung dieser Art bearbeiteten Achtklässler in einer Studie von Mevarech und Kramarsky (1997). Dabei hatten einige Schüler Schwierigkeiten, eine „je mehr, desto mehr“-Aussage in Form einer ansteigenden Linie darzustellen. Stattdessen zeichneten sie ein Diagramm, das genau einen einzelnen Punkt (in Form einer Säule) enthielt.

3.5.2.3 Integration

In Abschnitt 2.2 wurde die Verwendung von Diagrammen als Informationsträger und damit als Lehrmittel im Biologieunterricht beschrieben. In vielen Fällen muss ein gegebenes Diagramm dabei zu zusätzlichen Informationen wie beispielsweise einem Lehrbuchtext in Beziehung gesetzt werden. Diese Bereitstellung birgt einerseits ein hohes Potential für den Wissenserwerb. Der als Multimedia-Effekt bekannte Befund, dass Lernende sich unter bestimmten Voraussetzungen mehr konzeptuelles Wissen aneignen, wenn sie im Vergleich zum reinen Text mit Graphiken und Text arbeiten, ist vielfach repliziert worden (Mayer, 2003). Mit dem hohen Potential geht jedoch eine besondere Anforderung einher, die daraus resultiert, dass zwei verschiedene externe Repräsentationen verarbeitet und aufeinander bezogen werden müssen (z. B. Ainsworth, 1999). Dies wird im Kompetenzstufen-Modell, das als Ergebnis der PISA 2000-Studie in Bezug auf die Lesekompetenz formuliert wurden, ebenfalls ausgedrückt. Nach diesem Stufenmodell stellt die Fähigkeit, verschiedene Texte aufeinander zu beziehen, eine höhere Kompetenzstufe dar als das Lesen eines einzelnen Textes (Artelt et al., 2001). Da Diagramme bei PISA zu einer bestimmten Art von Text gerechnet werden („nicht-kontinuierliche Texte“, vgl. 3.6.3), können unter verschiedenen Texten ebenfalls Text und Diagramm im Sinne der vorliegenden Arbeit verstanden werden.

Wie durch das integrative Modell der Text- und Diagrammverarbeitung (Schnotz, 2001) erklärt wird (vgl. 3.5.1), erfordert das Herstellen von Bezügen zwischen Text und

Diagramm die Fähigkeit zur Integration von Text- und Diagramminformationen in eine gemeinsame mentale Repräsentation. Die Externalisierung der letzteren kann, wie in Abschnitt 3.5.1 beschrieben, zur Produktion von Text als auch zur Produktion eines Diagramms oder einzelner Teile eines solchen führen. Insbesondere bei der Verwendung von Diagrammen als Lernmittel (vgl. 2.2), wenn die Schüler von der eigenständigen Konstruktion eines Diagramms profitieren sollen, beschreibt ein von den Lernenden selbst konstruiertes Diagramm das Resultat einer Integration von Text- und Diagramminformationen. Damit bezieht sich Integration in der vorliegenden Arbeit nicht nur auf die Entnahme von Informationen aus einem Diagramm und deren Bezug zu einem zugehörigen Text, der in einer textlichen Äußerung resultiert (*Resultat Text*), sondern auch auf die Konstruktion eines Diagramms basierend auf Informationen aus Text und gegebenem Diagramm (*Resultat Diagramm*).

Die besondere Anforderung beim Integrieren von Information besteht dabei zum einen darin, dass ein Lernender selbst erkennen muss, welche Informationen aus dem Diagramm abzulesen bzw. in ein Diagramm einzutragen sind. Dies wird deutlich anhand der Studien von Seufert (2003), in denen es ihr durch direkte Hinweise darauf, welche Strukturen in verschiedenen Repräsentationen aufeinander zu beziehen waren, gelang, die Integrationsleistung von Studenten zu erhöhen.

Grundsätzlich wird die Fähigkeit, basierend auf einer Fragestellung gerichtete Suchprozesse zum Auffinden der relevanten Information in einem Diagramm durchzuführen, ebenfalls auf ein Diagrammschema zurückgeführt (Lohse, 1993). Es wird allerdings keine Aussage darüber getroffen, ob dieses Diagrammschema auch anleitet, in einem gegebenen Szenario selbständig eine *Fragestellung* zur Informationsentnahme aufzufinden. Somit ist offen, ob sich beispielsweise die Fähigkeit, auf eine explizite Frage nach der Wassertemperatur ab einer bestimmten Wassertiefe (vgl. Abbildung 3.3) eine gleichbleibende Temperatur von 4°C abzulesen, von der Fähigkeit unterscheidet, bei einer zusätzlich gegebenen Information, wie der Erklärung über die Dichteanomalie des Wassers, die oben beschriebene Information aus dem Diagramm abzulesen und zu dieser Erklärung in Beziehung zu setzen.

Bei der bisherigen Beleuchtung der Informationsentnahme, Konstruktion und Integration wurden Befunde sowohl zu Linien- und Säulendiagrammen als auch zu anderen Achsendiagrammen wie beispielsweise Streudiagrammen aufgeführt. Im folgenden Abschnitt werden Studien herangezogen, die Unterschiede in der Verarbeitung der beiden in der vorliegenden Arbeit fokussierten Diagrammtypen Linien- und Säulendiagramm untersuchen.

3.5.3 Diagrammtypen

Linien- und Säulendiagramme sind, wie in Abschnitt 3.2.2 dargelegt, strukturell ähnlich. Sie gehören zu den Achsendiagrammen, stellen somit quantitative Zusammenhänge zwischen zwei Variablen dar, die in ihnen dargestellte abhängige Variable (AV) ist metrisch. Bezogen auf den Aufbau unterscheiden sie sich dahingehend, welches Skalenniveau die unabhängige Variable (UV) besitzt. Im Liniendiagramm wird eine metrische UV dargestellt, im Säulendiagramm eine ordinale oder nominale UV.

Wie bereits im Hinblick auf die Verwendung von Linien im Fall einer kategorialen UV angesprochen (vgl. Abschnitt 3.2.2), ist die Angemessenheit des Diagrammtyps deshalb relevant, weil sich gezeigt hat, dass die gewählte Darstellung in Form der Linien oder Säulen die Informationsentnahme beeinflusst. So werden aus Liniendiagrammen vorrangig Trendinformationen abgelesen, während anhand von Säulendiagrammen eher Vergleiche einzelner Werte vorgenommen werden (Zacks & Tversky, 1999; Shah, Mayer & Hegarty, 1999). Diese Tendenz wurde von Zacks & Tversky (1999) insbesondere auch dann beobachtet, wenn das entsprechende Skalenniveau der zugrundeliegenden UV das Ableiten eines Trends aus den Daten nicht zuließ. So zeigte sich, dass einige Studenten aufgrund einer „fälschlichen“ Verwendung von Linien Trendaussagen über Daten, denen eine nominal skalierte UV zugrundeliegt, aus einem Diagramm ableiten. Bezogen auf ein Diagramm, in dem die Blütendurchmesser verschiedener Pflanzenarten durch mit einer Linie verbundene Punkte (anstelle einzelner Säulen) dargestellt sind, wäre eine entsprechende Aussage „Je mehr Seerose, desto größer der Blütendurchmesser“, die durch die gewählte Darstellungsform des Liniendiagramms intendiert würde.

Über den Einfluss des Diagrammtyps (Säulen- oder Liniendiagramm) auf die Art der abgelesenen Informationen hinaus existieren einige empirische Befunde zur Frage, ob der Diagrammtyp auch die Schwierigkeit des Ablesens von Informationen beeinflusst. In den Studien von Jacobs (1994; 1995a, b³⁶) wurde ein Einfluss des Diagrammtyps auf die Schätzung einzelner Größenwerte (1995a) sowie das Erkennen des Verlaufs einer Datenreihe (1994, 1995b) untersucht. In Bezug auf die jeweiligen Fragestellungen wurden die benötigte Zeit und der Prozentsatz der richtigen Antworten betrachtet. Jacobs fand keine signifikanten Unterschiede zwischen Linien- und Säulendiagramm in Hinblick auf das Abschätzen einzelner Größenwerte (Jacobs, 1995a), das dem Ablesen erster Ordnung nach Schnotz (1994) entspricht³⁷. Das Erkennen eines Verlaufs gelang den Versuchspersonen in Linien- und in Säulendiagrammen vergleichbar schnell, wenn nur eine Datenreihe dargestellt wurde (Jacobs, 1994). Beim Vorliegen mehrerer (zwei bis vier)

³⁶ Jacobs untersuchte insbesondere verschiedene Anordnungen der Datenreihen (Juxta- und Superposition). In der vorliegenden Arbeit werden die Ergebnisse zur Superposition, das bedeutet, zur Auftragung mehrerer Datenreihen in einem Diagramm, herangezogen.

³⁷ In aktuellen Studien von Ratwani und Trafton (in press) hingegen zeigten sich in Bezug auf das Ablesen eines Punktwertes für Säulendiagramme kürzere Ablesezeiten als für Liniendiagramme.

Datenreihen wurden Verläufe im Liniendiagramm schneller erkannt als im Säulendiagramm (Jacobs, 1994, 1995b). In einem Fall ging damit ein signifikant höherer Prozentsatz richtiger Antworten bezogen auf das Liniendiagramm einher (Jacobs, 1995b). In Hinblick auf das Erkennen von Verläufen, das in den unterschiedlichen Studien einem Ablesen zweiter Ordnung (Jacobs, 1994) respektive zweiter und dritter Ordnung (Jacobs, 1995b) entspricht, zeigten sich damit leichte Vorteile des Liniendiagramms. Diese sind zwar zum Teil statistisch signifikant, jedoch die Effektstärken so gering, dass Jacobs selbst deren praktische Bedeutsamkeit diskutiert (1995b).

Ausgehend von diesen Befunden ist anzunehmen, dass im Vergleich von Linien- und Säulendiagramm der Diagrammtyp kein schwierigkeitsbestimmendes Merkmal ist, zumindest bezogen auf einzelne Aspekte der Informationsentnahme. Eine diagrammtyp-übergreifende Modellierung der Aspekte der Informationsentnahme ist damit denkbar. Aufgrund des Mangels an empirischen Befunden zum Einfluss des Diagrammtyps bezogen auf Konstruktions- und Integrationsprozesse wird diese übergreifende Modellierung auf diese Bereiche ausgeweitet. Das im folgenden Abschnitt vorgestellte Kompetenzstrukturmodell bezieht sich damit umfassend auf Linien- und Säulendiagramme.

3.6 Vorschlag eines Kompetenzstrukturmodells

In Abschnitt 3.5.1 wurde das Modell zum Text- und Diagrammverstehen der Arbeitsgruppe um Schnotz (2001, Schnotz & Bannert, 2003) vorgestellt. Bezugnehmend auf dieses Modell wurden die Fähigkeiten zum Umgang mit Diagrammen auf die drei Bereiche Informationsentnahme, Konstruktion und Integration aufgeteilt, welche in den vorangegangenen Abschnitten (vgl. 3.5.2) durch differenzierte empirische Befunde näher beleuchtet wurden. Weiterhin stellt diese Einteilung die Grundlage für das nachfolgend vorgeschlagene Kompetenzstrukturmodell der Diagrammkompetenz dar. Dort werden Informationsentnahme, Konstruktion und Integration aufgegriffen – sie werden im Folgenden als Komponenten der Diagrammkompetenz bezeichnet. Wie in Abschnitt 3.4.2 erwähnt, können Kompetenzkomponenten in Unterkomponenten und diese wiederum in Aspekte gegliedert sein. Die in Abschnitt 3.5.2 dargelegten Befunde zu Informationsentnahme, Konstruktion und Integration werden im Weiteren herangezogen, um die feinere Modellierung der drei Komponenten in Unterkomponenten und Aspekte detailliert zu beschreiben.

3.6.1 Informationsentnahme

Als wichtiger Teil der Informationsentnahme wurde in Abschnitt 3.5.2.1 die Identifizierung herausgestellt (Bertin, 1974; Carpenter & Shah, 1998). Diese bezeichnet genauer die Identifizierung der dargestellten Variablen und deren Verknüpfung mit den Diagrammstrukturen wie beispielsweise einem Linienverlauf. Carpenter und Shah (1998) beschreiben diese Verknüpfung als einen bei der Verarbeitung eines Diagramms auftretenden Prozess, der erst *nach* der Analyse einer graphischen Diagrammstruktur abläuft. Bertin (1974) hingegen trennt die Identifizierung von der Verarbeitung der Datenpunkte in einem Diagramm, dem Ablesen. Sie findet nach Bertin *vor* dem Ablesen statt und liefert die Antwort auf die Frage „Welche Relation wird im Diagramm dargestellt?“ (Bertin, 1974). *Wann* der Prozess der Identifizierung bei der Diagrammverarbeitung stattfindet, ist eine wahrnehmungspsychologische Fragestellung, die an dieser Stelle nicht näher beleuchtet werden kann. Gemäß der Theorie der Diagrammschemata kann allerdings angenommen werden, *dass* ein Schema *existiert*, welches speziell die Identifizierung anleitet. Zur Modellierung der Struktur der Diagrammkompetenz wird deshalb die Identifizierung als eine Unterkomponente der Informationsentnahme aufgefasst.

Die Fähigkeiten zur **Identifizierung** beziehen sich genauer auf folgende Prozesse: das Erkennen der dargestellten Relation, die Zuordnung der dargestellten Variablen zur jeweiligen Achse, die Zuordnung der einzelnen Datenreihen zu den Symbolen (beim Vorhandensein mehrerer abhängiger Variablen oder Datenreihen) sowie die Beachtung der Skalenreichweiten (vgl. 3.5.2.1). Diese sind in unten stehender Übersicht (Tabelle 3.1) untereinander eingetragen, wodurch allerdings keinerlei Stufung suggeriert werden soll.

Als weitere Unterkomponente wird in Anlehnung an Bertin (1974) das Ablesen von der Identifizierung abgegrenzt. In Abschnitt 3.5.2.1 wurden differenzierte Kategorisierungen verschiedener Stufungen und Strukturen des Ablesens dargestellt. Wie sie in die hier vorgestellte Modellierung der Informationsentnahme eingeflossen sind, wird nachfolgend begründet.

Hinsichtlich des **Ablesens** von Informationen scheint eine Kategorisierung anhand einer vorgegebenen Fragestellung nach Bertin (1974) problematisch, da der Anforderungsgrad des Ablesens stark durch die Datenpunkte eines Diagramms sowie den Verlauf eines Graphen oder die Unterschiede in den Säulenhöhen bestimmt wird und somit durch die Fragestellung nicht immer eindeutig festgelegt werden kann (Lohse, 1993). Aus diesem Grund erscheint es für die Modellierung der Ablesefähigkeiten sinnvoll, die einzelnen Kategorien durch die Charakteristika der kognitiven Operationen zu definieren, so wie es die Kategorisierungen von Curcio (1987) und Schnotz (1994) bzw. Maichle (1994) vorsehen. In Abschnitt 3.5.2.1 ist dargestellt worden, dass sich die dreistufigen

Einteilungen der Ableseprozesse von Schnotz und Curcio überschneiden. Hinsichtlich der ersten Kategorie stimmen sie weitgehend überein: Curcios Lesen der Daten entspricht (mit Ausnahme der Identifizierungsanteile) dem Ablesen erster Ordnung nach Schnotz. Hinsichtlich des Lesens zwischen den Daten nach Curcio stellt die Kategorisierung nach Schnotz in das Ablesen zweiter und dritter Ordnung eine feinere Unterteilung dar. Diese wird für die Modellierung des Ablesens herangezogen und um das Extrapolieren und Vorhersagen nach Curcio (Lesen über die Daten hinaus), das bei Schnotz nicht explizit berücksichtigt wird, erweitert. Für das Kompetenzmodell wird dabei die Terminologie nach Schnotz verwendet und das Extrapolieren und Vorhersagen als Ablesen vierter Ordnung bezeichnet, wie in der folgenden Tabelle aufgeführt.

Tabelle 3.1: Strukturmodell der Diagrammkompetenz. Durch * gekennzeichnete Operationen beziehen sich nur auf Liniendiagramme.

Informationsentnahme		Konstruktion		
Identifizierung	Erkennen der dargestellten Relation	Wahl des passenden Diagrammtyps		Aufbau des Rahmens
	Zuordnung der Variablen zu den Achsen	Zuordnung der Variablen zu ihren Achsen		
		Beschriftung der Achsen		
	Zuordnung der Datenreihen zu den Symbolen ("Legende")	Zeichnen einer Legende		
Beachten der Skalenreichweite	Zeichnen der Skalen			
Ablesen	1. Ordnung	Ablesen eines "Funktionswertes"	Eintragen der Punktwerte	Dateneintrag
	2. Ordnung	Vergleich zweier Werte oder Erkennen eines Trends (qualitativ/quantitativ)	Skizzierung einer Verbindungslinie zwischen Punkten oder freie Skizzierung einer Trendlinie*	
	3. Ordnung	Vergleich mehrerer Werte oder Vergleichen von Trends (qualitativ/quantitativ)	freie Skizzierung mehrerer Trends*	
	4. Ordnung	Extrapolieren/Vorhersagen		
Integration				

Die dargestellten vier Ordnungen des Ablesens charakterisieren die Binnenstruktur der Unterkomponente Ablesen. Unter Bezugnahme auf den in Abschnitt 3.5.2.1 dargestellten Befund von Pereira-Mendoza und Mellor (1991) kann vermutet werden, dass die einzelnen Ablesenordnungen in ihrer Schwierigkeit ansteigen und die Bewältigung der Aufgaben höherer Ordnungen einer höheren Ablesekompetenz bedarf. Insbesondere wenn man die feinere Unterteilung des Ablesens zweiter und dritter Ordnung in quantitative und qualitative Vergleiche (Maichle, 1994) einbezieht, weist der folgende Befund allerdings auf die Notwendigkeit für die empirische Überprüfung dieser Binnenstruktur hin: In einer Studie von Gillan und Lewis (1994) stellten die Probanden beim qualitativen Vergleichen zweier Punktwerte lediglich die Punkthöhen gegenüber, ohne die absoluten Werte zu bestimmen. Während für den quantitativen Vergleich zweier Werte und damit zum Ablesen zweiter Ordnung die Fähigkeit zum Ablesen erster Ordnung explizit benötigt wird, konnten die Schüler in der obigen Studie eine Aufgabe zum Ablesen zweiter Ordnung bewältigen, ohne dass sichergestellt wäre, dass sie die Fähigkeit zum Ablesen erster Ordnung besaßen.

3.6.2 Konstruktion

Die in Abschnitt 3.5.2.2 aufgeführten Befunde beschreiben einzelne Anforderungen beim Konstruieren von Achsendiagrammen. Da es in der Literatur an Modellen zur systematischen Beschreibung von Schülerfähigkeiten hinsichtlich des Konstruierens fehlt, werden diese Einzelbefunde herangezogen, um die Teilaspekte der Kompetenz von Schülern in Bezug auf die Konstruktion von Diagrammen zu modellieren. Durch Bezüge zur Informationsentnahme (s.u.) werden sie an geeigneter Stelle ergänzt.

Als ein Konstruktionsaspekt wird die Auswahl des Diagrammtyps (vgl. Baker et al., 2001) aufgefasst. Damit werden die Fähigkeiten von Schülern beschrieben, zu gegebenen Daten den angemessenen Diagrammtyp (Linien- oder Säulendiagramm, vgl. 3.2.2) zu bestimmen.

Als weiterer Aspekt der Konstruktion wird die Achsenbelegung (vgl. Padilla et al., 1986) aufgegriffen. Sie steht für die Fähigkeiten von Schülern, aus der Aufgabenstellung die relevanten Variablen zu erfassen, die unabhängige Variable (UV) und die abhängige Variable (AV) zu identifizieren und die UV (möglichst³⁸) an der Abszisse einzutragen.

Ferner wird die Skalierung (vgl. Padilla et al., 1986; Wavering, 1989; Baker et al., 2001; Baker et al., 2002; Kerslake, 1981) als Aspekt der Konstruktion einbezogen. Sie umfasst die Fähigkeiten von Schülern, im metrischen Fall eine gleichmäßige, ausreichend lange Skala zu erstellen und zu beschriften, im nominalen oder ordinalen Fall die Kategorienamen vollständig und an markierten Stellen der Achse einzutragen.

³⁸ Ein Eintragen der UV auf der Ordinate entspricht zwar nicht der Konvention (vgl. 3.2.1), ist aber nicht als falsch zu betrachten (vgl. Balkendiagramme).

Darüber hinaus wird das Eintragen der im Text vorgegebenen Werte auf der Diagrammfläche als Konstruktionsaspekt aufgegriffen. Es beschreibt die Fähigkeiten von Schülern, die Markierungen an den der gegebenen Skalierung entsprechenden Stellen mit hinreichender Genauigkeit vorzunehmen. Die in Abschnitt 3.5.2.2 aufgeführten Befunde zum Eintragen einzelner Werte (Kerslake, 1981; Padilla et al., 1986) beziehen sich auf Datenpunkte zu zwei metrischen Variablen. In Bezug auf Säulendiagramme muss an dieser Stelle berücksichtigt werden, dass zum Eintragen eines Wertes zusätzlich zur Festlegung der Säulenhöhe durch einen Punkt das Zeichnen der zweidimensionalen Säule erforderlich ist.

Eine weitere Konstruktionsleistung kann mit „Skizzierung“ betitelt werden; sie bezieht sich auf Liniendiagramme. Mit Skizzierung ist generell das Zeichnen von Kurvenverläufen gemeint. Es kann dabei zwischen dem Zeichnen einer Verbindungslinie zwischen vorgegebenen Punkten (vgl. Kerslake, 1981) und der freien Skizzierung eines generellen Trends unterschieden werden (Leinhardt et al., 1990). Im letzteren Fall beschreibt die Skizzierung die direkte Übersetzung einer Information wie „Die Anzahl der Marienkäfer nimmt zu.“ in die Form einer ansteigenden Linie im Diagramm (vgl. Mevarech & Kramarsky, 1997).

Leinhardt und ihre Kollegen (1990) stellen heraus, dass die gerade dargestellte Unterscheidung nicht nur im Hinblick auf die Konstruktion zutrifft, sondern sich auch im Hinblick auf die Informationsentnahme³⁹ wieder findet. Dort können, ähnlich den Beschreibungen von Maichle (1994), quantitative Betrachtungen und Vergleiche einzelner Punkte und qualitative Beschreibungen von Linienverläufen unterschieden werden (Leinhardt et al., 1990). Darüber hinaus sehen sie beispielsweise die Skalierung als relevanten Aspekt sowohl für die Informationsentnahme als auch für die Konstruktion an. Diese Parallelisierung einzelner Aspekte der Informationsentnahme und Konstruktion wird für das vorliegende Kompetenzmodell ausgeweitet, indem ergänzend zwei Konstruktionsaspekte aus den Strukturen abgeleitet werden, die bei der Informationsentnahme als bedeutsam herausgearbeitet wurden. So kann gemäß den Unterkategorien der Identifizierung davon ausgegangen werden, dass es speziell bei der vollständigen Beschriftung der Achsen zu Schwierigkeiten kommen kann. Die Achsenbeschriftung wird daher als weiterer Konstruktionsaspekt aufgegriffen. Sie bezieht sich auf die Schülerfähigkeiten, sowohl den korrekten Namen der Variablen als auch deren Maßeinheit an den Diagrammachsen einzutragen. Analog zur Unterscheidung verschiedener Datenreihen auf der Seite der Identifizierung kann in Bezug auf Diagramme mit mehreren Datenreihen das Zeichnen einer Legende als zusätzlicher Aspekt der Konstruktion eingestuft werden.

³⁹ Leinhardt und Kollegen (1990) gebrauchen den Begriff Interpretation.

Ausgehend von diesen Überlegungen werden in der Tabelle 3.1 die Aspekte der Informationsentnahme und Konstruktion einander zugeordnet. Insbesondere kann die freie Skizzierung dabei je nachdem, ob ein durchgehender Trend gezeichnet oder mehrere Trends zueinander in Beziehung gesetzt werden sollen, sowohl dem Ablesen zweiter als auch dritter Ordnung gegenübergestellt werden.

Die einzelnen Aspekte der Konstruktion werden nicht als Kompetenzstufen angesehen, sondern als unterschiedliche Strukturen aufgefasst, deren Relationen untereinander allerdings nicht geklärt sind. Es bedarf empirischer Forschung, die Relationen und Stufungen dieser Binnenstrukturen der Konstruktion offen zu legen und Hinweise darauf zu geben, inwieweit sie als voneinander unabhängig angesehen werden können.

In Anlehnung an die zwei Unterkomponenten der Informationsentnahme lassen sich aus den genannten Aspekten zwei Unterkomponenten der Konstruktion formieren, die als *Aufbau (des Rahmens)* und *Dateneintrag* bezeichnet werden. Erstere Unterkomponente bezieht sich auf diejenigen Konstruktionsaspekte, die der Erstellung des vollständigen Diagrammrahmens dienen, letztere auf das Eintragen von Werten und die Skizzierung.

Ausgehend von der Unterscheidung globaler und lokaler Diagrammschemata (vgl. 3.5.1) könnten in einer alternativen Modellierung der Diagrammkompetenz Identifizierung und Aufbau sowie Ablesen und Dateneintrag zu jeweils einer Komponente zusammengefasst werden. Dies würde bedeuten, dass innerhalb dieser Komponenten die Bezüge zwischen Informationsentnahme und Konstruktion sehr stark wären. Grundsätzlich finden sich bisher in der Literatur keine Angaben zur systematischen Untersuchung des Zusammenhangs zwischen Informationsentnahme und Konstruktion (Friel et al., 2001). Padilla und Kollegen (1986) erfassen zwar sowohl Aspekte der Informationsentnahme als auch der Konstruktion, beziehen die Ergebnisse allerdings nur dahingehend aufeinander, dass die Schwierigkeiten der Einzelaspekte verglichen werden.

Trotz gewisser Parallelen in den Aspekten von Informationsentnahme und Konstruktion (s. o.) sehen Leinhardt und ihre Kollegen (1990) die Konstruktion als der Informationsentnahme deutlich gegenübergestellt an. Sie begründen dies damit, dass sich die Informationsentnahme auf etwas Gegebenes (das Diagramm) bezieht, also eine Reaktion auf etwas Vorhandenes darstellt, während Konstruktion das Erschaffen von etwas Neuem, nicht Gegebenen bedeutet. Auch in der vorliegenden Arbeit wurden Informationsentnahme und Konstruktion unter Rückgriff auf das integrative Modell des Text- und Diagrammverstehens (Schnotz, 2001) als gegenläufige Übersetzungsprozesse gekennzeichnet (vgl. 3.5.1). Davon ausgehend beschreibt das hier vorgeschlagene Kompetenzstrukturmodell Informationsentnahme und Konstruktion als getrennte Komponenten. Die Relationen zwischen diesen beiden Komponenten sowie zwischen deren Unterkomponenten (s. o.) bleiben zu klären.

3.6.3 Integration

In Abschnitt 3.5.2.3 ist bereits erwähnt worden, dass im Vergleich zum Lesen eines einzelnen Textes die Fähigkeit, verschiedene Texte aufeinander zu beziehen, im Stufenmodell der Lesekompetenz der PISA 2000-Studie als eine höhere Kompetenzstufe angesehen wird. Gemäß der Konzeption von Lesekompetenz werden dabei kontinuierliche, fortlaufende Texte und nicht-kontinuierliche Texte, zu denen u. a. Diagramme zählen, unterschieden, so dass entsprechend das Aufeinanderbeziehen von Text und Diagramm (als verschiedenen Textsorten) im Vergleich zur Informationsentnahme aus einem einzelnen Diagramm (als einem nicht-kontinuierlichen Text) als höhere Kompetenz eingeordnet werden kann. Die Verarbeitung von Diagramm und zugehörigem Text wird in der vorliegenden Arbeit als Teil der Integration verstanden. Sie erfordert den Aufbau einer zusammengehörigen internen Repräsentation (Schnotz, 2001). Deren Externalisierung kann, wie in Abschnitt 3.5.2.3 dargestellt, zu zwei verschiedenen externen Repräsentationen führen: Dies können einerseits sprachliche Äußerungen in Form von Text sein (*Resultat Text*), andererseits kann ein Diagramm produziert werden (*Resultat Diagramm*), in welchem die Informationen der zusammengehörigen internen Repräsentation wiedergegeben werden.

Die erste Form der Externalisierung (*Resultat Text*) erfordert im Hinblick auf die Verarbeitung des Diagramms Prozesse der Informationsentnahme. Die zweite Form der Externalisierung (*Resultat Diagramm*) erfordert ebenfalls eine Verarbeitung der im anfänglichen Diagramm gegebenen Informationen, darüber hinaus allerdings auch Diagrammkonstruktionsleistungen. In der PISA 2000-Studie war die resultierende externe Repräsentation, mit der die Integrationsleistung abgeprüft wurde, Text, so dass entsprechend den genannten Lesekompetenzstufen die Integration von Text und Diagramm als höhere Stufe der Informationsentnahme bezeichnet werden kann. Da in der vorliegenden Arbeit Integration weiter gefasst wird und, wie bereits erwähnt, im Hinblick auf ein resultierendes Diagramm auch Konstruktionsprozesse beinhalten kann, wird sie im hier beschriebenen Kompetenzstrukturmodell nicht als höhere Stufe der Informationsentnahmekompetenz angenommen. Es wird gemäß den kognitionspsychologischen Überlegungen (vgl. 3.5.1) postuliert, dass sich die Integration als eigenständige Komponente der Diagrammkompetenz nachweisen lässt. Insbesondere liefert auch die Theorie der Diagrammschemata (vgl. Lohse, 1993) keine explizite Erklärung dafür, Integration als höhere Stufe einzig der Informationsentnahme aufzufassen.

Grundsätzlich wäre denkbar, dass sich entsprechend der zwei genannten resultierenden externen Repräsentationen (*Resultat Text* oder *Resultat Diagramm*) zwei Unterkomponenten der Integration nachweisen lassen. Wie oben angedeutet, geht allerdings auch der Externalisierung in Form eines Diagramms eine Entnahme der

Informationen aus dem Diagramm voraus, so dass für das hier beschriebene Kompetenzmodell keine Trennung der beiden Integrationsformen angenommen wird. Die Integration wird somit als Komponente ohne Unterkomponenten angesehen. Allerdings sind dabei Bezüge der Integration sowohl zur Informationsentnahme als auch zur Konstruktion aus den oben genannten Gründen naheliegend. Dies wird in der Übersicht des Kompetenzmodells (vgl. Tabelle 3.1) dadurch angedeutet, dass die Integration als quer liegend zu den beiden anderen Komponenten dargestellt ist.

3.7 Lernerhintergrund und Lernermerkmale

Im Folgenden werden empirische Befunde angeführt, die auf verschiedene Einflussfaktoren beim Umgang mit Diagrammen hinweisen. Zum einen wird dabei auf die Faktoren Geschlecht und Klassenstufe eingegangen, die im Folgenden als Lernerhintergrund bezeichnet werden. Darüber hinaus werden verschiedene Lernermerkmale, wie beispielsweise das Fachwissen der Lernenden, behandelt.

3.7.1 Lernerhintergrund

3.7.1.1 Klassenstufe

Wie in den aktuellen Schulleistungsstudien dokumentiert wurde, unterscheiden sich die Schüler in verschiedenen Klassenstufen im Hinblick auf verschiedene Kompetenzen. Beispielsweise zeigten in der PISA-2003-Erhebung der naturwissenschaftlichen Kompetenz, die den Umgang mit verschiedenen Graphiken einschließt, 44 % der untersuchten Schüler am Ende der zehnten Klasse signifikant höhere Testergebnisse als am Ende der neunten Klasse (PISA-Konsortium Deutschland, 2006b). Speziell für den Umgang mit Diagrammen wurden von Padilla und Kollegen (1986) mehrere Jahrgänge hinsichtlich verschiedener Konstruktions- und Interpretationsleistungen verglichen. In ihren Studien waren die Zehntklässler den Neuntklässlern signifikant überlegen. In der Studie von Curcio (1987) wurden Viert- und Siebtklässler verglichen. Sie unterschieden sich zum einen in ihren Fähigkeiten zur Informationsentnahme aus Diagrammen, zum anderen hinsichtlich ihres Vorwissens zu abstrakten Diagrammen. Letzteres drückt sich nach Curcio beispielsweise dadurch aus, in einem Säulendiagramm, das den Zusammenhang zwischen zwei abstrakten Variablen X und Y darstellt, die Säule bestimmen zu können, die die größte Quantität angibt. Besonders interessant ist, dass sich für die Schüler der vierten Klasse dieses Wissen zu abstrakten Diagrammen als Vorhersagewert für das Diagrammverständnis identifizieren ließ, für die Schüler der siebten Klasse jedoch nicht (Curcio, 1987). Dies deutet auf unterschiedliche Fähigkeitsstrukturen zur Informationsentnahme von Viert- und Siebtklässlern hin. Im Hinblick auf die Modellierung der Diagrammkompetenz ergibt sich daraus die Frage, ob sich die Schüler verschiedener Klassenstufen lediglich im Ausmaß ihrer Fähigkeiten oder

zudem in ihren Kompetenzstrukturen unterscheiden, d. h. insbesondere, ob ein Kompetenzmodell für mehrere Klassenstufen empirisch validiert werden kann.

3.7.1.2 Geschlecht

Neben der Klassenstufe ist in verschiedenen Studien ein Einfluss des Geschlechts vor allem auf das Verständnis von Diagrammen nachgewiesen worden. Beispielsweise wurden hinsichtlich der in den PISA-Studien erhobenen Lesekompetenz, die sich u.a. auf Diagramme bezieht, geschlechtsspezifische Unterschiede festgestellt. Im Gesamttest Lesen erzielten die Mädchen in allen Teilnehmerstaaten in der PISA 2000-Studie signifikant höhere Werte als die Jungen (Stanat & Kunter, 2002). Die Überlegenheit der Mädchen zeigte sich dabei besonders deutlich für fortlaufend geschriebene (kontinuierliche) Texte, die in PISA von den nicht-kontinuierlichen Texten, das bedeutet bildhaften Darstellungen, zu denen auch Diagramme zählen, abgegrenzt werden. Dennoch ist auch im Hinblick auf nicht-kontinuierliche Texte der Unterschied zwischen Mädchen und Jungen signifikant (Stanat & Kunter, 2002). Im Naturwissenschaftstest der PISA 2000-Studie hingegen waren die Schülerinnen den Schülern darin unterlegen, aus graphischen Repräsentationen (Diagrammen, Schemazeichnungen, Tabellen und Karten) die richtige Information abzuleiten (Prenzel et al., 2001).

Im Gegensatz dazu wurden in den Studien von Curcio (1987), die speziell auf das Verständnis von Diagrammen ausgerichtet sind, in Bezug auf Viert- und Siebtklässler keine geschlechtsspezifischen Unterschiede gefunden. Analog zu den Hinweisen bei der Klassenstufe ist offen, ob diese möglichen Kompetenzunterschiede zwischen Schülerinnen und Schülern bei der Modellierung von Diagrammkompetenz berücksichtigt werden müssen.

3.7.2 Lernermerkmale

3.7.2.1 Fachwissen

Der Einfluss des inhaltlichen Vorwissens auf die Nutzungseffizienz von bildhaften Darstellungen in Verbindung mit Text ist vielfach beschrieben worden (Schnotz, Zink & Pfeiffer, 1996; Schnotz & Bannert, 2003; Lewalter, 1997; Mayer & Gallini, 1990). Dabei zeigte sich, dass besonders Lernende mit geringem Vorwissen vom Einsatz von Bildern profitierten. Auch im Hinblick auf Diagramme ist das Vorwissen als bedeutsamer Einflussfaktor insbesondere für die Informationsentnahme identifiziert worden (Friel et al., 2001; Curcio, 1987). In verschiedenen Modellen der Diagrammverarbeitung wird das Vorwissen explizit als Einflussfaktor aufgenommen (Freedman & Shah, 2002; Mayer, 2003). In einer Studie von Preece und Janvier (1992) wurden die Interpretationen von Schülern hinsichtlich strukturell gleicher Liniendiagramme, die sich auf verschiedene Inhalte beziehen, verglichen. Es zeigte sich, dass Schüler, die keinerlei Vorwissen über das

jeweilige Inhaltsgebiet besaßen, keine inhaltliche Interpretation des Diagramms vollziehen konnten, sie beschrieben lediglich die oberflächlichen Merkmale des Linienverlaufs. Inhaltliches Vorwissen kann somit als Voraussetzung für eine erfolgreiche Interpretation angesehen werden. Darauf weisen auch Befunde von Roth und Bowen (2003) hin, nach denen Wissenschaftlern trotz Erfahrung im Diagrammgebrauch die Interpretation von Diagrammen in ihnen unvertrauten Inhaltsgebieten deutlich schwerer fiel.

Umgekehrt wurde allerdings bei vorhandenem Vorwissen eine Beeinflussung der Interpretation nachgewiesen (Preece & Janvier, 1992). In den Studien von Shah wurde deutlich, dass Versuchspersonen, die mit bestimmten inhaltlichen Vorstellungen („je mehr betrunkene Fahrer, desto mehr Unfälle“) ein Diagramm betrachteten, vorrangig Trends ablesen, die mit diesen Vorstellungen konform gingen, selbst wenn die Daten nicht konsistent mit diesem generellen Trend waren (Shah, 2001). Dieser Befund deckt sich mit dem *Confirmation Bias*, der aus der Literatur zum Inquiry Learning bekannt ist: Lernende ignorieren, interpretieren um oder verwerfen Daten, die zum Zurückweisen der aufgestellten Hypothese führen, und bevorzugen dagegen solche Ergebnisse, die ihre Hypothese bestätigen (Klahr & Dunbar, 1988; de Jong & van Joolingen, 1998).

Die dargestellten Befunde legen nahe, das **biologische Vorwissen** der Lernenden bei der Nutzung von Diagrammen als erklärenden Faktor zu berücksichtigen. Ein korrektes oder inkorrektes Ablesen von bestimmten Inhalten beispielsweise kann unter Umständen gar nicht auf höhere oder geringere Fähigkeiten zum Umgang mit Diagrammen zurückzuführen sein, sondern auf ein höheres oder niedrigeres inhaltliches Vorwissen sowie darauf, dass sich die Vorstellungen und Erwartungen von Lernenden mehr oder minder mit den im Diagramm dargestellten Informationen decken.

3.7.2.2 Umgang mit abstrakten Diagrammen

Im Zusammenhang mit Fähigkeiten beim Umgang mit Diagrammen im konkreten Anwendungssituationen wie im Biologieunterricht (vgl. 3.2.3) wird an manchen Stellen auf die Fähigkeiten zum Umgang mit abstrakten Diagrammen hingewiesen. Als abstrakt werden in diesem Zusammenhang Diagramme bezeichnet, die keine konkreten Variablen darstellen, sondern beispielsweise die Bezeichnungen X und Y als Variablennamen verwenden.

In den Studien von Curcio (1987) zeigten sich die Fähigkeiten zum Umgang mit solchen abstrakten Diagrammen lediglich für die Viertklässler als Prädiktor für die Fähigkeiten zur Informationsentnahme (vgl. 3.7.1.1). Dennoch ist es möglich, dass die Fähigkeiten zum Umgang mit abstrakten Diagrammen auf für die in der vorliegenden Arbeit untersuchte Zielgruppe der Neunt- und Zehntklässler im Zusammenhang mit der Diagrammkompetenz in konkreten Anwendungsfällen stehen.

3.7.2.3 Lesekompetenz

Wie in den grundlegenden Überlegungen zu den Anforderungen, die sich beim Einsatz von Diagrammen ergeben, dargelegt wurde, wird Diagrammkompetenz in der vorliegenden Arbeit bezugnehmend auf bestimmte Textverarbeitungsfähigkeiten verstanden. Diese beziehen sich auf die Achsenbeschriftungen und Beschriftungen der Datenreihen, jedoch insbesondere auf den beim Konstruieren und Integrieren gegebenen Text. Im letzteren Fall müssen die hinsichtlich einer gegebenen Aufgabenstellung relevanten Informationen im Text aufgefunden werden. Es kann somit angenommen werden, dass ein gewisses Maß an Lesefähigkeiten zur erfolgreichen Bearbeitung der entsprechenden Aufgaben zur Diagrammkompetenz beiträgt. Diese Annahme ist empirisch gestützt durch Befunde von Curcio (1987). In ihren Studien wurden Lesefähigkeiten als bedeutender Vorhersagewert für das Verstehen von Diagrammen identifiziert.

3.7.2.4 Rechenfähigkeiten

Wie bereits bei den Ausführungen zum Ablesen angeführt, erfordern es Aufgabenstellungen zum Ablesen zweiter Ordnung, beispielsweise beim quantitativen Vergleich von Werten, eine Differenz zu berechnen. Entsprechend werden Schwierigkeiten bei einem solchen Ablesen häufig auf mangelnde Rechenfähigkeiten zurück geführt (Pereira-Mendoza & Mellor, 1991). Auch in größeren Schulleistungsstudien wurden Belege dafür gefunden, dass die Anforderung, bestimmte Berechnungen bei der Informationsentnahme durchzuführen, zur Schwierigkeit der entsprechenden Testaufgaben beiträgt (Åberg-Bengtsson, 1999; Prenzel et al., 2002). Inwieweit beispielsweise für das Festlegen einer den Daten angemessenen Skala beim Konstruieren von Diagrammen Rechenfähigkeiten erforderlich sind, ist offen. Somit ist es zu klären, ob auch dieser Faktor bei der Erfassung der Diagrammkompetenz berücksichtigt werden muss.

3.7.2.5 Anstrengung

Bei der Bearbeitung von Aufgaben zur Kompetenzmessung kann mangelnde Motivation oder auch Anstrengungsbereitschaft zu schlechteren Testergebnissen führen, als der jeweiligen Kompetenz der Schüler entsprechen. So berichten beispielsweise die Versuchspersonen in den Studien von Adams und Shrum (1990) von einer zunehmenden Ermüdung bei der Testbearbeitung, die in geringerer Anstrengungsbereitschaft beim Zeichnen von Diagrammen resultierte. Die entsprechenden Aufgaben wurden nur noch oberflächlich bearbeitet oder ausgelassen. Da somit die von den Schülern aufgebrauchte Anstrengung bei der Testbearbeitung die Testergebnisse beeinflussen kann, liegt es nahe, sie bei der Untersuchung der Diagrammkompetenz als erklärenden Faktor zu berücksichtigen.

4 Forschungsfragen

In den einleitenden fachdidaktischen Überlegungen zum Einsatz von Diagrammen im naturwissenschaftlichen Unterricht wurde die Vermittlung von Diagrammkompetenz als direktes Ziel dieses Unterrichts beschrieben (vgl. Kapitel 2). Als Grundlage dieser Vermittlung wird ein fundiertes Modell der Diagrammkompetenz benötigt, das die relevanten Kompetenzstrukturen sowie deren Zusammenhänge darlegt. Zum einen gibt ein solches Modell Aufschluss darüber, über welche Fähigkeiten Schüler verfügen müssen, um die Anforderungen, die der Einsatz von Diagrammen im naturwissenschaftlichen Unterricht mit sich bringt, bewältigen zu können. Zum anderen können anhand der Zusammenhänge der einzelnen Strukturen Hinweise auf mögliche Ansätze zur Vermittlung und Förderung der Kompetenz gewonnen werden.

Zielsetzung der vorliegenden Forschungsarbeit ist die Entwicklung eines solchen Strukturmodells für die Diagrammkompetenz sowie dessen empirische Validierung. Dabei werden insbesondere mögliche Einflussfaktoren untersucht, die bei der Diagnose sowie der Förderung von Diagrammkompetenz berücksichtigt werden sollten. Eine qualitative Analyse der auftretenden Fehler beim Umgang mit Diagrammen dient darüber hinaus der Aufdeckung gezielter Förderungsansätze.

Die Entwicklung des Modells wurde ausführlich im theoretischen Teil der Arbeit dargestellt. Im Folgenden werden die Forschungsfragen zur empirischen Überprüfung des Modells (4.1), zur Analyse der strukturellen Zusammenhänge, die der differenzierten Beschreibung der Modellstruktur dient (4.2), zur Untersuchung der Einflussfaktoren (4.3) sowie zur Analyse der auftretenden Fehler (4.4) aufgeführt.

4.1 Überprüfung der Modellstruktur

Entsprechend der Zielsetzung der vorliegenden Forschungsarbeit beziehen sich die Forschungsfragen zunächst auf die Überprüfung der im Kompetenzmodell formulierten Struktur der Diagrammkompetenz. Die übergeordnete Forschungsfrage lautet: Lassen sich die theoretisch abgeleiteten Strukturen der Diagrammkompetenz empirisch bestätigen?

Bezugnehmend auf die theoretischen Hintergründe, die zur Modellentwicklung herangezogen wurden, kann diese übergeordnete Frage in zwei Teilfragen differenziert werden. Die erste bezieht sich darauf, ob die im Kompetenzstrukturmodell beschriebenen Komponenten der Diagrammkompetenz in den empirischen Daten identifiziert werden können (4.1.1). Die zweite fokussiert, ob in der Modellstruktur eine Differenzierung der Diagrammtypen berücksichtigt werden muss (4.1.2).

4.1.1 Komponenten

Basierend auf dem integrativen Modell der Text- und Diagrammverarbeitung (Schnotz, 2001) wurden die Schülerkompetenzen zum Umgang mit Diagrammen in drei Bereiche eingeteilt: Informationsentnahme, Konstruktion und Integration. Diese sind in Abschnitt 3.5.2 näher beleuchtet und als Komponenten der Diagrammkompetenz in das in Abschnitt 3.6 vorgeschlagene Kompetenzstrukturmodell aufgenommen worden. Informationsentnahme bezieht sich auf die Fähigkeiten von Schülern, aus einem gegebenen Diagramm Informationen verschiedener Komplexität entnehmen zu können (vgl. 3.6.1). Konstruktion beschreibt die Fähigkeiten, anhand gegebener Textinformationen ein Diagramm zeichnen zu können (vgl. 3.6.2). Unter Integration werden die Fähigkeiten verstanden, bei vorgegebenem Diagramm und zugehörigem Text die Informationen aus beiden Quellen zu entnehmen und aufeinander zu beziehen (vgl. 3.6.3). Die **Hypothese 1a** lautet demnach:

Als Strukturen der Diagrammkompetenz lassen sich die drei Komponenten Informationsentnahme, Konstruktion und Integration unterscheiden.

Wie in Abschnitt 3.6.1 dargelegt wurde, werden im Kompetenzstrukturmodell innerhalb der Informationsentnahme die Unterkomponenten Identifizierung und Ablesen voneinander abgegrenzt. Hinsichtlich der Konstruktion wurden verschiedene Aspekte formuliert, die im Kompetenzstrukturmodell in den zwei Unterkomponenten *Aufbau des Rahmens* und *Dateneintrag* zusammengefasst werden (vgl. 3.6.2). Im Hinblick auf diese Unterkomponenten lautet die **Hypothese 1b**:

Als Strukturen der Diagrammkompetenz lassen sich die Komponenten Identifizierung, Ablesen, Aufbau, Dateneintrag und Integration identifizieren.

Zu beachten ist an dieser Stelle, dass die beiden genannten Hypothesen sich nicht gegenseitig ausschließen, da die in der Hypothese 1b postulierte Struktur der Diagrammkompetenz nicht im Widerspruch zur Struktur nach Hypothese 1a steht, sondern diese verfeinert.

4.1.2 Diagrammtypen

Das theoretisch abgeleitete Strukturmodell der Diagrammkompetenz bezieht sich sowohl auf Linien- als auch auf Säulendiagramme. Dabei stellt sich die Frage, ob diese übergreifende Struktur beibehalten werden kann oder ob eine Differenzierung der Diagrammtypen für das Modell nötig ist. In Abschnitt 3.5.3 wurde aufgezeigt, dass hinsichtlich der Schnelligkeit bestimmter Ableseprozesse kaum Unterschiede zwischen Linien- und Säulendiagramm bestehen. Hinsichtlich der Informationsentnahme kann somit

angenommen werden, dass die Fähigkeit zum Ablesen sich auf beide Diagrammtypen erstreckt. Aufgrund der strukturellen Ähnlichkeit von Linien- und Säulendiagrammen wird vermutet, dass sich die Fähigkeiten zu den Konstruktionsaspekten und zur Integration ebenfalls übertragen lassen. Im Hinblick auf die Kompetenzmodellierung wird damit insgesamt angenommen, dass die Differenzierung nach Diagrammtypen für das beschriebene Kompetenzstrukturmodell die empirischen Daten nicht besser erklären kann als das Modell, das beide Diagrammtypen umfasst. Entsprechend lautet die **Hypothese 2**:

Das Modell der Diagrammkompetenz ist für beide Diagrammtypen, Linien- und Säulendiagramm, gleichermaßen geeignet.

4.2 Differenzierte Beschreibung der Modellstruktur

An die Überprüfung der strukturellen Komponenten der Diagrammkompetenz schließen sich verschiedene Forschungsfragen an, die auf die differenzierte Beschreibung der Modellstruktur abzielen. Hinsichtlich dieser Fragestellungen wurden keine Hypothesen formuliert, da die Forschungslage keine eindeutigen Schlüsse zulässt. Sie haben folglich eher explorativen Charakter.

4.2.1 Relationen der Komponenten

Zur Beschreibung eines Kompetenzstrukturmodells ist nicht nur die Identifizierung der relevanten Strukturen wichtig, sondern darüber hinaus deren Zusammenhänge. In Abschnitt 3.6.2 wurde herausgestellt, dass bisher keine Studien zur systematischen Analyse des Zusammenhangs zwischen Informationsentnahme und Konstruktion bekannt sind. Es stellt sich damit die Frage danach, wie sich der genannte Zusammenhang für das in der vorliegenden Arbeit beschriebene Kompetenzstrukturmodell darstellt. Insbesondere ergibt sich daraus die Frage, welche Zusammenhänge sich jeweils zwischen der Integration und der Informationsentnahme respektive der Konstruktion zeigen.

Bei der Entwicklung des Kompetenzmodells wurde zudem erklärt, dass die Unterkomponenten Identifizierung (auf der Seite der Informationsentnahme) und Aufbau (auf der Seite der Konstruktion) sowie Ablesen und Dateneintrag insofern ähnliche Strukturen aufweisen, als dass sich erstere Unterkomponenten auf den Rahmen eines Diagramms, letztere Unterkomponenten auf die im Diagramm eingetragenen oder einzutragenden Werte beziehen. In Bezug zum Verhältnis von Informationsentnahme und Konstruktion ist zu prüfen, ob sich diese Ähnlichkeiten in starken Zusammenhängen zwischen diesen jeweiligen Unterkomponenten widerspiegeln.

Die zusammenfassende Forschungsfrage lautet:

Welche Zusammenhänge bestehen zwischen den einzelnen Komponenten der Diagrammkompetenz respektive ihren Unterkomponenten?

4.2.2 Schwierigkeiten der Komponenten

In direktem Bezug zur obigen Forschungsfrage steht die Frage danach, in welchem Verhältnis die einzelnen Komponenten sowie deren Unterkomponenten hinsichtlich ihrer Schwierigkeit zueinander stehen. Als weitere Forschungsfrage wird somit formuliert:

Unterscheiden sich die Komponenten und Unterkomponenten der Diagrammkompetenz in ihrer Schwierigkeit?

4.2.3 Binnenstrukturen der einzelnen Komponenten

Wie in Abschnitt 3.6 aufgeführt, sind die Komponenten Informationsentnahme und Konstruktion durch feinere Unterteilungen in Unterkomponenten und Aspekte gegliedert. Das genaue Verhältnis dieser Aspekte zueinander ist allerdings offen. Hinsichtlich des Ablesens stellt sich die Frage, ob sich die Ableseordnungen als Kompetenzstufen identifizieren lassen. Da keine umfassenden Untersuchungen zur detaillierten Erfassung der Konstruktionsleistungen bekannt sind, ergibt sich hinsichtlich der Konstruktion die Frage danach, wie die einzelnen Aspekte der Konstruktion miteinander zusammenhängen. In Hinblick auf die Integration ist offen, ob die verschiedenen resultierenden Repräsentationen (*Resultat Text* und *Resultat Diagramm*) sich in zu differenzierenden Kompetenzstrukturen widerspiegeln. Die umfassende Forschungsfrage lautet diesbezüglich:

Welche Binnenstrukturen können für die einzelnen Komponenten der Diagrammkompetenz identifiziert werden?

4.2.4 Vergleich der Schwierigkeiten bezogen auf die Diagrammtypen

In Abschnitt 3.5.3 wurden einzelne Befunde zur Schnelligkeit von Ableseprozessen aus Diagrammen dargestellt. Nach diesen unterscheiden sich Linien- und Säulendiagramm nicht beachtenswert. Da keine vergleichbaren empirischen Befunde hinsichtlich der Konstruktion und Integration bekannt sind, ergibt sich die Forschungsfrage:

Zeigen sich in Abhängigkeit vom Diagrammtyp Unterschiede in den Schwierigkeiten der einzelnen Komponenten?

4.3 Lernerhintergrund und Lernermerkmale

Verschiedene mögliche Einflussvariablen auf die Diagrammkompetenz von Schülern wurden in Abschnitt 3.7 erläutert. Da sich die herangezogenen Befunde meist auf einzelne Komponenten beziehen, können aus dieser Literatur keine umfassenden Hypothesen abgeleitet werden. Auch die folgenden Forschungsfragen werden daher explorativ untersucht.

4.3.1 Einfluss des Lernerhintergrunds

Hinsichtlich des Lernerhintergrundes wurden in Abschnitt 3.7.1 Hinweise darauf gegeben, dass sich Neunt- und Zehntklässler respektive Schülerinnen und Schüler in ihren Fähigkeiten in Bezug auf die Informationsentnahme unterscheiden. Offen ist, ob dies auch in Bezug auf die anderen Komponenten Konstruktion und Integration sowie die einzelnen Unterkomponenten der Informationsentnahme und Konstruktion gilt. Es ergibt sich die Forschungsfrage:

a) Zeigen sich hinsichtlich der identifizierten Komponenten Fähigkeitsunterschiede zwischen Neunt- und Zehntklässlern sowie zwischen Schülerinnen und Schülern?

Zur differenzierten Beschreibung der Kompetenzstruktur werden in der vorliegenden Arbeit die Zusammenhänge zwischen den einzelnen Komponenten respektive Unterkomponenten untersucht (vgl. 4.2.1). Dabei ist es relevant zu prüfen, ob diese Zusammenhänge auch für beide Klassenstufen sowie für beide Geschlechter gelten. Unterschiede in den Relationen der einzelnen Komponenten zueinander würden nahe legen, bei der Vermittlung der Kompetenz unterschiedliche Maßnahmen für die einzelnen Gruppen zu entwickeln. Als weitere explorative Forschungsfrage wird somit formuliert:

b) Zeigen sich für Neunt- und Zehntklässler respektive Schülerinnen und Schüler Unterschiede hinsichtlich der Relationen der Komponenten zueinander?

4.3.2 Einfluss von Lernermerkmalen

Die Befunde über einen Einfluss der in Abschnitt 3.7.2 dargestellten Lernermerkmale sind nicht einheitlich. So wird beispielsweise ein gewisses Fachwissen der Lernenden als notwendige Voraussetzung zum Verstehen von Diagrammen beschrieben (z. B. Preece & Janvier, 1992). Gleichzeitig kann das inhaltliche Vorwissen die Interpretation in der Hinsicht beeinflussen, dass Informationen entnommen werden, die mit den Erwartungen der Leser übereinstimmen, aber nicht im Diagramm dargestellt sind (vgl. Shah, 2001). Darüber hinaus beziehen sich die aufgeführten empirischen Befunde aller möglichen Lernermerkmale meist nur auf einzelne der im Strukturmodell der

Diagrammkompetenz formulierten Komponenten, so dass lediglich spekuliert werden kann, ob ein Einflussfaktor auf *eine* Komponente der Diagrammkompetenz auch Einfluss auf die anderen Komponenten hat. Als explorative Forschungsfrage ergibt sich somit:

Zeigen sich Zusammenhänge zwischen den genannten Lernermerkmalen (biologisches Vorwissen, Fähigkeiten zum Umgang mit abstrakten Diagrammen, Lesekompetenz, Rechenfähigkeiten und aufgebrachte Anstrengung) und den Fähigkeiten der Schüler in den einzelnen (Unter-)Komponenten der Diagrammkompetenz?

4.4 Qualitative Fehleranalyse

Zur Beschreibung der einzelnen Komponenten der Diagrammkompetenz wurden detailliert empirische Befunde zu den jeweils auftretenden Schwierigkeiten und Fehlern aufgeführt (vgl. 3.5.2). Eine Analyse der Schwierigkeiten von Schülern stellt daneben die Grundlage für gezielte Förderungsansätze dar. Nach Shah und Hoeffner (2002) fehlt es insbesondere hinsichtlich der Konstruktion von Diagrammen an solchen umfassenden Analysen. So sind zu einigen Aspekten, beispielsweise der Achsenbeschriftung, bisher keine empirischen Beschreibungen von Schülerfehlern bekannt. Es stellt sich die Frage, welche der bereits bekannten Fehler in der vorliegenden Untersuchung ebenfalls eine Rolle spielen, und welche Fehler darüber hinaus festgestellt werden können. Als übergreifende Forschungsfrage wird deshalb formuliert:

Welche Fehler treten beim Konstruieren von Diagrammen auf?

5 Methoden

Im diesem Kapitel wird im ersten Abschnitt der testtheoretische Hintergrund der vorliegenden Arbeit dargelegt (5.1). Im zweiten Abschnitt (5.2) werden die eingesetzten Erhebungsverfahren und im dritten Abschnitt (5.3) die Auswertungsmethoden vorgestellt.

5.1 Testtheoretischer Hintergrund

In der vorliegenden Arbeit wurden Verfahren der probabilistischen Testtheorie verwendet. Die Gründe für die Auswahl werden nachfolgend dargelegt, indem einleitend die probabilistische Testtheorie kurz von der klassischen Testtheorie abgegrenzt wird. Im Anschluss daran werden die Grundsätze der probabilistischen Testtheorie vorgestellt (5.1.1) und das eindimensionale Rasch-Modell als das grundlegende Modell der probabilistischen Testmodelle erläutert (5.1.2.1). Die in der vorliegenden Arbeit zum Einsatz gekommenen Erweiterungen des Rasch-Modells werden in den darauf folgenden Abschnitten (5.1.2.2 bis 5.1.2.4) nach ihren Anwendungen charakterisiert. Im Anschluss daran wird kurz auf die in der vorliegend beschriebenen Studie genutzte Parameterschätzung, die auf den probabilistischen Analysen basiert, eingegangen (5.1.2.5).

Anhand eines Tests sollen bestimmte Eigenschaften von Personen gemessen werden. Dies können beispielsweise Einstellungen und Überzeugungen zu bestimmten Themen oder, wie im vorliegenden Fall, bestimmte Fähigkeiten sein. Die klassische Testtheorie geht davon aus, dass sich die Messwerte einer Person, d. h. die manifesten Antworten, die sie in einem Test gibt, aus einem „wahren Wert“ für ihre Fähigkeit und einem Messfehler zusammensetzen. Die richtigen Antworten bzw. die Zahl der richtigen Antworten (Summenscore), werden bzw. wird damit als direktes Maß der Fähigkeit einer Person angesehen. Im Gegensatz dazu wird in der probabilistischen Testtheorie angenommen, dass die beobachtbaren Antworten einer Person aus einer latenten Fähigkeit resultieren, aber nicht *direktes* Maß für diese Fähigkeit sind. Die Fähigkeit ist demnach nicht direkt beobachtbar, sondern muss durch Analysen erst geschätzt werden (Rost, 1996, S. 10).

Bei dieser Schätzung werden in der probabilistischen Testtheorie *Personenparameter*, die das Ausmaß der Fähigkeit einer Person bezeichnen, und *Itemparameter* unterschieden. Letztere geben Aufschluss über die Schwierigkeit eines Testitems. Genau in dieser getrennten Schätzung von Personenfähigkeiten und Itemschwierigkeiten liegt das große Potential der probabilistischen Testmodelle zur systematischen Analyse der Fähigkeitsstrukturen. Während klassische Teststatistiken (beispielsweise die Itemschwierigkeit als Prozentsatz der richtig antwortenden Probanden) von der Stichprobe abhängen (Hambleton & Jones, 1993), lässt die probabilistische Modellierung Aussagen über das Messinstrument zu, die – bei Geltung des Modells – unabhängig von der

untersuchten Stichprobe sind (Baker, 2001, S. 51, 45f.) und damit hervorragend zur Testkonstruktion herangezogen werden können. Diese Eigenschaft probabilistischer Testmodelle wird als *Stichprobeninvarianz* (Rost, 1996; Baker, 2001) bezeichnet. Sie bietet insbesondere die Grundlage dafür, Testaufgaben definierten Fähigkeitsstufen zuweisen zu können (Hambleton & Jones, 1993) und damit valide Instrumente zur Beurteilung von Schülerfähigkeiten zu entwickeln. Die Konstruktion solcher Diagnoseinstrumente gewinnt insbesondere vor dem Hintergrund der nationalen Bildungsstandards (z. B. KMK, 2005a) und der Überprüfung deren Erreichens an Bedeutung.

Im folgenden Abschnitt werden die Grundlagen der probabilistischen Testtheorie zunächst vertieft, bevor die Modelle zur Berechnung der Personen- und Itemparameter vorgestellt werden.

5.1.1 Grundlagen der probabilistischen Testtheorie

5.1.1.1 Itemcharakteristik-Kurve (ICC)

Obwohl Fähigkeiten der Personen nicht direkt messbar sind, wird in der probabilistischen Testtheorie davon ausgegangen, dass es eine Skala gibt, auf der sie sich anordnen lassen. Diese Skala hat einen Nullpunkt und eine theoretische Reichweite von negativ unendlich bis positiv unendlich, aus praktischen Überlegungen werden allerdings die Werte üblicherweise in einem Intervall von -3 bis $+3$ angegeben (Baker, 2001, S. 5f).

Die Fähigkeit einer Person v wird im Allgemeinen mit dem griechischen Buchstaben Theta Θ_v bezeichnet. Von einem probabilistischen Standpunkt aus ist es von Interesse zu wissen, mit welcher Wahrscheinlichkeit eine Person v mit einer bestimmten Fähigkeit Θ_v ein Testitem i auf eine bestimmte Art x beantwortet. Diese Wahrscheinlichkeit wird vollständig mit $P(x_{vi}|\Theta_v)$ und abkürzend mit $P(x_{vi})$ bezeichnet. In jedem Testmodell, das diese Wahrscheinlichkeiten modelliert, werden dabei bestimmte Annahmen über den Zusammenhang zwischen der Personenfähigkeit und der Wahrscheinlichkeit einer bestimmten Itemantwort getroffen. Für dichotome Items, das heißt für solche mit genau zwei Antwortkategorien 0 und 1, lässt sich dieser Zusammenhang in Form einer *Itemcharakteristik-Kurve (item characteristic curve, ICC)* wie in der Abbildung 5.1 darstellen. Dabei wird $P(x_{vi}=1)$, die Wahrscheinlichkeit der 1-Antwort, die traditionellerweise als die „korrekte“ Antwort interpretiert wird, in Abhängigkeit von der Personenfähigkeit Θ_v aufgetragen.

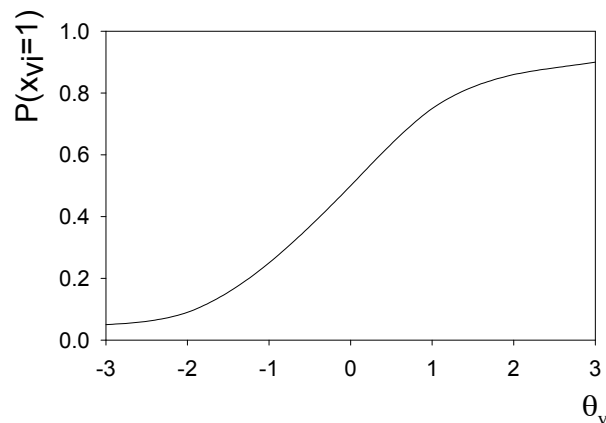


Abbildung 5.1: Itemcharakteristik-Kurve (item characteristic curve, ICC) für ein Item i ; monoton steigend.

Probabilistische Testmodelle nehmen einen kontinuierlichen Verlauf dieser ICC an⁴⁰. Sollen *Fähigkeiten* modelliert werden, die zum Lösen von Aufgaben eines bestimmten Typs benötigt werden, werden im Allgemeinen Modelle gewählt, die eine *monoton steigende Funktion* annehmen, da die Wahrscheinlichkeit, das Item korrekt zu beantworten, mit steigender Fähigkeit steigen und nicht etwa ab einem bestimmten Punkt wieder abnehmen sollte⁴¹. Meist wird dabei ein sigmoider Verlauf der Itemcharakteristik wie in der Abbildung 5.1 angenommen, der die Annahme widerspiegelt, dass die Wahrscheinlichkeit, ein Item richtig zu beantworten, bei einer geringen Personenfähigkeit nahe 0 liegt, und dass mit steigender Fähigkeit die Wahrscheinlichkeit einer richtigen Antwort erst langsam, dann stärker ansteigt, bevor sie sich bei den höchsten Fähigkeitswerten dem Wert 1 annähert.

Abgesehen vom prinzipiellen sigmoiden Verlauf ist die genaue Form der Itemcharakteristik für jedes Item spezifisch. Die Unterschiede der ICCs verschiedener Items manifestieren sich dabei in zwei Kennwerten: ihrer Lage auf der horizontalen Fähigkeitsskala und ihrer Steigung. Durch diese zwei charakteristischen Werte der ICC können gleichfalls zwei entscheidende technische Eigenschaften des zugehörigen Items, seine Schwierigkeit und seine Trennschärfe, beschrieben werden (Baker, 2001, S. 7).

5.1.1.2 Itemschwierigkeit

Abbildung 5.2 illustriert drei ICCs gleicher Steigung, die sich in ihrer Lage auf der Skala der Personenfähigkeit unterscheiden. Die Lage auf der Fähigkeitsskala ist dabei ein Maß für die Schwierigkeit des jeweils zugehörigen Items. Je weiter rechts die Kurve liegt, desto schwieriger das Item. So repräsentiert die Kurve A ein leichtes Item, da insbesondere

⁴⁰ Ein Beispiel für einen nicht-kontinuierlichen Verlauf wäre die Guttman-Charakteristik, vgl. dazu Rost, 1996, S. 104ff.

⁴¹ Prinzipiell wäre beispielsweise auch eine nicht-monotone, eingipflige Funktion denkbar, die z. B. die Zustimmung zu einer Aussage in einem mittleren Bereich einer Einstellungsskala modelliert.

Personen mit geringen Fähigkeitswerten (beispielsweise -1) die richtige Lösung des Items mit recht hohen Wahrscheinlichkeiten (hier etwa $0,6$) gelingt. Die Kurve C hingegen gehört zu einem schwierigeren Item: erst ab einem Fähigkeitswert von 2 liegt die Wahrscheinlichkeit einer richtigen Antwort über $0,6$.

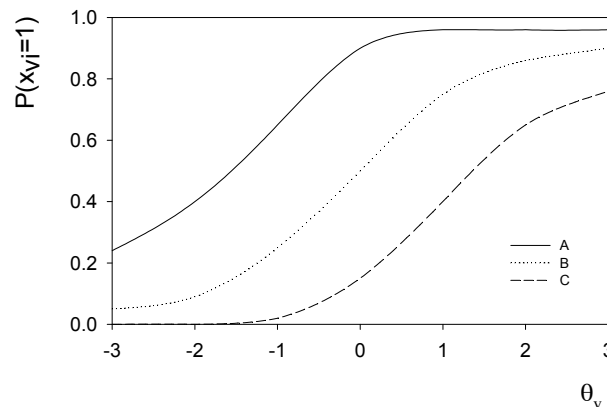


Abbildung 5.2: ICCs dreier Items mit unterschiedlicher Schwierigkeit. Je weiter rechts, desto schwieriger das Item.

Per Konvention wird die **Itemschwierigkeit** als derjenige Fähigkeitswert definiert, der einer Wahrscheinlichkeit einer korrekten Lösung von $0,5$ entspricht (Rost, 1996, S. 102). Für Kurve B in der Abbildung 5.2 ist dies genau 0 , für Kurve C etwa $1,3$.

Um Verwechslungen vorzubeugen, die aufgrund der Verwendung des Begriffs „Itemschwierigkeit“ in der klassischen Testtheorie entstehen könnten, sei an dieser Stelle auf Folgendes hingewiesen: Im Sinne der probabilistischen Testtheorie bedeutet eine Itemschwierigkeit von 2 nicht, dass Personen, deren Fähigkeitswert über 2 liegt, das Item korrekt lösen, und Personen, deren Fähigkeitswert unter 2 liegt, das Item falsch beantworten. Sie besagt, dass Personen mit einem Fähigkeitswert über 2 das Item mit einer Wahrscheinlichkeit von über 50% und Personen mit einem Fähigkeitswert unter 2 das Item mit einer Wahrscheinlichkeit von unter 50% korrekt lösen.

5.1.1.3 Trennschärfe

Die Trennschärfe eines Items ist ein Maß dafür, wie gut sich Personen unterschiedlicher Fähigkeiten unterscheiden lassen (Rost, 1996, S. 102) d. h. ob sich bei unterschiedlichen Fähigkeiten auch ihre Wahrscheinlichkeiten, das Item richtig zu beantworten, unterscheiden. In der klassischen Testtheorie wird die Trennschärfe als (punkt-biserielle) Korrelation zwischen der Beantwortung des Items und dem Gesamtestscore berechnet

(vgl. Fisseni, 1997, S. 48). In probabilistischen Testmodellen wird die Steigung⁴² der ICC eines Testitems als Maß für die Trennschärfe genutzt: je steiler die ICC, desto trennschärfer das Item. In der Abbildung 5.3 werden drei Itemfunktionen gleicher Itemschwierigkeit, jedoch mit unterschiedlichen Trennschärfen dargestellt.

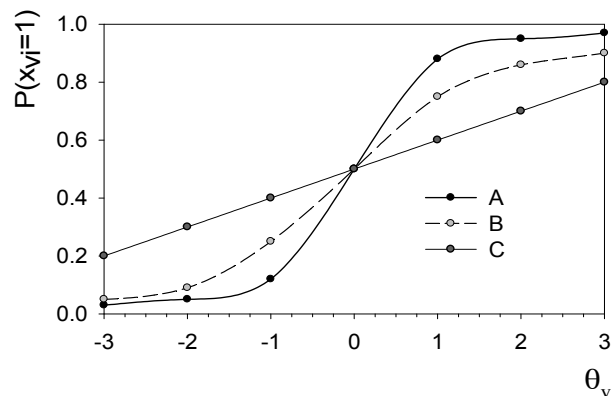


Abbildung 5.3: ICCs dreier Items unterschiedlicher Trennschärfe. Je steiler die Kurve, desto trennschärfer das Item.

5.1.2 Probabilistische Testmodelle

Im Folgenden wird das Rasch-Modell als grundlegendes Modell der probabilistischen Testtheorie erläutert. Wenngleich Erweiterungen dieses Modells bei der Auswertung zum Einsatz gekommen sind, soll an dieser Stelle darauf verzichtet werden, die jeweiligen Modellgleichungen aufzuführen und zu erklären. Sie bedürfen einer komplexen mathematischen Herleitung, die der einschlägigen Literatur entnommen werden kann (z. B. Rost, 1996; Adams, Wilson & Wang, 1997). Stattdessen werden die verwendeten Modelle basierend auf der ausführlichen Beschreibung der Modellparameter und deren Interpretation kurz charakterisiert und so ihre Anwendungen in der vorliegenden Arbeit dargelegt.

5.1.2.1 Das grundlegende Modell: Das eindimensionale Rasch-Modell für dichotome Items

Durch die ICC eines Items i wird der Zusammenhang zwischen der Fähigkeit Θ_v einer Person v und der Wahrscheinlichkeit $P(x_{vi}=1)$, dass sie Item i richtig löst, dargestellt. Eine Möglichkeit, eine monoton steigende ICC mit dem gewünschten sigmoiden Verlauf mathematisch zu beschreiben, liefert die logistische Funktion. Der Däne Georg Rasch beschrieb im Jahre 1960 als erster ein probabilistisches Testmodell, das nach ihm benannte

⁴² Genauer müsste es die Steigung im Bereich des steilsten Anstiegs der Kurve, d. h. der Tangente im Wendepunkt, heißen.

Rasch-Modell, bei dem die logistische Funktion zur Beschreibung der ICC verwendet wurde. Die Modellgleichung des Rasch-Modells lautet:

$$P(x_{vi}) = \frac{e^{(x_{vi}(\Theta_v - \sigma_i))}}{1 + e^{(\Theta_v - \sigma_i)}},$$

wobei v für eine Person, i für ein Item steht und x_{vi} die Antwort der Person v bezüglich Item i bezeichnet. Θ_v ist die Fähigkeit der Person v und σ_i die Schwierigkeit des Items i .

Die Modellgleichung beschreibt die Wahrscheinlichkeit einer bestimmten Antwort in Abhängigkeit von ausgewählten *Modellparametern*. Im Fall des Rasch-Modells ist auffällig, dass in die Berechnung der Lösungswahrscheinlichkeit ausschließlich die Parameter für die *Personenfähigkeit* Θ_v und die *Itemschwierigkeit* σ_i einfließen, nicht aber der für die Trennschärfe eines Items. Tatsächlich wird im Rasch-Modell angenommen, dass alle Items die gleiche Trennschärfe haben (Rost, 1996, S. 125). Unter dieser Annahme unterscheiden sich die Itemfunktionen der Items nur durch ihre Lage auf der Fähigkeitsskala und sind damit alle parallel (vgl. Abbildung 5.2). Daneben leitet sich aus dieser Annahme paralleler Itemfunktionen eine grundlegende Eigenschaft des Rasch-Modells ab: Die Wahrscheinlichkeit einer bestimmten Antwort kann allein durch die Differenz aus Fähigkeitsparameter Θ_v einer Person v und Schwierigkeit σ_i eines Items i bestimmt werden. So löst beispielsweise eine Person mit einem Fähigkeitswert von 1 ein Item mit Schwierigkeit 2 zu 27% richtig, eine Person mit einem Fähigkeitswert von -1 ein Item mit Schwierigkeit von 0 ebenfalls zu 27%. In beiden Fällen ist die Lösungswahrscheinlichkeit gleich, da auch die Differenz aus Personenfähigkeit und Itemschwierigkeit gleich ist. Insbesondere gilt die Gleichung

$$\log \frac{P(x_{vi} = 1)}{P(x_{vi} = 0)} = \Theta_v - \sigma_i,$$

die durch direkte Umformung der Gleichung des Rasch-Modells hergeleitet werden kann. Gleichzeitig wird durch den linken Teil der Gleichung die Einheit erklärt, in der im Rasch-Modell Personenparameter und Itemparameter angegeben werden; sie wird als **Logit der Wahrscheinlichkeit** bezeichnet (Rost, 1996, S. 122ff). Die Personen- und Itemparameter werden im Rasch-Modell auf *einer* Skala (oder Dimension), der **Logit-Skala**, dargestellt und können somit direkt miteinander verglichen werden (Hambleton & Jones, 1993). Diese Eigenschaft des Rasch-Modells wird auch als *latente Additivität* bezeichnet (Rost, 1996; Walter, 2005).

Das Rasch-Modell ist sehr gut geeignet, um die grundlegenden Begriffe der probabilistischen Testtheorie (ICC, Personenfähigkeit, Itemschwierigkeit) zu illustrieren. Seine Anwendbarkeit ist jedoch beispielsweise auf dichotome Items oder auf Sammlungen von Items begrenzt, die alle die gleiche latente Personenfähigkeit abprüfen sollen. Den Restriktionen des Rasch-Modells wird mit vielfältigen Weiterentwicklungen bzw. Verallgemeinerungen begegnet, die es ermöglichen, die jeweiligen empirischen Bedingungen zu berücksichtigen. Da diese Modelle die beschriebenen Eigenschaften des Rasch-Modells erhalten, werden sie oft unter dem Namen **Rasch-Familie** zusammengefasst (Walter, 2005). Die Modellierungen, die mit ihnen durchgeführt werden, werden dementsprechend als **Rasch-Modellierungen** bezeichnet. Nachfolgend werden die für die vorliegende Arbeit relevanten Modelle der Rasch-Familie aufgeführt und danach charakterisiert, welche Erweiterungen des Einsatzbereichs sie mit sich bringen.

5.1.2.2 Das Partial-Credit-Modell

Das **Partial-Credit-Modell** nach Masters (1982) ermöglicht die Analyse von Items mit mehr als zwei (ordinal geordneten) Antwortkategorien. In der vorliegenden Arbeit sind dies offene Aufgaben, deren Antworten nicht nur nach richtig und falsch bewertet werden (dichotome Kodierung), sondern z. B. nach vollständig gelöst, teilweise gelöst und nicht gelöst. Die Bezeichnung des Modells geht auf die englische Bezeichnung für „teilweise gelöst“ zurück: So erhält jemand, der eine Aufgabe teilweise gelöst hat, *partial credit*, im Vergleich zum *full credit* bei vollständiger Lösung⁴³.

Bei einer dichotomen Kodierung ist es üblich, die Kategorie der falschen Antwort mit „0“, die der richtigen mit „1“ zu benennen. Traditionellerweise werden auch bei einer ordinalen Kodierung die Zahlen, mit denen verschiedene Kategorien beschrieben werden, so verwendet, dass eine numerisch höhere Kategorie eine qualitativ höherwertige Antwort beschreibt. Dementsprechend bezeichnet in der vorliegenden Arbeit die Kategorie 2 bei ordinal skalierten Items immer die Kategorie „vollständig gelöst“, die Kategorie 1 beschreibt die Kategorie „teilweise gelöst“ und die Kategorie 0 steht für „nicht gelöst“.

Beim Vorliegen eines solchen ordinalen Items ist es sinnvoll, den Begriff der *Itemschwierigkeit* anders als für dichotome Items zu definieren. Im dichotomen Fall ist die Itemschwierigkeit nach Definition genau der Fähigkeitswert, an dem die Wahrscheinlichkeit einer richtigen Lösung (in dem Fall der 1-Kategorie) 0,5 beträgt. Im Falle eines ordinalen Items existieren mindestens *zwei* Antwortkategorien, die eine (teilweise) richtige Lösung bezeichnen, weshalb sich die obige Definition der

⁴³ Items mit auf diese Weise geordneten Antwortkategorien werden auch *ordinale Items* genannt und das Partial-Credit-Modell kann entsprechend als das *ordinale Rasch-Modell* bezeichnet werden (Rost, 1996, S. 206). Da es allerdings noch weitere ordinale Rasch-Modelle gibt, so etwa das Ratingskalenmodell, wird in der vorliegenden Arbeit der Terminus Partial-Credit-Modell (PCM) verwendet.

Itemschwierigkeit nicht ohne Weiteres übertragen lässt. Aus diesem Grund werden im PCM nicht Itemschwierigkeitsparameter, sondern **Schwellenparameter** geschätzt. Sie geben die Fähigkeitswerte an, ab denen die Wahrscheinlichkeit einer Antwort in der höheren Kategorie wahrscheinlicher ist als die einer Antwort in der niedrigeren Kategorie.

Am Beispiel eines ordinalen Items mit drei Kategorien 0, 1 und 2 lässt sich das Prinzip der Schwellenparameter gut illustrieren. Die Abbildung 5.4 zeigt die *Kategorienfunktionen* (Rost, 1996, S. 197) für ein solches Item. Eine Kategorienfunktion gibt für *jede Kategorie* x die Wahrscheinlichkeiten, dass diese Kategorie gewählt wird, in Abhängigkeit von der Personenfähigkeit an (die bisher bekannte ICC für dichotome Items mit den Kategorien 0 und 1 ist damit ebenfalls eine Kategorienfunktion, und zwar die der 1-Kategorie).

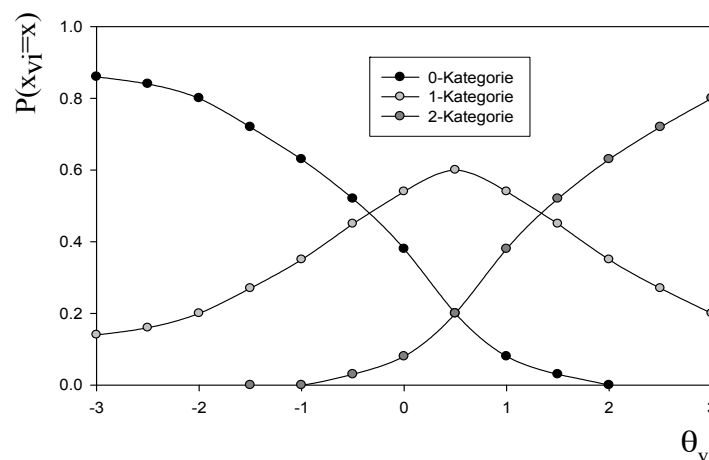


Abbildung 5.4: Kategorienfunktionen eines ordinalen Items mit drei Kategorien.

Diese Kategorienfunktionen spiegeln die Annahme wider, dass in Bereichen niedriger Fähigkeitsausprägungen die Wahrscheinlichkeit für eine falsche Antwort (0-Antwort) am höchsten ist, in Bereichen mittlerer Fähigkeiten die Wahrscheinlichkeit für eine teilweise korrekte Antwort (1-Antwort) überwiegt und in Bereichen hoher Fähigkeitsausprägungen die Wahrscheinlichkeit einer vollständig richtigen Lösung (2-Antwort) dominiert. Während die Wahrscheinlichkeit einer falschen Antwort mit zunehmender Fähigkeit kontinuierlich abnimmt, und die Wahrscheinlichkeit einer vollständig richtigen Antwort immer mehr zunimmt, steigt die Wahrscheinlichkeit für eine teilweise korrekte Lösung zuerst an, nimmt aber mit zunehmender Fähigkeit wieder ab, da dann die Wahrscheinlichkeit der vollständig korrekten Lösung überwiegt.

Im Beispiel von Abbildung 5.4 schneiden sich die Kategorienfunktionen der 0- und 1-Kategorie bei einer Personenfähigkeit von $-0,4$; ab diesem Wert ist die Wahrscheinlichkeit für eine 1-Antwort höher als die für eine 0-Antwort. Ab einer Personenfähigkeit von $1,4$ ist es für die Kategorienfunktionen der 1- und 2-Kategorie

ebenso: vorher ist die Wahrscheinlichkeit für die 1-Kategorie höher als die der 2-Kategorie, hinterher verhält es sich umgekehrt.

Die Schnittstellen der Kategorienfunktionen jeweils *zweier benachbarter Kategorien* werden auch als **Schwellen** (Rost, 1996, S. 199) bezeichnet. Ein dreikategorielles Item besitzt damit zwei Schwellen s_1 und s_2 . Diese geben Aufschluss darüber, wie schwierig eine Kategorie im Vergleich zur nächsthöheren ist: Je höher der Schwellenparameter, desto höher muss die Personenfähigkeit sein, damit die Wahrscheinlichkeit der höheren Kategorie größer ist als die der niedrigeren Kategorie. Verhalten sich die Kategorienfunktionen wie in der Abbildung 5.4 angezeigt, so ist s_1 kleiner als s_2 (vgl. 5.3.2.1, s. Schwellenparameter).

Als Maß für die **Schwierigkeit des gesamten ordinalen Items** wird dann der Mittelwert s_M der Schwellen berechnet. Diese *mittleren Schwellenparameter* stehen darüber in Bezug zu den Itemschwierigkeiten für dichotome Items, dass für letztere die Itemschwierigkeit gerade genau den Schwellenwert zwischen der 0- und der 1-Kategorie (und damit auch den Mittelwert der Schwellenparameter, da nur eine Schwelle existiert) darstellt. Ist im Folgenden somit vom Schwierigkeitsparameter oder von der Itemschwierigkeit die Rede, handelt es sich bei ordinalen Items um die gemittelte Schwelle.

5.1.2.3 Das Ordered-Partition-Modell

In manchen Fällen kann es sinnvoll und wünschenswert sein, in den Daten ebenfalls die Information darüber festzuhalten, auf welche Weise jemand beispielsweise einen Punkt erlangt hat: dann werden feinere Rubriken (etwa 1a und 1b) geschaffen, die beide mit einem Punkt bewertet werden. Somit existieren im vorliegenden Fall drei verschiedene Kategorien von Punktzahlen (0, 1, 2), aber vier verschiedene Antwortrubriken (0, 1a, 1b, 2). Damit ist das Partial-Credit-Modell nicht mehr anwendbar, da dazu jede Rubrik genau einer Punktzahl entsprechen müsste. Abhilfe kann in diesem Fall eine weitere Ergänzung des Rasch-Modells schaffen, das **Ordered-Partition-Modell** nach Wilson (1992). Die Anwendung dieses Modells erlaubt es, mit Items umzugehen, bei denen mehrere verschiedene Rubriken einer Punktzahl zugeordnet werden.

5.1.2.4 Das mehrdimensionale Rasch-Modell

Die beschriebenen Modelle haben die Annahme der Eindimensionalität der Personenfähigkeit gemein. Damit wird angenommen, dass alle eingesetzten Testitems *eine* latente Fähigkeit messen. Insbesondere vor dem Hintergrund der Frage nach Kompetenzstrukturen gewinnt jedoch die Messung mehrerer unterschiedlicher Komponenten einer Kompetenz innerhalb eines Tests an Bedeutung. An dieser Stelle greift

das *multidimensional random coefficients multinomial logit model (MRCMLM)* nach Adams, Wilson und Wang (1997), mit Hilfe dessen Items zu mehreren verschiedenen Fähigkeitskomponenten verarbeitet werden können. Es wird in der vorliegenden Arbeit abkürzend als **mehrdimensionales Rasch-Modell** bezeichnet. Allerdings bezieht sich der Begriff der *Mehrdimensionalität* spezifisch auf die Rasch-Modellierung und darf nicht mit dem Dimensionsbegriff, der häufig für Kompetenzmodelle gebraucht wird, verwechselt werden: Eine Dimension im mehrdimensionalen Rasch-Modell entspricht je nach Modellierung einer Komponente oder Unterkomponente im Kompetenzstrukturmodell (vgl. 3.4.2).

Durch verschiedene Typen des mehrdimensionalen Rasch-Modells, dem *between-item-Modell* und dem *within-item-Modell* (Adams et al., 1997), können ferner unterschiedliche Kompetenzstrukturen modelliert werden. Im *between-item-Modell* (siehe Abbildung 5.5a) werden sämtliche Items a priori zu disjunkten Gruppen zusammengefasst, die jeweils den angenommenen Komponenten entsprechen. Es wird damit modelliert, dass die Lösung eines Items jeweils nur von *einer* der betrachteten Fähigkeiten abhängt. Im *within-item-Modell* (siehe Abbildung 5.5b) kann jedes Item einer oder mehreren Dimensionen zugeordnet werden. Dadurch wird der Fall modelliert, dass zur Lösung des jeweiligen Items eine oder mehrere Fähigkeitskomponenten benötigt werden.

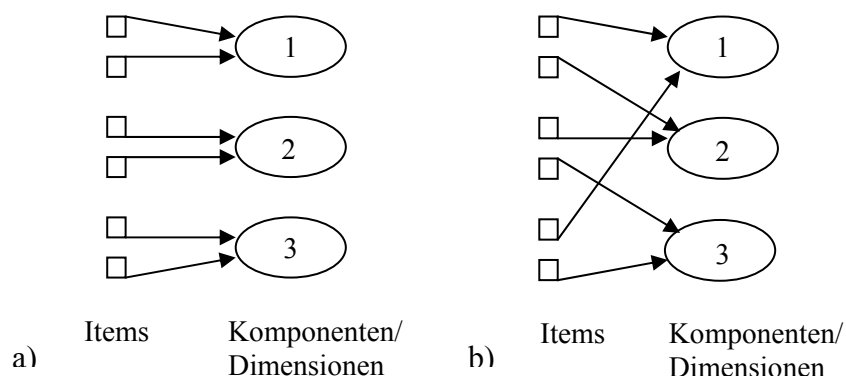


Abbildung 5.5: Schematische Darstellung der zwei mehrdimensionalen Rasch-Modelle (verändert nach Adams et al., 1997). Komponenten/Dimensionen 1 bis 3. a) *between-item-Modell*: jedes Item ist genau einer Fähigkeitskomponente/Modelldimension zugeordnet. b) *within-item-Modell*: jedes Item ist einer oder mehreren Fähigkeitskomponenten/Modelldimensionen zugeordnet.

Das mehrdimensionale Rasch-Modell schließt die Grundgleichungen des oben dargestellten Partial-Credit-Modells (vgl. 5.1.2.2) sowie Ordered-Partition-Modells (vgl. 5.1.2.3) ein. Durch die Implementation dieses Modells in die in der vorliegenden Arbeit verwendete Software ConQuest (Wu, Adams & Wilson, 1998) können somit Partial-Credit-Items und insbesondere Items mit mehreren Antwortrubriken zu einer Punktzahl in einer mehrdimensionalen Rasch-Modellierung verarbeitet werden.

5.1.2.5 Parameterschätzung

Bei der Berechnung probabilistischer Testmodelle werden als Personenparameter nicht die absoluten Testscores verwendet, sondern basierend auf den Antworten der Personen deren Fähigkeitswerte geschätzt (vgl. 5.1), diese Schätzwerte sind quasi die „Messwerte“ der Personen. Dazu stehen verschiedene Schätzer zur Verfügung, wie beispielsweise der *Maximum-Likelihood-Schätzer (MLE)*, der *Bayes-Modal-Estimator (BME)*, der *Weighted-Likelihood-Schätzer (WLE)* oder der *Expected-a-posteriori-Schätzer (EAP)* (für einen Überblick siehe Walter, 2005). In der vorliegenden Arbeit wird der ***Weighted-Likelihood-Schätzer (WLE)*** (Warm, 1989) zur Berechnung der individuellen Personenparameter verwendet, der als bester unter den genannten Schätzern gilt (für detailliertere Informationen siehe Walter, 2005) und auch in den PISA-Studien zum Einsatz kam (OECD, 2005, S. 278). In Modellen mit mehreren Dimensionen (vgl. 5.1.2.4) wird dabei ein WLE pro Dimension geschätzt.

5.1.3 Zusammenfassung

Die Itemcharakteristik-Kurve (ICC) kann als zentrales Element der probabilistischen Testmodelle angesehen werden. Sie beschreibt den Zusammenhang zwischen der Wahrscheinlichkeit einer richtigen Itemantwort und der Fähigkeit einer Person. Durch die ICC werden für jedes Item zwei Parameter, Itemschwierigkeit und Trennschärfe, festgelegt.

Die Modelle der Rasch-Familie, die das dichotome Rasch-Modell sowie Erweiterungen auf ordinale Items und mehrere Fähigkeitsdimensionen umfasst, verwenden zur Modellierung der ICC eine logistische Funktion und nehmen ferner eine gleiche Trennschärfe aller Items an. Bei ihrer Berechnung wird damit die Itemschwierigkeit als einziger Itemparameter geschätzt. Zudem werden die Fähigkeiten der Probanden als Personenparameter (in der vorliegenden Arbeit als Werte des Weighted-Likelihood-Schätzers WLE) geschätzt. Item- und Personenparameter können auf einer gemeinsamen Logit-Skala dargestellt und so direkt miteinander verglichen werden. Anhand der Differenz der Fähigkeit einer Person und der Schwierigkeit eines Items kann leicht die Wahrscheinlichkeit ermittelt werden, mit der die Person das Item richtig löst.

5.2 Erhebungsverfahren

Die vorliegende Forschungsarbeit umfasst mehrere Teile. Der erste Teil ist die Entwicklung eines theoretischen Modells für die Diagrammkompetenz. Diese ist im theoretischen Teil der Arbeit dargestellt. Der zweite Teil, die empirische Überprüfung und Weiterentwicklung des Kompetenzstrukturmodells, ist seinerseits in zwei Studien gegliedert. Sie werden im nachfolgenden Abschnitt näher beschrieben.

5.2.1 Überblick über die Studien

Die empirische Überprüfung eines Kompetenzstrukturmodells erfordert die Konstruktion eines Testinstruments zur Erfassung der durch das Modell formulierten Komponenten respektive deren Unterteilungen in Unterkomponenten und Aspekte. Aus diesem Grund ging den empirischen Erhebungen die modellbasierte Entwicklung von Testitems voraus. Die Konstruktion dieser Items erfolgte unter zwei Maßgaben: Einerseits sollten sie die theoretisch abgeleiteten Strukturen vollständig repräsentieren, um das theoretische Modell überprüfbar zu machen, andererseits mussten sie sich zu einem ökonomischen Test zusammenstellen lassen, der mit vertretbarem Aufwand in der Schule verwendet werden kann. Dazu wurde ein umfangreicher Pool an Items entworfen. Beispiele der Testitems werden in Abschnitt 5.2.4.1 vorgestellt. Sämtliche Items wurden in der ersten empirischen Studie, die nachfolgend als **Pilotierung** bezeichnet wird, hinsichtlich ihrer Schwierigkeit, ihrer statistischen Güte und ihrer Verwertbarkeit in einem Test zur Diagrammkompetenz evaluiert. Über Rasch-Modellierungen wurden dabei insbesondere die Schwierigkeiten der Items geschätzt (vgl. 5.1.2). Anhand statistischer Gütekriterien sowie inhaltlicher Merkmale wurden Items für das endgültige Testinstrument, den **Haupttest**, ausgewählt. Die Kriterien für diese Itemselektion sind ausführlich unter 5.3.2.1 aufgeführt sowie die Auswahl unter 5.3.2.2 dokumentiert. Der Haupttest (siehe 5.2.4.4) wurde in der zweiten empirischen Studie, der **Haupttestung**, zur Erhebung eingesetzt. Diese Studie beschreibt die Überprüfung und Weiterentwicklung des Kompetenzstrukturmodells. Die zugehörigen Auswertungsverfahren werden unter 5.3.3 beschrieben.

Neben der Strukturanalyse der Diagrammkompetenz wurden explorativ verschiedene mögliche Einflussfaktoren auf die Kompetenz respektive ihre Komponenten untersucht. Items, die diese Faktoren erfassen sollten, wurden in den Test integriert und ebenfalls in der Pilotierung erprobt sowie für die Haupttestung einer Auswahl unterzogen. Die entsprechenden Variablen werden in Abschnitt 5.2.4.2 näher beschrieben.

5.2.2 Stichproben

Die Zielgruppe der vorliegenden Untersuchung waren Schüler der neunten und zehnten Klassenstufe in Gymnasien. Die Stichprobe für die Pilotierung setzte sich aus 152 Schülern aus sieben Klassen von drei Gymnasien in Schleswig-Holstein zusammen. Die Verteilung der Schüler nach Geschlecht und Klassenstufe kann Tabelle 5.1 entnommen werden. Da die Testaufgaben auf sechs unterschiedliche Testhefte verteilt waren, wurde jede Aufgabe durchschnittlich von 24-26 Personen bearbeitet. Eine Ausnahme stellen dabei die Ankeritems (vgl. 5.2.4.3) dar, die allen 152 Schülern vorgelegt wurden.

Tabelle 5.1: Stichprobengröße in der Pilotierung.

Pilotierung	Schüler	Schülerinnen	Gesamtzahlen
9. Klasse	19	18	37
10. Klasse	67	48	115
Gesamtzahlen	86	66	152

An der Haupttestung nahmen insgesamt acht Gymnasien teil. Dabei wurden 289 Schüler aus 22 Kursen untersucht. Bei dieser Testung wurde genau eine Testversion eingesetzt, so dass alle Aufgaben allen Schülern vorlagen. Aufgaben, insbesondere am Ende des Tests, die von weniger als 50 % der Schüler bearbeitet wurden, wurden jedoch von den Analysen ausgeschlossen.

Die Verteilung auf Schülerinnen und Schüler sowie auf die neunte und zehnte Klassenstufe war ausreichend balanciert. Die genauen Gruppengrößen in der Haupttestung werden in der unten stehenden Tabelle 5.2 aufgeführt.

Tabelle 5.2: Stichprobengröße in der Haupttestung. Die durch * gekennzeichneten Werte sind jeweils um eins höher als die Zeilensummen, da eine Person der neunten Klasse keine Angabe zum Geschlecht gemacht hat.

Haupttestung	Schüler	Schülerinnen	Gesamtzahlen
9. Klasse	63	70	134*
10. Klasse	76	79	155
Gesamtzahlen	139	149	289*

5.2.3 Organisation der einzelnen Testung

Die Pilotierung der Aufgaben wurde im Zeitraum Januar/Februar 2007 durchgeführt, die Haupttestung im Zeitraum Juni/Juli 2007.

In beiden Studien fand die Testung während des Schulbetriebs, wenn möglich im Biologieunterricht, statt. Zum Ausfüllen der Tests hatten die Schüler jeweils eine Schulstunde (45 Minuten) Zeit. Anlass und Zweck der Untersuchung wurden ihnen im Vorwege von der Lehrkraft mitgeteilt. Am Tag der Testung arbeiteten die Schüler nach einer kurzen, etwa dreiminütigen Einführung durch den Versuchsleiter⁴⁴ einzeln und selbständig.

5.2.4 Material

Die in der vorliegenden Arbeit eingesetzten Tests wurden basierend auf dem Modell der Diagrammkompetenz eigens für die vorgestellte Forschungsarbeit zu dessen Überprüfung entwickelt. Die Items wurden derart konstruiert, dass sie alle beschriebenen Komponenten (Informationsentnahme, Konstruktion oder Integration) und deren Unterkomponenten sowie feinere Aspekte abdeckten. Insbesondere wurde eine hohe „inhaltliche Relevanz“ (vgl. Reulecke & Rollett, 1976, S. 183) der Items angestrebt. Das bedeutet, dass die Items möglichst nah die tatsächlich im Schulunterricht auftretenden Anforderungen widerspiegeln sollten.

Um möglichst viele Items in den Test aufnehmen zu können, wurde eine Kombination aus Items im offenen Format und Single-Choice-Items gewählt. Erstere wurden entweder dichotom oder ordinal kodiert (vgl. 5.3.1).

Insgesamt wurden 157 Items entwickelt, die sich auf 40 Diagramme zu unterschiedlichen Themen verteilen. Dadurch wurden die einzelnen Fähigkeiten zum Umgang mit Diagrammen jeweils an mehreren Inhalten und bezogen auf zwei unterschiedliche Diagrammtypen abgeprüft. In vielen Fällen wurden mehrere Items anhand eines Diagramms bearbeitet. Das entsprechende Diagramm ist in der Benennung des Items gekennzeichnet, so zeigt beispielsweise der Anteil „Re“ in den Items *A_Re*, *Id_Re* und *K_Re_W* an, dass alle Items sich auf das Diagramm „Rentiere“ beziehen.

Eine Auflistung der Items, nach Diagrammen geordnet, befindet sich im Anhang (s. Tabelle 12.1). Im folgenden Abschnitt 5.2.4.1 werden ausgewählte Diagramme mit zugehörigen Items sowie in Abschnitt 5.2.4.2 exemplarische Items zur Erfassung des Lernerhintergrunds und der Lernermerkmale vorgestellt. Im Anschluss wird der Aufbau der Tests in Pilotierung (5.2.4.3) und Haupttestung (5.2.4.4) beschrieben.

⁴⁴ In Einzelfällen erfolgte die Einführung durch die Lehrkraft in einer vom Versuchsleiter vorgegebenen Form.

5.2.4.1 Ausgewählte Testitems zur Diagrammkompetenz

Informationsentnahme

Im Strukturmodell der Diagrammkompetenz wurden zwei Unterkomponenten der Informationsentnahme formuliert: Identifizierung und Ablesen. Zur Identifizierung wurden insgesamt 18 Items entwickelt, die sich beispielsweise auf das Erkennen des im Diagramm dargestellten Zusammenhangs oder das Erkennen der Skalenreichweite beziehen (vgl. 3.6.1). Die Abbildung 5.6 zeigt das Diagramm *Maus* mit einem zugehörigen Item, welches die dargestellte Relation abfragt.

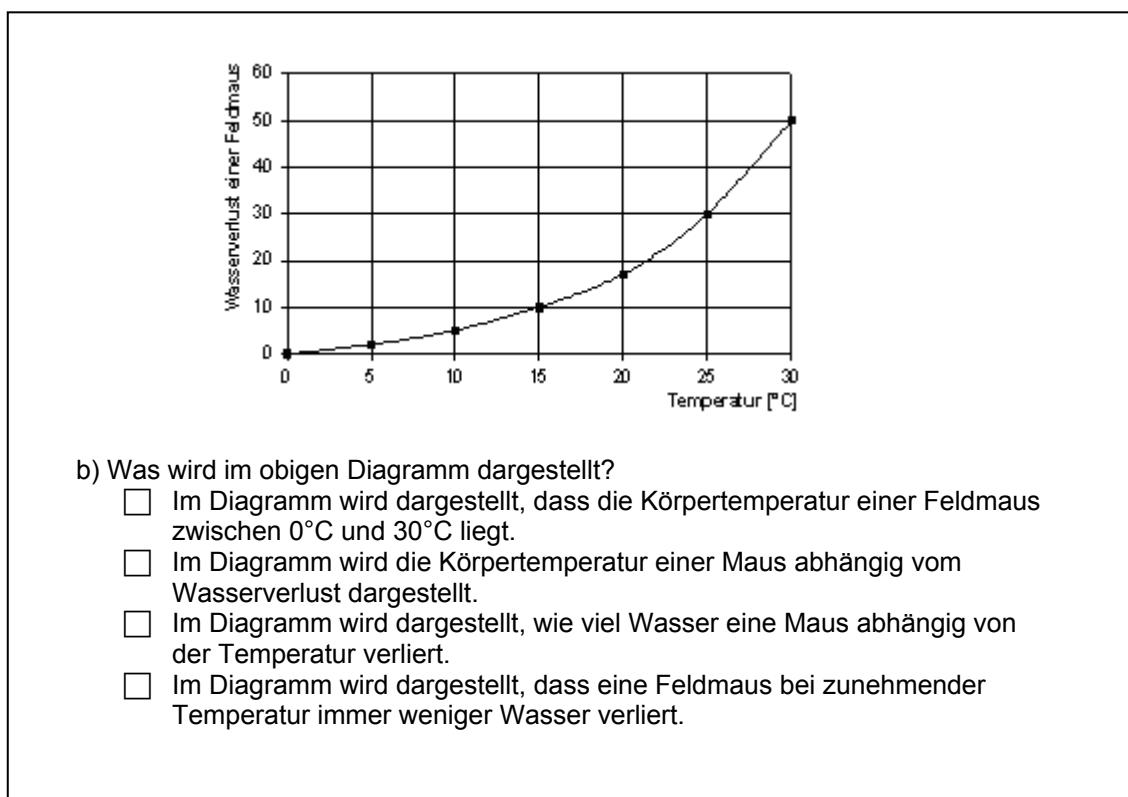


Abbildung 5.6: Diagramm *Maus* mit exemplarischem Item zur Identifizierung, welches das Erkennen der dargestellten Relation abfragt⁴⁵.

Die folgende Abbildung 5.7 stellt das Diagramm *Rentiere*⁴⁶ mit zugehörigem offenen Item zur Identifizierung dar, das sich auf die Veränderungen in einer Rentierpopulation auf der St. Matthew-Insel bezieht. Das Item „Mit welchem Jahr beginnt die Zeitskala im obigen Diagramm?“ zielt auf die Betrachtung der Skalenreichweite ab und erfragt die richtige Extrapolation der Skala.

⁴⁵ Das Fehlen der Einheit für den Wasserverlust ist beabsichtigt, da das Item auf das Erkennen der unvollständigen Achsenbeschriftung abzielt.

⁴⁶ Werte in Anlehnung an Bayrhuber & Kull, 2005, S. 93.

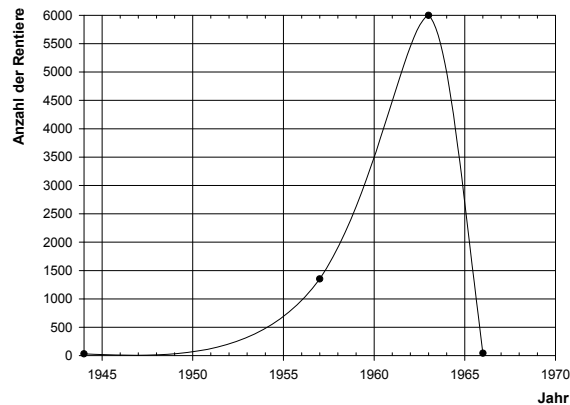


Abbildung 5.7: Diagramm *Rentiere* (im Testbogen ohne eingezeichnete Datenpunkte sowie Interpolationslinie).

Hinsichtlich der zweiten Unterkomponente der Informationsentnahme, dem Ablesen, wurden insgesamt 47 Items zu acht verschiedenen Diagrammen formuliert. Das Ablesen umfasst gemäß dem vorgeschlagenen Kompetenzstrukturmodell verschiedene Ordnungen. Im Folgenden sind beispielhaft Items aufgeführt, die diese verschiedenen Ordnungen repräsentieren. Sie beziehen sich auf das Diagramm *Papageitaucher*, welches im unten stehenden Abschnitt *Integration* abgebildet ist.

a) Welche Flügellänge hat die Rasse A? _____

c) Wie unterscheiden sich die Flügellängen der Rassen A und C?

- Die Flügel der Papageitaucher der Rasse A sind länger als die der Rasse C.
- Die Flügellängen der beiden Papageitaucherarten unterscheiden sich nicht.
- Die Vögel der Rasse C haben kürzere Flügel als die der Rasse A.
- Die Flügel der Papageitaucher der Rasse A sind kürzer als die der Rasse C.

d) Um wie viele Millimeter unterscheiden sich die Flügellängen der Rassen B und D?

- Um 15 Millimeter.
- Um 27 Millimeter.
- Um 30 Millimeter.
- Um 37 Millimeter.

g) Vergleiche den Unterschied in den Flügellängen der Rassen B und C mit dem der Rassen C und D!

- Der Unterschied zwischen den Rassen B und C beträgt 15 mm und ist daher größer als der zwischen den Rassen C und D.
- Der Unterschied zwischen den Rassen B und C ist um 20 mm geringer als der Unterschied zwischen den Rassen C und D.
- Der Unterschied der Rassen B und C ist kleiner als der Unterschied der Rassen C und D, der 23 mm beträgt.
- Der Unterschied zwischen den Rassen B und C ist genau so groß wie der zwischen den Rassen C und D.

Abbildung 5.8: Items zum Diagramm *Papageitaucher*, die verschiedene Ordnungen des Ablesens repräsentieren.

Item a) ist ein offenes Item zum Ablesen eines Punktwertes (einer Koordinate) und damit zum Ablesen erster Ordnung. Item c) fordert ein Ablesen zweiter Ordnung, da zwei Flügellängen verglichen werden müssen. Der hier nötige Vergleich ist qualitativer Natur (vgl. 3.5.2.1, s. Ablesen), da keine konkreten Flügellängen als Werte der abhängigen Variablen oder Differenzen zwischen diesen angegeben sind. Im Gegensatz dazu beschreibt das Item d) ein Ablesen zweiter Ordnung mit einem quantitativen Vergleich zweier Flügellängen. Das Item g) stellt ein Item zum Ablesen dritter Ordnung dar, bei dem Differenzen zwischen je zwei Flügellängen überwiegend quantitativ miteinander verglichen werden müssen.

Ein Item zum Ablesen vierter Ordnung bezieht sich auf das Diagramm *Rentiere* (vgl. Abbildung 5.7). Im zugehörigen Text sind die Anzahlen an Rentieren auf der Insel bis zum Jahr 1966 angegeben, so dass das Item „Die Winter nach 1966 waren auf der Insel mild. Wie groß war die Population schätzungsweise im Jahr 1970?“ auf eine Vorhersage bezüglich der Entwicklung der Rentierpopulation abzielt.

Konstruktion

Der Itempool umfasst insgesamt 47 Items zur Konstruktion. Dazu gehören einerseits Aufgabenstellungen, die sich jeweils auf einzelne Aspekte der Konstruktion beziehen. Ein Beispiel für eine solche Aufgabenstellung zeigt Abbildung 5.9. Anhand des dort abgebildeten Diagramms *Eier* wird die *Achsenbeschriftung* als ein Konstruktionsaspekt der Unterkomponente *Aufbau* abgeprüft. Gefordert ist der Eintrag der zwei Variablen *Tierart* an der Abszisse und *Eierzahl* an der Ordinate. Das Eintragen der zwei Tierarten an die jeweiligen Säulen (das als Konstruktionsleistung die Skalierung der nominalen UV beinhaltet) wird in Form eines Items zur Integration erfasst und unten im entsprechenden Abschnitt erläutert.

Im unten stehenden Diagramm sollen die Eierzahlen eines Sandhais und einer Python dargestellt werden (wobei die Python mehr Eier produziert).

Trage die fehlenden Beschriftungen ein!

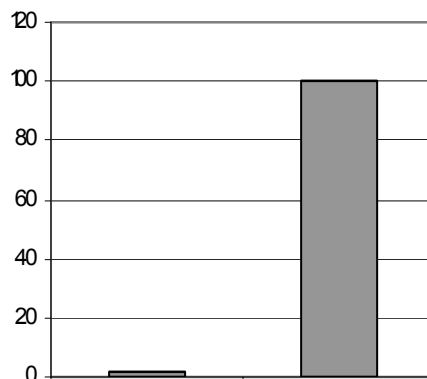


Abbildung 5.9: Diagramm *Eier*⁴⁷ mit Aufgabenstellung zur Konstruktion respektive Integration.

Neben dem Erfragen einzelner Aspekte umfassen die Items zur Konstruktion andererseits Aufgabenstellungen, die die Konstruktion eines vollständigen Diagramms fordern. Ein Beispiel für eine solche Aufgabenstellung ist in Abbildung 5.10 dargestellt.

Der Kiefernspinner ist ein Schmetterling, der bevorzugt in Kiefernwäldern vorkommt. Die Umgebungstemperatur hat einen großen Einfluss darauf, ob aus den abgelegten Eiern der Weibchen auch eine Raupe schlüpfen kann. Die höchste Schlupfrate zeigt sich bei einer Umgebungstemperatur von 20°C, in diesem Fall schlüpfen aus ca. 98% der Eier wirklich junge Raupen. Je kälter es ist, desto weniger Raupen schlüpfen relativ gesehen zur Zahl der Eier. Bei einer Temperatur von 10°C können noch 37% der Raupen schlüpfen, bei 5°C und darunter kann sich keine Kiefernspinnerraupe entwickeln. Ein ähnliches Bild zeigt sich bei Temperaturen über 20°C: die Schlupfrate sinkt kontinuierlich bis auf 0% bei 37°C.

Zeichne ein Diagramm, das den Einfluss der Temperatur auf die Schlupfrate des Kiefernspinners wiedergibt!

Abbildung 5.10: Aufgabenstellung zur Konstruktion des Diagramms *Kiefernspinner*⁴⁸. Zum Zeichnen wurde Millimeterpapier bereitgestellt.

Die einzelnen in der vorliegenden Arbeit untersuchten Aspekte der Konstruktion wurden gemäß dem Modell als unterschiedliche Items so kodiert, dass sie unabhängig voneinander sind, d. h. dass sich beispielsweise ein Fehler in der Achsenbeschriftung nicht ebenfalls als Fehler in der Achsenbelegung niederschlägt. Bezüglich der obigen Aufgabenstellung

⁴⁷ Werte aus Flindt, 2000, S. 92f.

⁴⁸ Werte in Anlehnung an Bayrhuber & Kull, 2005, S. 51.

wurden dabei die folgenden Items unterschieden, die jeweils durch die Skizzierung eines Erwartungshorizonts charakterisiert werden.

Auswahl des Diagrammtyps: Im Fall des Diagramms *Kiefernspinner* liegen zwei metrische Variablen vor. Der angemessene, zu zeichnende Diagrammtyp ist somit ein Liniendiagramm.

Achsenbelegung: Die Variable *Umgebungstemperatur* muss als unabhängige Variable identifiziert und der Abszisse zugeordnet werden, die Variable *Schlupfrate* als abhängige Variable der Ordinate zugeordnet werden.

Achsenbeschriftung: Die Beschriftung der Achsen umfasst die Angaben *Umgebungstemperatur* mit der Einheit *Grad Celcius* sowie *Schlupfrate* mit der Einheit *Prozent*.

Skalierung: An beiden Achsen wird eine metrische Skalierung erwartet. Dabei wird insbesondere darauf Wert gelegt, dass die gewählten Maßstäbe ausreichend groß sind, um einen exakten Werteintrag zuzulassen. Beispielsweise wird ein Abstand von 1 cm zwischen zwei Markierungen, die eine Differenz von 20 % in der Schlupfrate darstellen, als nicht ausreichend groß eingestuft, da die Markierung des Wertes 37 % auf Millimeterpapier nicht exakt möglich ist. Ein Abstand von 2 cm für 20 % wäre ein möglicher ausreichend großer Abstand.

Werteintrag: Die im Text genannten Werte müssen mit hinreichender Genauigkeit (genauere Angaben dazu befinden sich im Kodierleitfaden) auf der Diagrammfläche markiert werden.

Skizzierung: Gemäß den Angaben im Text beispielsweise über ein kontinuierliches Absinken der Schlupfrate bei Erhöhung der Umgebungstemperatur soll eine Interpolationslinie skizziert werden, die die eingezeichneten Datenpunkte berührt.

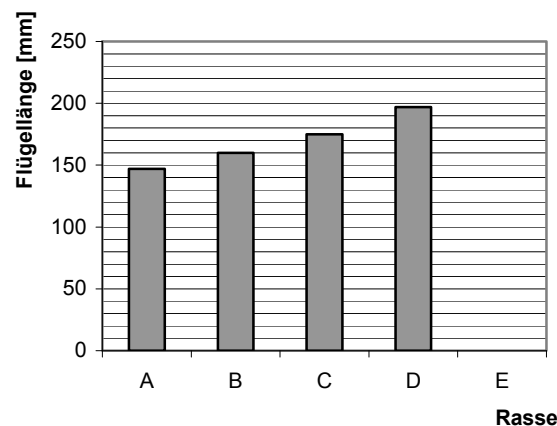
Dem Kompetenzmodell (vgl. 3.6, Tabelle 3.1) entsprechend werden die ersten vier Aspekte der Unterkomponente Aufbau und die letzten zwei Aspekte der Unterkomponente Dateneintrag zugeordnet.

Integration

Bezogen auf die Integration wurden insgesamt 30 Items entwickelt. Es lassen sich zwei Typen von Items unterscheiden: Diejenigen, bei denen zur Verknüpfung der Informationen aus Text und Diagramm eine (zumindest teilweise) Diagrammkonstruktion gefordert ist (*Resultat Diagramm*, vgl. 3.6.3), und diejenigen, bei denen die genannte Verknüpfung in Form von Text offen gelegt werden muss (*Resultat Text*). Ein Beispiel für ein Item des

ersten Typs bezieht sich, wie oben bereits erwähnt, auf das Diagramm *Eier* (siehe Abbildung 5.9), bei dem die Skalierung der nominalen Skala für die Variable *Tierart* als Integrationsleistung bewertet wird. Dabei muss die im Text gegebene Information, dass eine Python mehr Eier als ein Sandhai produziert, auf die zwei unterschiedlich hohen Säulen des Diagramms bezogen werden, wobei ein Vergleich der beiden Säulenhöhen und damit quasi ein Ablesen zweiter Ordnung nötig ist. Eine erfolgreiche Verknüpfung der Information aus beiden Quellen besteht somit darin, die Kategorie Python unter die höhere und die Kategorie Sandhai unter der niedrigeren Säule einzutragen. Die Integration umfasst somit die Verarbeitung der textlich gegebenen Information, einen Ableseprozess sowie die Skalierung als Konstruktionsleistung.

Zwei Itembeispiele des zweiten Typs zeigt die Abbildung 5.11 zum Diagramm *Papageitaucher*. Das textliche Resultat der Integration ist in diesem Fall nicht selbst zu produzieren, sondern als Auswahl anzukreuzen. Die richtige Antwort erfordert dabei die Verknüpfung von im Text gegebenen Informationen über das Verbreitungsgebiet und von im Diagramm gegebenen Informationen über die Flügellänge der jeweiligen Rassen. Item a) verlangt dabei ein Ablesen erster Ordnung, Item b) einen Vergleich mehrerer Werte und somit ein Ablesen dritter Ordnung.



Das oben stehende Diagramm zeigt die Flügelängen verschiedener Rassen des Papageitauchers. Der folgende Text beschreibt das Vorkommen der vier Papageitaucherrassen hinsichtlich der Breitengrade nördlicher Breite. Dabei bedeutet eine höhere Gradzahl nördlicher Breite eine nördlichere Lage.

Die Rasse A kommt in der Bretagne vor, die sich unter 50° nördlicher Breite erstreckt. Rasse B findet man an Küsten Irlands, Englands und Schottlands (etwa 50° bis 59° nördlicher Breite). Die Rasse C kann an den Küsten Südgrönlands, Islands und Norwegens angetroffen werden. Diese liegen zwischen etwa 60° und 70° nördlicher Breite. Das Vorkommen der Rasse D ist auf Gebiete über 70° nördlicher Breite beschränkt, so etwa Nord-Norwegen und Spitzbergen.

a) Wo kommt die Rasse mit einer Flügelänge von 160 mm vor?

- An den Küsten Südgrönlands.
- An den Küsten der Bretagne.
- In Gebieten zwischen etwa 50° und 59° nördlicher Breite.
- In Gebieten zwischen 70° und 90° nördlicher Breite.

b) Welche Aussage kann man aus den gegebenen Daten ableiten?

- Die Rasse mit den längsten Flügeln kommt im südlichsten Gebiet vor.
- Die Rasse mit den zweitlängsten Flügeln kommt im nördlichsten Gebiet vor.
- Die Rasse mit den kürzesten Flügeln kommt im südlichsten Gebiet vor.
- Die Rasse mit den kürzesten Flügeln kommt im nördlichsten Gebiet vor.

Abbildung 5.11: Diagramm *Papageitaucher*⁴⁹ mit zusätzlichem Text und zwei Items zur Integration.

5.2.4.2 Itembeispiele zur Erfassung des Lernerhintergrunds und der Lernermerkmale

Wie aus dem theoretischen Hintergrund abzuleiten ist, wurden das *Geschlecht* und die *Klassenstufe* der Teilnehmer als Variablen zur Erfassung des **Lernerhintergrunds** erfragt. Zudem wurden verschiedene **Lernermerkmale** aufgrund der Literaturhinweise auf einen

⁴⁹ In Anlehnung an Bayrhuber & Kull, 2005, S. 132.

möglichen Einfluss auf die Testleistung der Schüler hinsichtlich des Umgangs mit Diagrammen erhoben. Diese werden nachfolgend anhand exemplarischer Items vorgestellt.

Biologisches Vorwissen. Zur Erfassung des biologischen Vorwissens wurden zehn Single-Choice-Items zu Inhalten konstruiert, die auch in den Diagrammen zur Testung der Diagrammkompetenz vorkommen. Ein Beispiel zeigt die folgende Abbildung 5.12.

<p>Eine Population ist...</p> <p><input type="checkbox"/> ...eine Gruppe von Tieren einer Art, die alle von einem Vorfahr abstammen.</p> <p><input type="checkbox"/> ...eine Gruppe von Tieren einer Art, die alle gleich alt sind.</p> <p><input type="checkbox"/> ...eine Gruppe von Tieren einer Art, die in einem definierten Gebiet leben.</p> <p><input type="checkbox"/> ...eine Gruppe von Tieren einer Art, die alle gleich groß sind.</p>

Abbildung 5.12: Beispiel für ein Item zum biologischen Vorwissen.

Neben dem biologischen Vorwissen wurde die letzte **Zeugnisnote** in **Biologie** erfragt. Darüber hinaus sollte von den Schülern die letzte Zeugnisnote in **Mathematik** angegeben werden.

Lesekompetenz. Zur Einschätzung der Lesekompetenz wurde der Text des Aufgabenbeispiels „Grippe“ der PISA 2000-Studie (OECD, 2000) in gekürzter Form verwendet. Es handelt sich dabei um eine Aufgabe zur Lesekompetenz, die der Subskala „Informationen ermitteln“ zugeordnet ist. Der Aufgabentext ist relativ kurz und enthält kaum konkurrierende Informationen, so dass die beiden für die vorliegende Arbeit übernommenen Single-Choice-Items hinsichtlich der Kompetenzstufen nach PISA den unteren Kompetenzstufen I und II (vgl. Artelt et al., 2001) zuzurechnen sind. Sie prüfen das Auffinden von Informationen im Text ab.

Rechnen. Zur Einschätzung der Rechenfähigkeiten der Schüler wurden fünf Items zu Subtraktion und Addition konstruiert, die den Rechenoperationen ähneln, die bei der Bearbeitung der Aufgaben zur Diagrammkompetenz durchgeführt werden mussten. Ein Beispiel ist die Aufgabe $910\,000 - 91\,000 = \underline{\hspace{2cm}}$.

Umgang mit abstrakten Diagrammen. Zu Erhebung der Fähigkeiten zum Umgang mit abstrakten Diagrammen wurden 18 Items, darunter sowohl Single-Choice-Items als auch offene Items, entwickelt. Abstrakt bedeutet in diesem Fall, dass die zu bearbeitenden Diagramme nicht mit einem konkreten biologischen Inhalt verknüpft sind, die Variablen werden beispielsweise als Faktor X und Faktor Y bezeichnet. Wie die Testitems zur Diagrammkompetenz sind auch die Items zum Umgang mit den abstrakten Diagrammen jeweils bestimmten Aspekten des Kompetenzstrukturmodells zugeordnet, sie bilden das Modell allerdings in ihrer Funktion einer Kontrollvariable nicht vollständig ab. Ein Beispiel für ein Item zum Umgang mit abstrakten Diagrammen zeigt die folgende Abbildung 5.13.

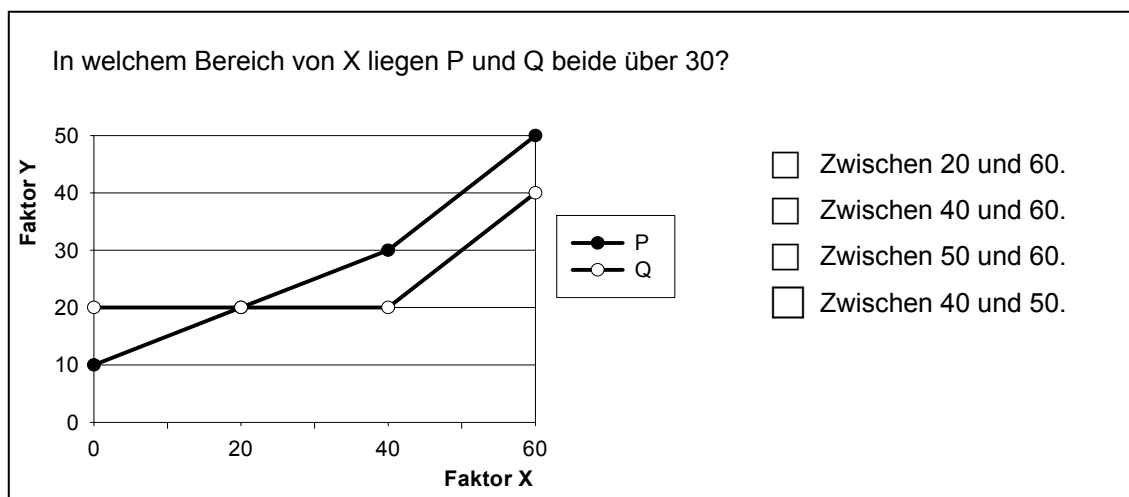


Abbildung 5.13: Beispiel für ein Item zum Umgang mit abstrakten Diagrammen.

Anstrengung. Im Haupttest wurde zudem eine modifizierte Version des „Anstrengungsthermometers 2“ aus PISA 2003 (PISA-Konsortium Deutschland, 2006a, S. 125f) eingesetzt, um einen Hinweis auf die Motivation und *Anstrengung* der Schüler bei der Testbearbeitung zu erhalten. Aufgabe der Schüler war es, ihre Anstrengung während der Bearbeitung des Tests auf einer zehnstufigen Likert-Skala einzuordnen, wobei der Wert zehn die in einer persönlich wichtigen Situation maximal aufwendbare Anstrengung bezeichnen sollte.

5.2.4.3 Testhefte der Pilotierung

Die Pilotierung wurde mit der Zielsetzung durchgeführt, die Items unter realen Bedingungen zu erproben und anhand verschiedener Kennwerte (vgl. 5.3.2.1) geeignete Aufgaben für den Haupttest auszuwählen. Die Analyse der Daten mit probabilistischen Methoden erlaubt dabei, Itemschwierigkeiten von Items zu vergleichen, die nicht von den gleichen Testpersonen beantwortet wurden, und ermöglicht ein Multi-Matrix-Design der

Tests. Dem folgend wurden sämtliche Items auf fünf Testhefte aufgeteilt, wobei einige Ankeritems, d. h. Items, die in allen Testheften vorkamen, integriert wurden. In jedem Testheft befand sich eine Auswahl von Items, durch die alle drei theoretisch postulierten Komponenten (Informationsentnahme, Konstruktion und Integration) sowie deren Unterkomponenten (Identifizierung und Ablesen sowie Aufbau und Dateneintrag) abgedeckt wurden. Die Items zum biologischen Vorwissen, zur Lesekompetenz und zu den Rechenfähigkeiten sowie die abstrakten Diagrammaufgaben bildeten ein sechstes Testheft. Auch in dieses Testheft wurden die Ankeritems aufgenommen. Die genaue Verteilung der Items auf die Testhefte findet sich in der Tabelle 12.1 im Anhang.

5.2.4.4 Haupttest

Die Kriterien zur Auswahl der Items für den Haupttest werden in Abschnitt 5.3.2.1 detailliert beschrieben und die Itemselektion wird in Abschnitt 5.3.2.2 (sowie im Anhang 12.4) dokumentiert. An dieser Stelle ist festzuhalten, dass das endgültige Testinstrument insgesamt 23 Items zur Informationsentnahme, 18 Items zur Konstruktion sowie zehn Items zur Integration enthält⁵⁰. Die Informationsentnahme umfasst sechs Items zur Identifizierung und 17 Items zum Ablesen, die Konstruktion umfasst zwölf Items zum Aufbau und sechs zum Werteintrag.

Um die benötigte Zeit zur Bearbeitung des Tests zu begrenzen, wurde die Itemzahl zur Erfassung des Lernerhintergrunds im Haupttest reduziert. Dabei wurden die gleichen Kriterien wie für die Auswahl der Items zur Diagrammkompetenz zugrunde gelegt, die Dokumentation der entsprechenden Kennwerte befindet sich im Anhang (s. 12.4). Im Haupttest wurden mit jeweils einen Item die Lesekompetenz und die Rechenfähigkeiten der Schüler kontrolliert. Das biologische Vorwissen sowie die Fähigkeiten zum Umgang mit abstrakten Diagrammen wurden mit jeweils vier Items erhoben. Sämtliche Haupttest-Items zu den Lernermerkmalen sind im Anhang (s. 12.3.2) aufgeführt.

5.3 Auswertungsverfahren

Die Darstellung der Auswertungsverfahren ist in drei Teile gegliedert. Im ersten Teil wird die Kodierung der Testaufgaben (5.3.1) vorgestellt. Im zweiten Teil werden die Kriterien der Itemselektion erläutert und die Auswahl der Items dokumentiert (5.3.2). Im dritten Teil werden die Methoden beschrieben, die zur Bearbeitung der Forschungsfragen eingesetzt wurden (5.3.3).

⁵⁰ In die Auswertungen konnten allerdings nicht alle dieser Items einbezogen werden (vgl. 5.3.2.2).

5.3.1 Kodierung

Wie oben beschrieben (vgl. 5.2.4), wurden in der vorliegenden Studie sowohl Single-Choice-Aufgaben als auch Aufgaben mit offenem Antwortformat verwendet. Die Schülerantworten wurden entweder dichotom als „richtig“ oder „falsch“ oder ordinal als „falsch“, „teilweise richtig“ und „vollständig richtig“ bewertet. Gemäß diesen Einteilungen wurden die Punktzahlen 0 und 1 bzw. 0, 1 und 2 vergeben und die jeweiligen Antworten auch mit *0-Antwort*, *1-Antwort*, *2-Antwort* oder *0-Kategorie*, *1-Kategorie*, *2-Kategorie* benannt.

Um qualitative Unterschiede in den Schülerfähigkeiten beim Umgang mit Diagrammen aufdecken zu können, wurden zum Teil mehrere inhaltliche Antwortkategorien formuliert, die *einer* Punktzahl zugeordnet wurden. Um Verwechslungen mit den Punktzahlkategorien 0, 1 und 2 zu vermeiden, werden diese inhaltlichen Kategorien im Folgenden als **Rubriken** bezeichnet. Die Implementation des Ordered-Partition-Modells (vgl. 5.1.2.3) in das mehrdimensionale Rasch-Modell erlaubt in dem Fall eine Modellierung der Daten unter Einbeziehung sämtlicher Rubriken.

Die Dateneingabe erfolgte auf der Ebene der Rubriken. Hierfür wurde für die Items in offenem Antwortformat in einer Kombination aus induktiver und deduktiver qualitativer Inhaltsanalyse (vgl. Mayring, 2000) ein ausführlicher Kodierleitfaden angefertigt, der genaue Angaben darüber enthält, wie die Antworten zu bewerten sind. Der deduktive Anteil bestand darin, dass im Vorwege Rubriken zu aus der Literatur bekannten häufigen Fehlern (so beispielsweise die nicht-metrische Skalierung, vgl. 3.5.2.2) formuliert wurden. Der induktive Anteil war die Entwicklung zusätzlicher Rubriken beim Auftreten von Schülerantworten, die mit den vorweg formulierten Rubriken nicht erfasst werden konnten, respektive die Systematisierung der Antworten zu den Aspekten, für die im Vorfeld keine literaturbasierten Rubriken formuliert werden konnten.

Durch das Aufgreifen aus der Literatur bekannter Fehler wird ein Vergleich von Ergebnissen der vorliegenden Arbeit mit anderen empirischen Ergebnissen ermöglicht. Darüber hinaus wird dadurch, dass die Distraktoren in Single-Choice-Aufgaben bekannte Fehlertypen aufgreifen, die Wahrscheinlichkeit, die richtige Lösung durch Raten zu finden, minimiert (Lind & Knoche, 2004).

Im Fall der Single-Choice-Aufgaben, die vier verschiedene Lösungsmöglichkeiten enthalten, steht jeder Distraktor 1 bis 4 für eine Rubrik 1 bis 4. Bei der Kodierung der offenen Aufgaben werden die Rubriken durch eine zweistellige Ziffernkombination benannt: die erste Stelle gibt an, welcher Punktzahl die jeweilige Rubrik zugeordnet wird, in der zweiten Stelle sind die Rubriken durchnummeriert. Die folgende Tabelle 5.3 zeigt am Beispiel des Kodierleitfadens zur Achsenbeschriftung (als einem Aspekt der Konstruktion) die Formulierung solcher Rubriken.

Tabelle 5.3: Kodierleitfaden für die Achsenbeschriftung⁵¹.

Rubrik 01	eine Achse nicht, zweite Achse nur durch Einheit beschriftet
Rubrik 02	eine Achse nicht, zweite Achse ohne Einheit beschriftet
Rubrik 03	eine Achse nicht, zweite Achse mit Einheit beschriftet
Rubrik 11	beide Achsen nur durch Einheit beschriftet
Rubrik 12	eine Achse nur durch Einheit, zweite Achse ohne Einheit beschriftet
Rubrik 13	eine Achse mit Einheit, zweite nur durch Einheit beschriftet
Rubrik 14	beide Achsen ohne Einheiten beschriftet
Rubrik 15	eine Achse ohne Einheit, zweite mit Einheit beschriftet
Rubrik 21	beide Achsen mit Einheiten beschriftet

Die Zuverlässigkeit des Kodierleitfadens wurde anhand der Interrater-Reliabilität überprüft. Da sich einzelne Rubriken auf eine Punktzahl beziehen, beschreiben die unterschiedlichen Rubriken keine ordinale Skala, sondern sind als Ausprägungen einer nominalen Variable aufzufassen. Deshalb wurde Cohens Kappa (Cohens K) als ein mögliches Maß für die Interrater-Reliabilität gewählt (vgl. Wirtz & Caspar, 2002). Dieses Maß setzt die Zahl der auftretenden Übereinstimmungen zweier bewertender Personen zu der Zahl der durch Zufall zu erwartenden Übereinstimmungen in Beziehung. Die resultierende Maßzahl zeigt an, wie stark die tatsächlich auftretenden Übereinstimmungen die nach Zufall zu erwartenden übersteigen, ein Wert von Cohens K > 0.75 wird dabei als Indikator für sehr hohe Übereinstimmung angesehen (Wirtz & Caspar, 2002, S. 59). Anhand einer zufällig ausgewählten Teilstichprobe von 25 Schülern wurde für die offenen Items der vorliegenden Arbeit ein Cohens K von 0.845 ermittelt. Der Kodierleitfaden ist damit als zuverlässig einzustufen.

5.3.2 Itemselektion

Die Entwicklung eines angemessenen Tests zur Erfassung der Diagrammkompetenz stellt ein zentrales Ergebnis der vorliegenden Arbeit dar. Notwendige Voraussetzung hierfür war die *Konstruktion eines geeigneten Pools an Items* basierend auf dem Modell der Diagrammkompetenz als theoretischem Rahmen. Sämtliche Items wurden in der *Pilotierung* unter realen Bedingungen erprobt und einer Rasch-Modellierung unterzogen. Zur Auswahl derjenigen Items, die im Haupttest eingesetzt werden sollten, wurde eine differenzierte Itemanalyse durchgeführt. Dabei wurden unterschiedliche statistische Kriterien beachtet und gegeneinander abgewogen. Diese Kriterien werden nachfolgend dargestellt (5.3.2.1), bevor die Itemselektion dokumentiert wird (5.3.2.2).

⁵¹ Zur kürzeren Darstellung werden die Ankerbeispiele sowie die Kodierregeln (vgl. Mayring, 2000) hier nicht aufgeführt.

5.3.2.1 Kriterien der Itemselektion

Itempassung: Weighted mean square

Der entscheidende Kennwert, anhand dessen die Güte eines einzelnen Items beurteilt wird, ist der WMNSQ-Wert (weighted mean square). Er dient als Maß dafür, wie gut die empirisch ermittelten ICCs (vgl. 5.1.1.1) mit den durch die geschätzten Itemparameter bestimmten ICCs übereinstimmen. Dazu wird die tatsächliche Varianz der Residuen, also der Differenzen zwischen beobachteten und geschätzten Werten, mit der unter Modellgeltung geschätzten Varianz verglichen. Werte nahe eins deuten auf eine gute Übereinstimmung und damit eine gute Itempassung hin (Wilson, 2005). Als Grenzen für den WMNSQ werden von Adams und Khoo (1996, zitiert in Wilson, 2005, S. 129) 0,75 und 1,33 empfohlen, wobei Werte kleiner als 1 im Allgemeinen eher als unproblematisch angesehen werden, da sie auf eine steile ICC und damit eine hohe Trennschärfe (im probabilistischen Sinne) hinweisen.

Als zugehörige statistische Größen wurden die t-Werte herangezogen, die eine Transformation des WMNSQ in eine Normalverteilung beschreiben. Sie geben Aufschluss darüber, ob die durch den WMNSQ angedeutete Abweichung signifikant ist oder nicht. t-Terte im Intervall $[-2,2]$ weisen auf eine nicht signifikante Abweichung hin. In der vorliegenden Arbeit wurden Items mit abweichendem WMNSQ als kritisch eingestuft bzw. direkt entfernt, wenn zusätzlich der zugehörige t-Wert außerhalb des angegebenen Intervalls lag.

Itemschwierigkeiten und Schwellenparameter

Anders als in der klassischen Testtheorie, in der die Itemschwierigkeit daraus bestimmt wird, welcher Prozentsatz der Probanden eine Aufgabe richtig beantworten kann, werden die Itemschwierigkeiten in den probabilistischen Modellen als Parameter getrennt von den Personenfähigkeiten geschätzt. Durch den Abgleich von Item- und Personenparametern kann sichergestellt werden, dass der Test für eine Stichprobe im Schwierigkeitsgrad angemessen ist, und so einem Decken- oder Bodeneffekt vorgebeugt werden (Rost, 2004, S. 92). Da bei den Analysen von ConQuest (Wu et al., 1998) der Mittelwert der Personenfähigkeit auf 0 festgelegt wird, sollten demnach auch die Itemschwierigkeiten um 0 herum variieren. Der nach Baker (2001, S. 5) dazu empfehlenswerte Bereich ist das Intervall $[-3;3]$. In Folge dessen wurden in der vorliegenden Arbeit vorrangig Items für den Haupttest ausgewählt, deren Schwierigkeit innerhalb dieses Bereichs lag.

Ein weiteres Kriterium zur Beurteilung der Itemgüte hinsichtlich der Schwierigkeit war die Überprüfung, ob die Schwellen nach aufsteigenden Kategorien geordnet sind, da eine niedrigere Kategorie nicht einfacher als eine höhere Kategorie sein sollte (OECD, 2005,

S. 125). Dazu ist sicherzustellen, dass die Schwelle s_1 zwischen den Kategorien 0 und 1 kleiner ist als die Schwelle s_2 zwischen den Kategorien 1 und 2. Eine andere Möglichkeit, die Ordnung der Schwellen zu überprüfen, besteht darin, die Abweichungen der jeweiligen Schwellen vom Schwellenmittelwert zu analysieren, welche vom ConQuest (Wu et al., 1997) bei der Itemanalyse als τ_1 und τ_2 ausgegeben werden (vgl. OECD, 2005, Appendix 15). Dabei sind diese Abweichungen τ_i für $i=1,2$ die Differenzen von s_i und s_M , so dass die Ordnung der Schwellen $s_1 < s_M < s_2$ darüber überprüft werden kann, dass τ_1 negativ und τ_2 positiv ist⁵².

Da τ_1 und τ_2 nach Definition den gleichen Abstand vom Schwellenmittelwert haben, sind sie vom Betrag her gleich und durch Multiplikation mit dem Faktor -1 ineinander überführbar, weshalb die Überprüfung der Negativität von τ_1 als Kriterium für die Ordnung der Schwellenparameter genügt. Zur Gewichtung dieses Kriteriums wird indes der Standardfehler SE_τ berücksichtigt: ist er größer als $|\tau_1|=|\tau_2|$, so kann ein positiver Wert für τ_1 vernachlässigt werden. Lediglich im Fall $|SE_\tau| < |\tau_1|$ weist ein positives Vorzeichen bei τ_1 auf eine falsche Ordnung der Schwellen hin.

An dieser Stelle wird darauf hingewiesen, dass die Schwellenparameter Kennzahlen auf Kategorienebene sind, im Vergleich dazu bezieht sich die Itemschwierigkeit auf das gesamte Item. Aufgrund der geringen Stichprobengröße pro Item sind die Schwellenparameter mit Vorsicht zu interpretieren und werden deshalb nicht direkt als Ausschlusskriterium angesehen. Stattdessen führen sie zur kritischen Beachtung des jeweiligen Items (vgl. 5.3.2.1, s. Zusammenfassung).

Klassische Trennschärfe

Als weiteres Selektionskriterium wurde die Trennschärfe im Sinne der klassischen Testtheorie herangezogen⁵³. Dabei wird die Trennschärfe des Items als Produkt-Moment-Korrelation (**PMK**) zwischen dem Itemscore und dem Gesamtscore der Skala ermittelt (Bortz, 1999, S. 215; Fisseni, 1997, S. 49). Sie sollte größer als 0,25 sein (OECD, 2005, S. 217).

Entsprechend den Schwellenparametern, die als Schwierigkeitsmaß für die Punktzahl-Kategorien eines ordinalen Items fungieren, kann eine Trennschärfe ebenfalls auf Kategorienebene und sogar auch auf Rubrikenebene (vgl. 5.3.1) berechnet werden. In diesem Fall wird für jede Rubrik eine dichotome Indikatorvariable mit den Ausprägungen „hat die Rubrik gewählt“ und „hat die Rubrik nicht gewählt“ ermittelt und die Trennschärfe als punkt-biseriale Korrelation (**PBK**) zwischen dieser Indikatorvariable und dem Gesamtscore berechnet (Wu et al., 1998, S. 148). In einem guten Item sollten die PBK für die jeweils höher bewerteten Rubriken höher als für niedriger bewertete und für falsche

⁵² Denn wegen $\tau_1 = s_1 - s_M$ ist τ_1 genau dann negativ, wenn $s_1 < s_M$ ist. Analog für τ_2 .

⁵³ Nicht zu verwechseln mit dem probabilistischen Konzept der Trennschärfe.

Rubriken insbesondere nicht positiv sein (OECD, 2005, S. 123). Entsprechend wurden in der vorliegenden Arbeit Items, die für eine Null-Rubrik eine signifikante⁵⁴ PBK_0 mit $r > 0,05$ aufwiesen, aussortiert. Durch die Voraussetzung der Signifikanz wird dabei der geringen Stichprobe Rechnung getragen⁵⁵. Die folgende Abbildung 5.14 zeigt ein Beispiel für die vom Programm ConQuest ausgegebene Itemstatistik, in der die punkt-biserialen Korrelationen (Pt Bis) für jede Rubrik und die Signifikanzwerte (p) abgelesen werden können. In diesem Beispiel sind lediglich die Korrelationen der Rubriken 1 und 3 signifikant, wobei davon die PBK (engl. PT Bis) der 0-Rubrik negativ ist – für dieses Item ist damit die PBK_0 kein Ausschlusskriterium.

```

item:1 (Id_K_a) .....
Cases for this item 150 Discrimination 0.26
Item Threshold(s) -1.54 Weighted MNSQ 1.06
-----
Label Score Count % of tot Pt Bis t (p)
-----
1 0.00 2 1.33 -0.24 -3.04 (.003)
2 0.00 24 16.00 -0.13 -1.65 (.101)
3 1.00 121 80.67 0.26 3.29 (.001)
4 0.00 1 0.67 -0.03 -0.42 (.676)

```

Abbildung 5.14: Auszug aus der Itemstatistik für ein Single-Choice-Item zur Identifizierung. Die Spalte *Label* enthält die Rubriken 1 bis 4, Spalte *Score* deren Punktzahl. Für jede Rubrik sind in der Spalte *Pt Bis* deren punkt-biserialen Korrelation sowie die zugehörige *t*-Statistik mit Signifikanzwert *p* aufgeführt.

Zusammenfassung der Kriterien

In der nachfolgenden Tabelle 5.4 sind die Kriterien zur Itemselektion in Kurzform aufgelistet. Sie werden dabei in *Ausschlusskriterien* und *Beachtungskriterien* unterteilt. Erstere führten zu einer Verwerfung des entsprechenden Items, sobald eines der Kriterien erfüllt war. Letztere führten nicht direkt zur Verwerfung sondern zur Einstufung als „kritisches“ Item. Ein solches Item wurde dennoch bevorzugt verworfen, wenn noch andere, nach den genannten Kriterien geeignete Items der gleichen inhaltlichen Komponente zur Auswahl standen.

⁵⁴ Als Signifikanzniveau wurde das 5 %-Niveau gewählt.

⁵⁵ Vgl. die Schwellenparameter im vorigen Abschnitt, deren Nichtordnung aufgrund der geringen Stichprobengröße lediglich als Beachtungskriterium gewertet wird.

Tabelle 5.4: Kriterien der Itemselektion.

Ausschlusskriterien
Der WMNSQ lag außerhalb des Intervalls]0,75;1,33[und der zugehörige t-Wert außerhalb des Intervalls]-2;2[.
Die Itemschwierigkeit lag außerhalb des Intervalls [-3;3].
Die Trennschärfe (gemessen als PMK) war kleiner als 0,25.
Eine signifikante PBK einer Null-Rubrik war größer als 0,05.
Beachtungskriterien
Nur der WMNSQ lag außerhalb des angegebenen Intervalls (nicht der t-Wert).
Die Schwellen waren nicht geordnet, d. h. τ_1 war positiv mit $ SE\tau < \tau_1 $.

5.3.2.2 Dokumentation der Itemselektion

In der Pilotierung wurden insgesamt 157 Items erprobt. Die Itemanalyse erfolgte unter Abwägung der genannten Kriterien sowie inhaltlicher Argumente, so dass sichergestellt wurde, dass innerhalb der drei postulierten Komponenten Informationsentnahme, Konstruktion und Integration auch alle einzelnen Aspekte durch Items repräsentiert wurden. Aus diesem Grund wurden die Kennwerte der Items vorrangig innerhalb der Komponenten miteinander verglichen.

Da mithilfe des in das Programm ConQuest (Wu et al, 1998) implementierten Algorithmus nur Items analysiert werden können, die eine echte Varianz aufweisen, die also nicht von allen Personen entweder richtig oder falsch beantwortet wurden, mussten in einem ersten Schritt bereits 39 Items von den Analysen ausgeschlossen werden⁵⁶. In weiteren Schritten wurden Items eliminiert, die eins der beschriebenen *Ausschlusskriterien* (vgl. Tabelle 5.4, Abschnitt 5.3.2.1) erfüllten. Items, die eins der *Beachtungskriterien* (vgl. Tabelle 5.4, Abschnitt 5.3.2.1) erfüllten, wurden lediglich als kritisch eingestuft und gegebenenfalls aus triftigen inhaltlichen Gründen für den Haupttest ausgewählt, wie nachfolgend erläutert wird. Die Itemkennwerte sind im Anhang (s. 12.4.1.2) aufgeführt.

Informationsentnahme

Sämtliche Items, die der Informationsentnahme zugeordnet wurden, wiesen eine ausreichende Itempassung hinsichtlich der WMNSQ- und t-Werte auf (vgl. Tabelle 12.2)

⁵⁶ Eine Liste dieser Items befindet sich im Anhang (s. 12.4.1.1).

im Anhang). Das Item *Id_K_b* wurde als einziges aufgrund einer zu geringen Schwierigkeit ausgeschlossen. Kein Item zeigte eine zu hohe signifikante PBK einer Null-Rubrik, so dass sich für die Items der Informationsentnahme die generelle (klassische) Trennschärfe als Hauptausschlusskriterium herausstellte.

Da sich in der Pilotierung gezeigt hatte, dass die Bearbeitung des Tests mehr Zeit als im Vorwege eingeschätzt einnahm, wurde die Itemanzahl weiter reduziert, indem einige Items (in der Spalte „selektiert“ der Tabelle 12.2 in grauer Schrift dargestellt) trotz passender Itemkennwerte aussortiert wurden. Das Anliegen, möglichst viele Items und damit möglichst viele verschiedene Diagramme in den Haupttest aufzunehmen, wurde gegen die Testlänge abgewogen. Aus diesem Dilemma heraus wurde beispielsweise dem Item *IdS_Ga* trotz einer zu geringen Trennschärfe der Vorzug vor dem nach Kennwerten besseren Identifizierungsitem *IdS_Cb* gegeben, da das Item *IdS_Ga* anhand des gleichen Diagramms wie das nach Kennwerten auszuwählende Item *IdS_Gb* bearbeitet werden konnte.

Hinsichtlich der Diagramme *Maus* und *LäuseKäfer* wurde jeweils ein Item gegenüber einem anderen bevorzugt, um die Abdeckung aller im Kompetenzstrukturmodell formulierten Aspekte der Informationsentnahme zu erreichen (s. Tabelle 12.2). Da die Trennschärfe der ausgewählten Items allerdings zu gering war, wurden die Distraktoren überarbeitet, um ihre Kennwerte im Haupttest möglichst zu verbessern. Zudem wurde das Item *A_Hu_a* aufgenommen, da es den direkten inhaltlichen Vergleich zu Item *A_Hu_b* ermöglicht. Ebenfalls aus inhaltlichen Gründen wurden die Items *A_P2c3*, *A_P2c4* und *A_RB_d* die bereits im ersten Schritt von den Itemanalysen ausgeschlossen worden waren (vgl. 12.4.1.1), unter Verschärfung der Bewertung aufgenommen, insbesondere da sie unter minimalem Mehraufwand an Zeit zu bearbeiten waren. Aus den gleichen Gründen wurde die Sammlung zudem um das neu konstruierte Item *A_Re* erweitert, das anhand des Diagramms „Rentiere“ zu bearbeiten war.

Konstruktion

Von den Items, die der Konstruktion zugeordnet wurden, zeigten die zwei Items *K_Pr_T* und *K_Pr_A* eine signifikante Nicht-Passung (vgl. Tabelle 12.3). Vier Items wiesen eine zu geringe Schwierigkeit auf. Da die Konstruktion im Vergleich zur Informationsentnahme einige ordinale Items umfasst, wird in der Tabelle 12.3 τ_1 als Abweichung der Schwellenparameter vom Schwellenmittelwert (vgl. 5.1.2.2) sowie deren Standardfehler SE_{τ} aufgeführt. Einige Items zeigten durch τ_1 nicht geordnete Schwellenparameter an, jedoch nur in drei Fällen (Items *K_PW_S*, *K_Ki_BC*, *K_Ki_S*) mit zu geringem Standardfehler. Diese sind in der Tabelle 12.3 in der Spalte τ_1 mit * markiert. Die Items *K_FH_A*, *K_Aa_A*, *K_PW_T*, *K_PW_A*, *K_Ak_S* und *KS_O_S* wurden aufgrund ihrer zu

geringen Trennschärfe ausgeschlossen, kein Item wies eine zu hohe signifikante PBK einer Null-Rubrik auf.

Zahlreiche Items der Konstruktion konnten in die engere Auswahl genommen werden, wie in Tabelle 12.3 in der Spalte „selektiert“ zu erkennen ist. Wie auch bei der Informationsentnahme wurde bei der Endauswahl eine möglichst hohe Anzahl an Items bei eingeschränkter Anzahl an Aufgaben und damit verschiedenen Diagrammen (zur Reduzierung der Testlänge) angestrebt. Dabei wurde den Aufgaben, bei denen die Konstruktion eines vollständigen Diagramms gefordert war, im Vergleich zu denen, bei denen nur ein Diagrammteil zu erstellen war, der Vorzug gegeben, da die vollständige Konstruktion den Realanforderungen im Unterricht näher kommt.

Über die Items hinweg wiesen die Diagramme *Rentiere*, *Kiefernspinner*, *Geschlechtsreife* und *Blütendurchmesser* insgesamt die besten Werte auf: sie beinhalten lediglich zwei kritische Items, deren kritische Werte sich auf die Ordnung der Schwellen beziehen, und ein Item, das durch eine zu geringe Itemschwierigkeit ausgeschlossen worden war. Dies wurde als vertretbar angesehen, da keine entsprechende Aufgabe ohne ein solches Item vorhanden war.

Es ist anzumerken, dass im Zuge des Ausschlusses vollständiger Konstruktionsaufgaben ein Item zum Zeichnen einer Legende (K_{Pr}_L) ausgemustert wurde. Ein weiteres Item (K_{Sc}_L) wurde aus ökonomischen Gründen nicht in den Haupttest übernommen, da es als einziges anhand eines einzelnen Diagramms hätte bearbeitet werden müssen. Aufgrund einer zu geringen Itemzahl konnte das Zeichnen einer Legende als Aspekt der Konstruktion in der Haupttestung nicht untersucht werden. Im Folgenden bezieht sich daher die Konstruktion auf die verbleibenden sechs Aspekte Auswahl des Diagrammtyps, Achsenbelegung, Achsenbeschriftung, Skalierung, Werteintrag und Skizzierung (vgl. Tabelle 3.1).

Integration

Mehrere Items, die der Integration zugeordnet wurden, erhielten in der Pilotierung gute Itemkennwerte, wie durch die starke Besetzung in der Spalte „selektiert“ in der Tabelle 12.4 angedeutet wird. Lediglich zwei Integrationsitems zeigten signifikant eine Nicht-Passung, die in beiden Fällen allerdings knapp signifikant war. Das Item I_{P2c} wurde deshalb trotz dieses Ausschlusswertes ausgewählt, da es mit geringem Mehraufwand an Zeit anhand eines für die Informationsentnahme ausgewählten Diagramms zu bearbeiten war. Eine zu hohe Schwierigkeit führte zum Ausschluss des Items I_{Bak} , aufgrund einer zu geringen PBK wurden die Items I_{P3}_L und I_{GW}_L ausgeschlossen. Ein signifikanter PBK_0 -Wert von 0,39 führte zum Ausschluss des Items I_{GW}_W (vgl. Tabelle 12.4). Die Auswahl zur Testkürzung erfolgte somit hauptsächlich nach inhaltlichen Kriterien. Es

wurde dabei eine Verteilung der Items auf die zwei in Abschnitt 3.6.3 beschriebenen Aspekte der Integration (*Resultat Text* und *Resultat Diagramm*) sowie eine ausgewogene Verteilung der jeweiligen Anforderungen hinsichtlich Informationsentnahme und Konstruktion angestrebt.

Dokumentation der Itemkennwerte für die Haupttestung

In die der Haupttestung folgenden Analysen wurden die Items zum Diagramm *Räuber-Beute* (*A_RB_d*, *I_RB_a*, *I_RB_d* und *I_RB_e*) nicht einbezogen, da sie von weniger als 50 % der Probanden bearbeitet wurden. Es verbleiben damit 22 Items zur Informationsentnahme, 18 Items zur Konstruktion und sieben Items zur Integration für die Analysen zur Bearbeitung der Forschungsfragen. Die Kennwerte dieser Items in der Haupttestung werden im Anhang (s. 12.4.3) dokumentiert.

5.3.3 Bearbeitung der Forschungsfragen

Nachfolgend wird vorgestellt, welche Auswertungsmethoden zur Bearbeitung welcher Forschungsfragen eingesetzt werden. Auf die Methodik der Modellüberprüfung (vgl. 4.1) wird dabei vertieft eingegangen (5.3.3.1). Hinsichtlich der differenzierten Beschreibung der Modellstruktur (vgl. 4.2) werden die Regressions- und Mediationsanalyse (5.3.3.2) ausführlicher dargestellt, da sie für Auswertungen zentral sind. Die restlichen, gängigen Auswertungsverfahren werden ohne vertiefende Betrachtung lediglich mit dem Hinweis auf die einschlägige Literatur aufgeführt.

5.3.3.1 Überprüfung der Modellstruktur

Die Modellstruktur wird zum einen hinsichtlich der strukturellen Komponenten, zum anderen hinsichtlich der Differenzierung nach Diagrammtypen untersucht.

Komponenten

In den Hypothesen 1a und 1b (vgl. 4.1.1) wurden zwei Strukturen der Diagrammkompetenz formuliert. Beide Strukturen werden jeweils durch ein mehrdimensionales Rasch-Modell (vgl. 5.1.2.4) repräsentiert und die entsprechenden Modelle in Anlehnung an die Namensgebung der Experimental- und Kontrollansätze beim Experimentieren als *Experimentalmodelle* bezeichnet. Des Weiteren werden in Form verschiedener *Kontrollmodelle* theoretisch begründete alternative Strukturen ausgedrückt. Nachfolgend werden zuerst die Experimentalmodelle und im Anschluss die Kontrollmodelle vorgestellt.

Experimentalmodelle:**3dim-Modell**

Identifizierung	Aufbau
Ablesen	Dateneintrag
Integration	

Gemäß dem in Abschnitt 3.6 beschriebenen Kompetenzstrukturmodell wird in diesem Modell angenommen, dass die Diagrammkompetenz aus den drei Komponenten Informationsentnahme (Identifizierung und Ablesen umfassend), Konstruktion (Aufbau und Dateneintrag umfassend) sowie Integration besteht, denen sich die Items ohne Überschneidung zuordnen lassen (vgl. Hypothese 1a).

5dim-Modell

Identifizierung	Aufbau
Ablesen	Dateneintrag
Integration	

Dieses Modell beschreibt die Diagrammkompetenz als aus fünf Komponenten bestehend. Es nimmt gegenüber dem 3dim-Modell eine feinere Unterteilung der Komponente Informationsentnahme in Identifizierung und Ablesen sowie der Komponente Konstruktion in Aufbau und Dateneintrag vor und ist damit ein Untermodell des 3dim-Modells (vgl. Hypothese 1b).

Kontrollmodelle:**1dim-Modell**

Identifizierung	Aufbau
Ablesen	Dateneintrag
Integration	

In diesem Modell wird angenommen, dass sich beim Diagrammgebrauch keinerlei einzelne Fähigkeiten differenzieren lassen.

2dim-Modell

Identifizierung	Aufbau
Ablesen	Dateneintrag
Integration	

Gemäß dem grundlegenden Modell von Schnotz (2001) sind an der Integration sowohl Modellinspektions- als auch -konstruktionsprozesse beteiligt. Entsprechend könnte auch vermutet werden, dass die Integration keine eigenständige Fähigkeit darstellt. Dieses Modell postuliert deshalb zwei Komponenten der Diagrammkompetenz, die Informationsentnahme und die Konstruktion. Es wird angenommen, dass sich die Items zu zwei disjunkten Skalen anordnen lassen, wobei die Items zur Integration je nach Art der resultierenden externen Repräsentation (*Resultat Text* oder *Resultat Diagramm*, vgl. 3.6.3) einer der beiden Komponenten zugeteilt werden.

2dim_within-Modell

Identifizierung	Aufbau
Ablesen	Dateneintrag
Integration	

Das Modell von Schnotz (vgl. 3.5.1) ließe weiterhin die Vermutung zu, dass bei einer Integrationsleistung sowohl Fähigkeiten zur Informationsentnahme als auch zur Konstruktion zum Tragen kommen. In diesem Modell werden diese beiden Komponenten Informationsentnahme und Konstruktion als Teile der Diagrammkompetenz angenommen, dabei aber anders als im 2dim-Modell die Items zur Integration jeweils beiden Komponenten zugeordnet. Es liegt von der Struktur her ein „*within*“-Modell (vgl. 5.1.2.4) vor.

2dim_asym-Modell

Identifizierung	Aufbau
Ablesen	Dateneintrag
Integration	

Weiterhin ist denkbar, dass die Fähigkeit zum Integrieren von Informationen zur Fähigkeit der Informationsentnahme gehört, da beim Vorliegen eines Diagramms vor dem Integrieren immer Informationen entnommen werden müssen. Anders als im Modell 2dim_within wird in diesem Modell die Integration vollständig der Informationsentnahme zugeordnet. Damit werden die Items zur Informationsentnahme und zur Integration in der nebenstehenden schematischen Modelldarstellung zu einer asymmetrischen Skala zusammengefasst und das Modell im Folgenden als 2dim_asym-Modell bezeichnet. Die Items zur Konstruktion bilden die zweite Komponente.

3dim_q-Modell

Identifizierung	Aufbau
Ablesen	Dateneintrag
Integration	

Aufgrund der inhaltlichen Bezüge kann angenommen werden, dass sich die Fähigkeiten, die sich auf den Diagrammrahmen beziehen, von den Fähigkeiten abgrenzen lassen, die direkt mit den im Diagramm enthaltenen Daten zu tun haben. Dementsprechend nimmt dieses Modell eine Komponente *Rahmen* an, die auf der Seite der Informationsentnahme die Items zur Identifizierung und auf der Seite der Konstruktion die Items zum Aufbau umfasst, sowie eine weitere Komponente *Werte*, die die Items zum Ablesen auf der Seite der Informationsentnahme und zum Dateneintrag auf der Seite der Konstruktion beinhaltet. Die dritte Komponente ist wie im Modell *3dim* die Integration. *Rahmen* und *Werte* liegen in diesem Modell *quer* zu den im *3dim*-Modell postulierten Komponenten, weshalb die Benennung *3dim_q*-Modell gewählt wurde.

3dim_z-Modell

Identifizierung	Aufbau
Ablesen	Dateneintrag
Inte-	gration

Das letzte Modell fungiert als „technischer Kontrollansatz“. Dabei soll überprüft werden, ob eine bessere Modellpassung allein durch eine Einteilung der Items in drei Gruppen und damit durch eine Erhöhung der Parameterzahl erreicht werden kann (vgl. Wu & Adams, 2006). In diesem Modell werden ebenfalls drei Komponenten angenommen, wobei sämtliche Items *zufällig*, entsprechend die Benennung mit *3dim_z*, auf diese drei Skalen verteilt werden, was in nebenstehender Abbildung durch die uneindeutigen Positionen der Worte angedeutet werden soll. Da die zufällige Einteilung der Items auf drei Skalen keine inhaltliche Struktur widerspiegelt, wird erwartet, dass dieser Ansatz die schlechteste Modellpassung aufweist.

In den hier aufgeführten Modellen werden verschiedene mögliche Strukturen der Diagrammkompetenz formuliert. Die einzelnen Teile eines jeweiligen Modells werden im Folgenden grundsätzlich als *Komponenten* bezeichnet (vgl. 3.4.2). Demzufolge wird von den *Komponenten* des *3dim*-Modells und ebenso von den *Komponenten* des *5dim*-Modells gesprochen.

Alle Komponenten des Modells *5dim* sind zueinander strukturell gleichwertig. Im inhaltlichen Bezug auf die Komponenten des *3dim*-Modells stellen sie jedoch, wie oben erwähnt, feinere Unterteilungen dar. In dem besonderen Fall, dass die Strukturen des *5dim*-Modells in Bezug zu den Strukturen des *3dim*-Modells gesetzt werden, werden einzelne *Komponenten* des *5dim*-Modells als *Unterkomponenten* bezeichnet. Identifizierung und

Ablesen sind folglich Komponenten des Modells *5dim*, aber Unterkomponenten der Komponente Informationsentnahme des *3dim*-Modells.

Zur Überprüfung der Forschungshypothesen und damit der Struktur des Kompetenzmodells wurden alle beschriebenen Modelle berechnet und miteinander verglichen. Gemäß den Hypothesen wurde erwartet, dass die beiden Experimentalmodelle den Kontrollmodellen überlegen sind. Zur Bestätigung der Hypothesen musste jedes der beiden Experimentalmodelle die folgenden vier Kriterien erfüllen.

1. Das Modell sollte im Vergleich mit den Kontrollmodellen die beste Modellpassung aufweisen.
2. Die Reliabilitäten der einzelnen Komponenten sollten zufrieden stellend sein.
3. Die Personenhomogenität sollte hinsichtlich der einzelnen Komponenten gewährleistet sein.
4. Die einzelnen Komponenten sollten einen Beitrag zur Aufklärung der Gesamtvarianz leisten.

Wie die einzelnen Kriterien operationalisiert wurden, ist in den folgenden Abschnitten beschrieben.

Modellpassung. Grundsätzlich können Daten mit einem Modell, das sehr viele Modellparameter umfasst, immer exakter beschrieben werden, als mit einem Modell, das nur sehr wenige Parameter enthält. Ziel einer wissenschaftlichen Modellierung ist es allerdings, mit *möglichst wenigen* Annahmen (Parametern) einen Sachverhalt *möglichst gut* zu beschreiben⁵⁷ (vgl. Rost, 1996, S. 324). Beim Vergleich verschiedener Modelle muss deshalb zur Beurteilung der Modellgüte abgewogen werden, wie gut ein Modell auf die Daten passt und mit wie vielen Parametern diese Passung erzielt wurde. Zu diesem Zweck werden in der vorliegenden Arbeit zwei informationstheoretische Indices herangezogen, die sowohl die Loglikelihood⁵⁸ als Anpassungsmaß als auch die Zahl der Modellparameter n_p berücksichtigen. Diese sind eine besondere Form des *Akaike's Information Criterion (AIC)*, namentlich der *AIC_c*-Index, und das *Best Information Criterion (BIC)*, deren Berechnungsformeln in der Abbildung 5.15 aufgeführt sind.

⁵⁷ Dies wird auch als „Einfachheitskriterium“ (Rost, 1996) oder „Sparsamkeitskriterium“ bezeichnet.

⁵⁸ Die Likelihood-Funktion L bezeichnet in der probabilistischen Testtheorie die gemeinsame Wahrscheinlichkeitsfunktion der Itemantworten unter der Bedingung der unbekannt Parameter eines Testmodells. In anderen Worten, sie „beschreibt die Wahrscheinlichkeit der Daten in Abhängigkeit von den Modellparametern unter der Annahme, dass das Modell gilt“ (Rost, 1996, S. 117). Sie ist ein Maß für die Passung eines Modells auf die Daten: je höher die Likelihood, desto besser passt das Modell. Meist werden allerdings die logarithmierten Likelihoodwerte (die Loglikelihoods, $\log L$) berichtet, da die unlogarithmierten Werte sehr klein sind (Rost, 1996, S. 326).

$$AIC_c = -2 \log L + 2n_p + \frac{2n_p(n_p + 1)}{N - n_{p-1} - 1} \quad BIC = -2 \log L + (\log N)n_p$$

Abbildung 5.15: Formeln für die Indices zum Modellvergleich. L = Likelihood, n_p = Zahl der Modellparameter (vgl. 5.1.2.4), N = Stichprobengröße.

Sowohl für AIC_c als auch für BIC gilt:

je geringer der Index, desto besser passt das Modell auf die Daten.

Die Differenz zwischen den Kennwerten der verschiedenen Modelle ist dabei entscheidend für eine bessere oder schlechtere Passung: In Anlehnung an Burnham und Anderson (2004) wird in der vorliegenden Arbeit ein Modell als *bedeutsam besser oder schlechter passend* als ein anderes interpretiert, wenn die Differenz der Indices größer oder gleich zehn ist.

Der AIC_c wird von Burnham und Anderson (2004) insbesondere dann empfohlen, wenn, wie in der vorliegenden Arbeit, das Verhältnis von Stichprobengröße und größter Parameterzahl (aller miteinander verglichenen Modelle) geringer als 40 ist⁵⁹. Da sich aufgrund des geringeren Faktors die Parameterzahl im AIC_c nicht so stark auswirkt wie im BIC , kann ein komplexeres Modell dabei nach dem AIC_c eine bessere, nach dem BIC eine schlechtere Passung aufweisen als ein sparsameres. Um dem Sparsamkeitskriterium in höherem Maße nachzukommen, wird in der vorliegenden Arbeit der BIC -Index höher bewertet als der AIC_c , wenn eine zwar bedeutsame, aber dennoch geringe bessere Passung eines Modells nach dem AIC_c durch eine unverhältnismäßige Erhöhung der Parameterzahl erzielt wird. In diesem Fall wird das sparsamere Modell vorgezogen.

Reliabilitäten. Die Überprüfung der Reliabilitäten dient der Absicherung der Modellstruktur dahingehend, dass die Komponenten auf reliablen Skalen basieren sollen, d. h. dass sichergestellt sein soll, dass alle Items einer Skala zuverlässig die gleiche Fähigkeit ansprechen und damit die Skala als eine Komponente sinnvoll interpretierbar ist.

In der vorliegenden Arbeit wird die Reliabilität dabei für die Personenparameter (WLEs) nach der Erwartungswertmethode (Rost, 2004, S. 380) berechnet, die im Rahmen der probabilistischen Auswertungsverfahren häufig eingesetzt wird und von ConQuest als *WLE person separation reliability* (nachfolgend als WLE-Reliabilität bezeichnet) ausgegeben wird. Da diese Methode die Reliabilitäten eher unterschätzt, insbesondere dann, wenn sich viele Personen mit Extremscores in der Stichprobe befinden (Rost, 2004, S. 283), wird zum Vergleich ein weiteres Reliabilitätsmaß zu Rate gezogen: die EAP/PV-Reliabilität (für nähere Informationen siehe Walter, 2005, und Rost, 2004 unter Bayes-

⁵⁹ Für $N \rightarrow \infty$ strebt dabei AIC_c gegen AIC .

Reliabilität). Da dieses Reliabilitätsmaß eine gute Übereinstimmung mit Cronbachs α aufweist (Rost, 2004, S. 382), kann im Falle einer starken Abweichung der beiden Maße auf die EAP/PV-Reliabilität zurückgegriffen und diese stärker gewichtet werden.

Personenhomogenität. Die Prüfung der Personenhomogenität gibt Aufschluss darüber, ob die Personen alle Items einer Skala aufgrund der gleichen Fähigkeit auf eine bestimmte Art bearbeiten, anders gesagt, ob die Skala bei allen Personen gleich misst. Dazu wird von Rost (1996, S. 342) ein graphischer Modelltest vorgeschlagen, bei dem die gesamte Stichprobe in zwei Gruppen geteilt und für jede dieser Gruppen eine eindimensionale Rasch-Skalierung durchgeführt wird. Falls Personenhomogenität gewährleistet ist, sollte bei Auftragung der Itemschwierigkeiten beider Teilstichproben in einem Streudiagramm eine Gerade mit Steigung 1 erkennbar sein, wie in Abbildung 5.16 für die zwei Teilstichproben 2A und 2B dargestellt wird.

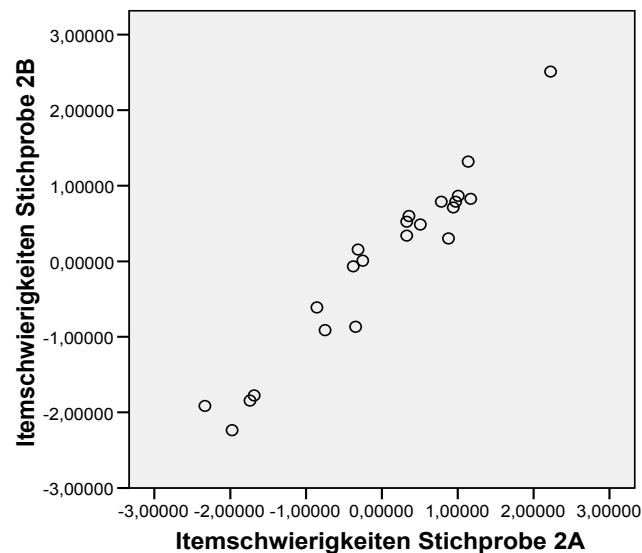


Abbildung 5.16: Beispiel für den graphischen Modelltest auf Personenhomogenität. Die Stichproben 2A und 2B bezeichnen zwei Teilstichproben einer zufälligen Zweiteilung einer fiktiven Gesamtstichprobe. Im dargestellten Fall ist Personenhomogenität gegeben.

In der vorliegenden Arbeit werden für jede Komponente drei zufällige Zweiteilungen der Stichprobe sowie eine Teilung nach Klassenstufe und eine nach Geschlecht vorgenommen.

Alternativ zur graphischen Auftragung kann der Zusammenhang zwischen den Itemschwierigkeiten beider Teilstichproben durch eine Korrelationsanalyse (Maßkorrelation nach Pearson, s. Backhaus, Erichson, Plinke & Weiber, 2003) ermittelt werden. Entsprechend der Steigung der Gerade sollen die Itemschwierigkeiten beider Teilstichproben jeweils mit mindestens $r = .9$ korrelieren. Aufgrund der insgesamt fünf Berechnungen pro Komponente wird diesem Verfahren in der vorliegenden Arbeit zur besseren Übersicht der Vorzug vor der graphischen Auswertung gegeben.

Vorhersagewerte der Komponenten. Durch die Struktur eines untersuchten Modells wird eine Zerlegung der Diagrammkompetenz in verschiedene Komponenten postuliert. Als letztes Kriterium zur Validierung einer bestimmten Modellstruktur wird überprüft, ob diese Komponenten einen Beitrag zur Messung der Diagrammkompetenz insgesamt leisten. Dazu wird mittels einer schrittweisen Regression analysiert (vgl. 5.3.3.2, Relationen der Komponenten zueinander), ob sich die Personenparameter der verschiedenen Komponenten als signifikante Vorhersagewerte für den Personenparameter des eindimensionalen Modells identifizieren lassen. Damit wird sichergestellt, dass die Komponenten strukturell wirklich bedeutsam sind. Dieses Kriterium wurde in Anlehnung an vergleichbare Studien (z. B. Senkbeil, Rost, Carstensen & Walter, 2005) formuliert, in denen in korrelationsstatistischen Analysen die Aufklärung der Varianz durch eine Komponente als Hinweis auf deren Bedeutsamkeit herangezogen wird. Dabei wird allerdings für jede Komponente einzeln die Korrelation zwischen dem WLE der Komponente und dem WLE des eindimensionalen Modells berechnet, so dass die Zusammenhänge einzelner Komponenten untereinander unberücksichtigt bleiben. Durch die schrittweise Regression werden diese Zusammenhänge jedoch herausgerechnet und damit die unabhängigen Zusammenhänge zwischen den einzelnen Komponenten und dem WLE des eindimensionalen Modells offen gelegt, weshalb diese Betrachtung in der vorliegenden Arbeit bevorzugt wird.

Diagrammtypen

In den vorangehend beschriebenen Analysen wird untersucht, ob sich die drei respektive fünf Komponenten der Diagrammkompetenz, wie sie im theoretisch abgeleiteten Modell postuliert werden, empirisch bestätigen lassen. Diese Komponenten beziehen sich sowohl auf Linien- als auch auf Säulendiagramme. Um zu überprüfen, ob diese umfassende Struktur durch eine andere Struktur ersetzt oder erweitert werden muss, die eine Differenzierung der Diagrammtypen vorsieht, werden drei mehrdimensionale Rasch-Modelle getestet, die nachfolgend vorgestellt werden. Sie werden entsprechend den oben genannten Kriterien begutachtet und mit dem Modell verglichen, das durch die Analysen zu den Komponenten (s. o.) als bestpassend ausgewiesen wird.

Das erste Modell nimmt eine Unterteilung nach den Diagrammtypen (DT) vor und postuliert keine weiteren Strukturen. Es hat damit zwei Rasch-Dimensionen und wird im Folgenden als *2dimDT* bezeichnet.

Ausgehend vom theoretischen Strukturmodell der Diagrammkompetenz nehmen die weiteren zwei Modelle eine Strukturierung nach Diagrammtypen vor und greifen zudem die drei respektive fünf Komponenten des Strukturmodells auf. Die Rasch-Modelle, die

diese Struktur repräsentieren, sind *within-item*-Modelle (vgl. 5.1.2.4). Jedes Item wird dabei einer der (Unter-)Komponenten der Diagrammkompetenz sowie dem Diagrammtyp zugewiesen. Zur Veranschaulichung ist in der Abbildung 5.17 das entsprechende Modell für die drei Komponenten Informationsentnahme, Konstruktion und Integration dargestellt. Bezugnehmend auf die Namensgebung in der Modelle zur Analyse der Komponenten der Diagrammkompetenz (vgl. 5.3.3.1, Komponenten) wird es als *2dimDTx3dim*-Modell bezeichnet. Entsprechend wird das andere Modell mit *2dimDTx5dim*-Modell benannt.

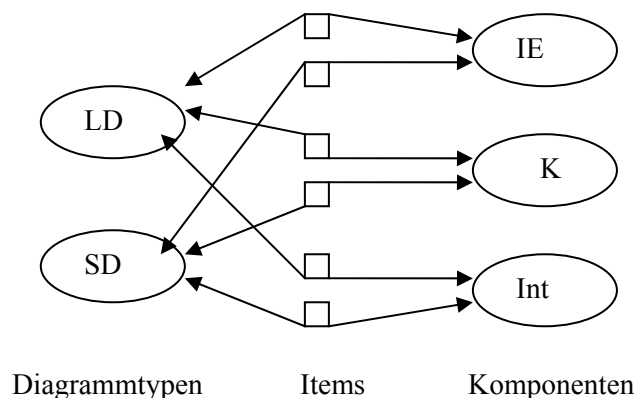


Abbildung 5.17: Schematische Darstellung des *within-item*-Modells⁶⁰ *2dimDTx3dim*, das drei strukturelle Komponenten Informationsentnahme (IE), Konstruktion (K) und Integration (Int) der Diagrammkompetenz sowie die Diagrammtypen Liniendiagramm (LD) und Säulendiagramm (SD) unterscheidet. Ein Item wird jeweils einem Diagrammtyp und einer Kompetenzkomponente zugeordnet.

5.3.3.2 Differenzierte Beschreibung der Modellstruktur

Gemäß den in Kapitel 4 dargelegten Forschungsfragen werden im Folgenden die Auswertungsmethoden in vier Abschnitten dargestellt. Im ersten Abschnitt wird aufgeführt, mit welchen Methoden die Relationen zwischen den einzelnen Komponenten und Unterkomponenten untersucht werden (vgl. 4.2.1). Im darauffolgenden Abschnitt wird dargelegt, wie die (Unter-)Komponenten hinsichtlich ihrer Schwierigkeiten miteinander verglichen werden (vgl. 4.2.2). Daran anschließend wird die Analyse der Binnenstrukturen einzelner Komponenten beschrieben (vgl. 4.2.3), bevor der Vergleich der Schwierigkeiten bezogen auf die Diagrammtypen aufgeführt wird (vgl. 4.2.4).

Relationen der Komponenten zueinander

Zur Analyse der Modellstruktur werden in einem ersten Schritt die Zusammenhänge zwischen den einzelnen Komponenten untersucht und dazu Korrelations-, Regressions- und Mediationsanalysen durchgeführt.

⁶⁰ Die Anordnung der verschiedenen Dimensionen des Modells auf beiden Seiten der Items dient der übersichtlicheren Darstellung und weicht deshalb von der Darstellung in Abschnitt 5.1.2.4 ab.

Grundlage für die hier vorgestellten Analysen sind die Personenparameter für die jeweiligen Komponenten, die bei der Rasch-Modellierung geschätzt wurden. Um die Zusammenhänge für jeweils zwei Komponenten aufzudecken, werden **Korrelationsanalysen** (Maßkorrelation nach Pearson, für detailliertere Informationen siehe Backhaus et al., 2003; Janssen & Laatz, 2005) berechnet. Die als Maß der Stärke des Zusammenhangs genutzten Korrelationskoeffizienten werden mit r betitelt (vgl. „*Pearsons r* “, Backhaus et al., 2003). Mithilfe von **Regressionsanalysen** (vgl. Backhaus et al., 2003) werden Zusammenhänge zwischen mehreren Komponenten untersucht. Im Folgenden werden die zur Berichterstattung genutzten Kennwerte dargelegt.

In linearen Regressionsanalysen wird der Zusammenhang verschiedener unabhängiger Variablen (**Prädiktoren**) und einer abhängigen Variablen in Form einer linearen Funktion modelliert, die Stärke des Zusammenhangs geben dabei die *Regressionskoeffizienten* an. Bei der Berichterstattung wird der Test auf globale Passung der Gleichung angegeben, bei dem über eine **F-Statistik** die Abweichungen der Modellgleichung von der „Nullfunktion“, bei der sämtliche Regressionskoeffizienten null sind, analysiert werden. Der F-Test kann dabei interpretiert werden als Signifikanzprüfung, ob die angenommenen unabhängigen Variablen einen Beitrag zur Erklärung der Varianz leisten (Janssen & Laatz, 2005, S. 419). Zudem wird das **korrigierte Bestimmtheitsmaß R^2_{kor}** herangezogen, das angibt, wie gut die Modellfunktion die Daten erklärt, d. h. wie hoch der Anteil der durch die Modellfunktion aufgeklärten Varianz an der Gesamtvarianz der abhängigen Variable ist. Insbesondere steht das Bestimmtheitsmaß in Beziehung zu der Effektgröße ϵ^2 , die nach Cohen (1988, zitiert in Bortz, 1999, S. 449) zur Unterscheidung von schwachen, mittleren und starken Effekten genutzt werden kann. Diese Klassifikation kann auf das Bestimmtheitsmaß übertragen und letzteres analog interpretiert werden. Ein schwacher Effekt ist demnach durch $0,0196 < R^2_{\text{kor}} < 0,1304$, ein mittlerer Effekt durch $0,1304 < R^2_{\text{kor}} < 0,2593$ und ein starker Effekt durch $0,2593 < R^2_{\text{kor}}$ gekennzeichnet (Bortz, 1999, S. 449).

Bei Signifikanz des F-Wertes und damit Passung der Gleichung wird mit einem t-Test für jeden einzelnen Prädiktor überprüft, ob dessen Regressionskoeffizient sich signifikant von null unterscheidet. Ist dies der Fall, wird der Prädiktor in der vorliegenden Arbeit als **signifikanter Prädiktor** oder **signifikanter Vorhersagewert** bezeichnet. Um verschiedene Regressionskoeffizienten vergleichen zu können, werden deren standardisierte Werte, die **Beta-Koeffizienten (β)**, verwendet (Backhaus et al., 2003, S. 61). In tabellarischen Übersichten werden gemäß den APA-Standards (APA, 2001) zudem die unstandardisierten Koeffizienten (b) und deren Standardfehler (SE_b) aufgeführt.

In der vorliegenden Arbeit werden insbesondere schrittweise lineare Regressionen (für det. Informationen siehe Bortz, S. 446ff) berechnet. Dabei werden nicht automatisch alle in die Berechnung eingegebenen Variablen als Prädiktoren in die Regressionsgleichung

aufgenommen. Die Auswahl erfolgt über einen Algorithmus, bei dem in jedem Schritt eine Variable draufhin getestet wird, ob sie im Vergleich zu den bereits im Modell enthaltenen weiter zur Varianzaufklärung des Modells beiträgt und ob bei Aufnahme dieser Variable durch Kollinearität eine andere redundant wird, d. h. dann keinen Beitrag mehr zur Varianzaufklärung leistet. Letztere würde wieder aus dem Modell entfernt. Durch dieses Verfahren wird gewährleistet, dass die ins Modell aufgenommenen Variablen keinen zu hohen Grad an gemeinsamer Varianz aufweisen. Dabei werden insbesondere nur signifikante Vorhersagewerte aufgenommen und so im Prinzip die „aussagekräftigeren“ Prädiktoren identifiziert.

Falls durch die Korrelationen ein Zusammenhang zwischen zwei Komponenten X und Y angezeigt wird, sich dieser bei Aufnahme einer zusätzlichen Komponente M als Prädiktor für Y aber verringert, ist es möglich, dass der Zusammenhang von X und Y im Grunde auf M zurückzuführen ist und damit durch M vermittelt wird. In diesem Fall werden die Zusammenhänge zwischen den Komponenten durch *Mediationsanalysen* noch detaillierter beleuchtet. Dabei wird ein Koeffizient berechnet, der den *indirekten Zusammenhang* von X auf Y beschreibt, und in der vorliegenden Arbeit durch den *Sobel-Test* auf seine Signifikanz hin überprüft. Das Vorgehen bei dieser Analyse wird nachfolgend kurz vorgestellt.

Neben der Berechnung des indirekten Zusammenhangs per se muss für die Durchführung einer Mediationsanalyse überprüft werden, ob vier Voraussetzungen (1) bis (4) erfüllt sind (Baron & Kenny, 1986; Urban & Mayerl, 2006). Zu beiden Zwecken sind insgesamt *drei* Berechnungen 1 bis 3 nötig, die insgesamt *vier* Regressionskoeffizienten liefern. Es handelt sich dabei um zwei einfache Regressionen, die in der Abbildung 5.18 als violetter respektive blauer Pfeil symbolisiert sind, und eine multiple Regression, die durch die schwarzen Pfeile dargestellt wird.

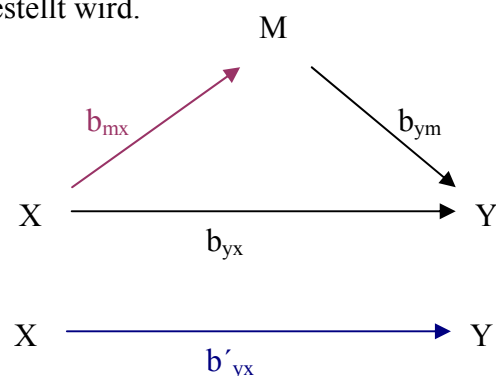
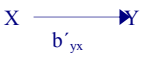
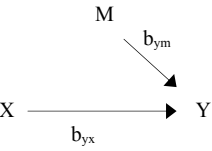
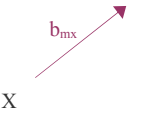


Abbildung 5.18: Zusammenhänge zwischen unabhängiger (X), medierender (M) und abhängiger Variable (Y). Zu berechnende Regressionen mit zugehörigen Koeffizienten in drei unterschiedlichen Farben.

In der Tabelle 5.5 sind die drei Berechnungen aufgeführt, die jeweils relevanten Regressionskoeffizienten hervorgehoben und die Voraussetzungen genannt, die mit diesen überprüft werden. Es sei darauf hingewiesen, dass für diese Analysen die *unstandardisierten* Regressionskoeffizienten verwendet werden.

Tabelle 5.5: Übersicht über die zu berechnenden Regressionen und zu prüfenden Voraussetzungen für eine Mediationsanalyse. * = signifikanter Koeffizient, $p < .05$ mindestens.

Berechnung mit resultierenden Koeffizienten		Zu prüfende Voraussetzung (in Worten und Symbolen)		
1	 einfache Regression mit X als Prädiktor für Y	b'_{yx}	(1) signifikanter Effekt des Prädiktors X auf Y	b'_{yx}^*
2	 multiple Regression mit X und M als Prädiktoren für Y	b_{yx}	(2) Effekt des Prädiktors X auf Y in der multiplen Regression mit M als zusätzlichem Prädiktor geringer als in der einfachen Regression	$b_{yx} < b'_{yx}$
		b_{ym}	(3) signifikanter Effekt des Prädiktors M auf Y	b_{ym}^*
3	 einfache Regression mit X als Prädiktor für M	b_{mx}	(4) signifikanter Effekt des Prädiktors M auf Y	b_{mx}^*

Zur Bestimmung des Koeffizienten $b_{mx \cdot ym}$ des indirekten Zusammenhangs zwischen X und Y werden nach Urban und Mayerl (2006) die zwei Werte b_{ym} und b_{mx} benötigt. Der indirekte Zusammenhang berechnet sich dann durch

$$b_{mx \cdot ym} = b_{mx} \cdot b_{ym}$$

Die Signifikanzprüfung dieses Koeffizienten erfolgt durch einen t-Test. Für diesen wird der Standardfehler des Koeffizienten bestimmt, der sich nach Sobel⁶¹ (1982, zitiert in Urban & Mayerl, 2006) wie folgt zusammensetzt:

$$SE_{b_{mx \cdot ym}} = \sqrt{b_{mx}^2 \times SE_{b_{ym}}^2 + b_{ym}^2 \times SE_{b_{mx}}^2}$$

Die jeweiligen Standardfehler $SE(b_{ym})$ und $SE(b_{mx})$ der Regressionskoeffizienten ergeben sich bei den zugehörigen Regressionsanalysen. Der t-Wert wird als Quotient aus dem Koeffizienten des indirekten Zusammenhangs $b_{mx \cdot ym}$ und seinem Standardfehler berechnet.

⁶¹ Grundsätzlich wären auch andere Formeln zur Berechnung der genannten Standardfehler möglich, wie sie beispielsweise von Goodman (1960, zitiert in Urban & Mayerl, 2006) vorgeschlagen wurden.

Wird der t-Test signifikant, so wird M als Mediator des Zusammenhangs zwischen X und Y bezeichnet.

Ein **totaler Mediator-Effekt** liegt vor, wenn in der multiplen Regression kein signifikanter Zusammenhang zwischen X und Y ermittelt wird, d. h. der Regressionskoeffizient b_{yx} nicht signifikant ist. Ist noch eine (geringe) signifikante Korrelation zwischen X und Y festzustellen, so spricht man von einem **partiellen Mediator-Effekt** (Urban & Mayerl, 2006).

Zusätzlich zu den letztgenannten Analysen kann die Stärke des Mediator-Effekts darüber ermittelt werden, in welchem Verhältnis die Koeffizienten bestimmter Zusammenhänge zueinander stehen. In Anlehnung an Urban & Mayerl (2006) wird dabei der Zusammenhang, der in Regression 1 zwischen X und Y ermittelt wird, als **totaler Zusammenhang** bezeichnet. Zur Bestimmung der Stärke des Mediator-Effekts wird der Koeffizient des indirekten Zusammenhangs $b_{mx\cdot ym}$ zum Koeffizienten des totalen Zusammenhang b'_{yx} in Beziehung gesetzt (vgl. Urban & Mayerl, 2006). Je höher der Quotient $b_{mx\cdot ym}/b'_{yx}$, desto höher ist der Anteil des Zusammenhangs zwischen X und Y, der über die Mediatorvariable vermittelt wird.

Verhältnis der Schwierigkeiten

Zur weiteren Strukturbeschreibung der Diagrammkompetenz wird mittels der deskriptiven Statistiken analysiert, in welchem Verhältnis die Schwierigkeiten der (Unter-)Komponenten zueinander stehen. Für den Vergleich werden die Itemparameter über alle Komponenten gemeinsam Rasch-skaliert und anschließend verglichen (vgl. Einhaus, 2007). Die graphische Darstellung der Schwierigkeiten erfolgt in einem Boxplot wie in der Abbildung 5.19 beispielhaft angegeben. Als abhängige Variable wird die Itemschwierigkeit auf einer Skala von -3 bis $+3$ angegeben, wobei die Items leichter sind, je geringer der Wert ist. Als zentraler Wert wird der Mittelwert eingetragen. Die obere und untere Grenze der Box zeigen das 75. und das 25. Perzentil, innerhalb der Box liegen damit 50% der Items der jeweiligen Komponente. Die „Whisker“ (vgl. Dytham, 2005, S. 47) sind bis zum Maximum und zum Minimum ausgedehnt und zeigen dadurch direkt die Spannweite der Werte an.

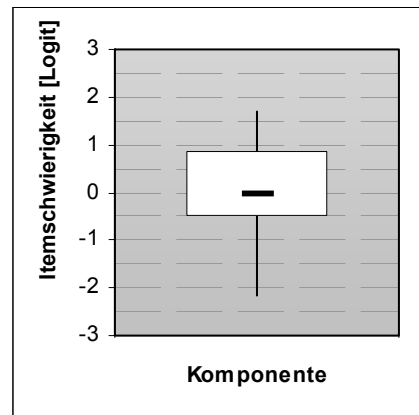


Abbildung 5.19: Beispiel für ein Boxplot.

Die Unterschiede zwischen den einzelnen (Unter-)Komponenten bezüglich der mittleren Itemschwierigkeiten werden über H-Tests nach Kruskal-Wallis auf Signifikanz untersucht. Dieses nichtparametrische Verfahren wurde gewählt, da die Komponenten unterschiedlich stark mit Items besetzt waren und infolge dessen keine Varianzhomogenität, welche Voraussetzung für das entsprechende parametrische Verfahren wäre, gewährleistet war (für weitere Informationen siehe bspw. Lücken, 2007).

Analyse der Binnenstrukturen

Zur Analyse der Binnenstrukturen werden Rasch-Modelle für die einzelnen Komponenten berechnet, die die jeweils formulierten Aspekte modellieren, und diese mit den ein-respektive zweidimensionalen Modellen der jeweiligen Komponente verglichen (analog zu 5.3.3.1). Im Fall der Informationsentnahme sind dies

- das *1dim IE* als eindimensionales Rasch-Modell, das nur die Items zur Informationsentnahme umfasst,
- das *2dim IE* als zweidimensionales Rasch-Modell, das die Unterkomponenten der Informationsentnahme (Identifizierung und Ablesen) unterscheidet,
- das *5dim IE* als fünfdimensionales Rasch-Modell, das die Identifizierung sowie die vier Ableseordnungen formuliert.

In Hinblick auf die Binnenstruktur der Konstruktion wurden folgende Modelle berechnet:

- das *1dim K* als eindimensionales Rasch-Modell, das nur die Items zur Konstruktion umfasst,
- das *2dim K* als zweidimensionales Rasch-Modell, das die Unterkomponenten der Konstruktion (Aufbau und Dateneintrag) unterscheidet,
- das *6dim K* als fünfdimensionales Rasch-Modell, das nach den sechs in der vorliegenden Arbeit untersuchten Aspekten der Konstruktion (vgl. 5.3.2.2) differenziert.

Hinsichtlich der Integration ergaben sich

- das *1dim Int* als eindimensionales Rasch-Modell der Integration und
- das *2dim Int* als zweidimensionales Rasch-Modell, das die beiden in Abschnitt 3.6.3 beschriebenen Formen der Integration nach den resultierenden externen Repräsentationsformen (*Resultat Text* und *Resultat Diagramm*) unterscheidet.

Neben der Analyse der geeignetsten Modellstruktur werden die Schwierigkeiten der Items für die jeweilig untersuchten Aspekte herangezogen. Schwierigkeitsunterschiede werden mit den entsprechenden nicht-parametrischen Verfahren (H-Test oder U-Test, s. Lücken, 2007) auf Signifikanz überprüft.

Vergleich der Schwierigkeiten bezogen auf die Diagrammtypen

Analog zur Analyse der Schwierigkeiten der Komponenten (s. o.) werden die mittleren Schwierigkeiten der Items in Abhängigkeit vom Diagrammtyp verglichen.

5.3.3.3 Lernerhintergrund und Lernermerkmale

Einfluss des Lernerhintergrunds

In Abhängigkeit der Klassenstufe respektive des Geschlechts werden (a) die Fähigkeiten der Lernenden in den einzelnen Gruppen verglichen. Darüber hinaus werden (b) die Relationen untersucht, die sich zwischen den einzelnen (Unter-)Komponenten der Diagrammkompetenz für die einzelnen Schülergruppen zeigen (vgl. 4.3.1).

a) Unterschiede in Bezug auf die Fähigkeiten je Komponente. Zur Analyse der Fähigkeitsunterschiede zwischen den beiden untersuchten Klassenstufen respektive Geschlechtern werden die Mittelwerte der Personenparameter in den einzelnen Komponenten herangezogen. Mithilfe von t-Tests (siehe bspw. Bortz, 1999) wird

untersucht, ob sich bezogen auf die Mittelwerte signifikante Unterschiede zwischen den Gruppen zeigen.

b) Unterschiede in Bezug auf die Relationen der Komponenten. In Bezug zu den Ergebnissen aus Abschnitt 6.2.1 wird beleuchtet, ob sich hinsichtlich der Relationen der Komponenten zueinander Unterschiede in Abhängigkeit von Klassenstufe und Geschlecht zeigen. Dazu werden die entsprechenden Zusammenhänge jeweils für die Teilstichproben ermittelt und die Korrelationskoeffizienten in Z-Werte transformiert (Transformation nach Fisher, vgl. Bortz, 1999, S. 209). Unter Berücksichtigung der Stichprobengröße beider Teilstichproben werden auftretende Unterschiede zwischen den Teilstichproben auf Signifikanz hin untersucht (Preacher, 2002).

Einfluss der Lernerkmale

Mithilfe von schrittweisen Regressionsanalysen (vgl. 5.3.3.2, Relationen) wird analysiert, welche der Lernerkmale biologisches Vorwissen, Umgang mit abstrakten Diagrammen, Lesekompetenz, Rechenfähigkeit und Anstrengung sich als Vorhersagewerte auf die Personenfähigkeiten in den verschiedenen Komponenten herausstellen lassen.

Die Reliabilitäten beider erstgenannter Variablen sind relativ gering, so dass für die späteren Auswertungen sowohl die Summenscores der jeweiligen Skalen als auch deren Einzelitems herangezogen wurden. Das Item *pop1* (s. 12.3.2.1) zum biologischen Vorwissen musste im Nachhinein aus den Analysen entfernt werden, da es durch im Test gegebene Informationen beantwortet werden konnte.

5.3.3.4 Qualitative Fehleranalyse

Zur Erfassung der in der vorliegenden Arbeit aufgetretenen Fehler beim Konstruieren von Diagrammen werden die relativen Häufigkeiten des Auftretens bestimmter Rubriken (vgl. 5.3.1) bestimmt. In dieser Hinsicht besitzen diese Analysen quantitativen Charakter. Der Fokus liegt allerdings weniger auf einer umfassenden Quantifizierung als auf der Aufdeckung (zum Teil exemplarischer) möglicher Fehler.

6 Ergebnisse

In diesem Kapitel werden die Ergebnisse der vorliegenden Arbeit in direktem Bezug zu den in Abschnitt 5.3.3 vorgestellten Methoden zur Bearbeitung der einzelnen Forschungsfragen vorgestellt. Zuerst werden die Ergebnisse der Analysen zur Überprüfung der im Strukturmodell der Diagrammkompetenz postulierten Kompetenzstruktur beschrieben (6.1). Im Anschluss werden die Ergebnisse der weiterführenden Analysen zur differenzierten Beschreibung der Modellstruktur aufgeführt (6.2). Daran schließt sich die Darstellung der Ergebnisse zum Einfluss des Lernerhintergrunds und der Lernermerkmale an (6.3). Die Darstellung der Ergebnisse schließt mit der Darstellung der qualitativen Analysen zu aufgetretenen Fehlern bei der Konstruktion (6.4).

6.1 Modellstruktur

Die Überprüfung der Modellstruktur erstreckt sich, wie in Abschnitt 5.3.3.1 beschrieben, auf die postulierten Komponenten der Diagrammkompetenz (6.1.1) sowie auf die zwei untersuchten Diagrammtypen Linien- und Säulendiagramm (6.1.2). Es wurden verschiedene Testmodelle berechnet, die jeweils andere Komponenten der Diagrammkompetenz postulierten, respektive die Diagrammtypen zusammenfassten oder differenzierten. Die Ergebnisse dieser Berechnungen sowie die Vergleiche der einzelnen Modelle werden nachfolgend beschrieben.

6.1.1 Komponenten

Die zur Überprüfung der Modellstruktur eingesetzten Modelle wurden in Abschnitt 5.3.3.1 (s. Komponenten) ausführlich vorgestellt. An dieser Stelle werden sie deshalb nur kurz aufgegriffen.

Das Modell *3dim* repräsentiert die im Strukturmodell der Diagrammkompetenz beschriebenen drei Komponenten Informationsentnahme, Konstruktion und Integration. Das *5dim*-Modell verfeinert die im *3dim*-Modell postulierte Struktur, indem die Informationsentnahme in die Unterkomponenten Identifizierung und Ablesen sowie die Konstruktion in die Unterkomponenten Aufbau und Dateneintrag unterteilt wird; insgesamt umfasst es somit insgesamt fünf Komponenten. Beide genannten Modelle greifen die im Modell der Diagrammkompetenz beschriebene Struktur auf und werden deshalb als Experimentalmodelle bezeichnet.

Die Experimentalmodelle werden mit den Kontrollmodellen verglichen. Das erste Kontrollmodell ist das *1dim*-Modell, welches keine Einzelkomponenten der Diagrammkompetenz annimmt. Im Modell *2dim* werden lediglich die beiden Komponenten Informationsentnahme und Konstruktion differenziert. Das Modell *2dim_within* nimmt darüber hinaus die Existenz der Komponente Integration an,

welche jedoch als eine Kombination aus Informationsentnahme und Konstruktion aufgefasst wird. Im Modell *2dim_asym* wird angenommen, dass Informationsentnahme und Integration eine Komponente bilden, die der Konstruktion gegenübersteht. Im *3dim_q*-Modell wird angenommen, dass sich die Fähigkeiten zur Integration, zum Lesen als auch Zeichnen eines Diagrammrahmens (Komponente *Rahmen*) sowie zum Ablesen als auch Eintragen von Daten (Komponente *Werte*) unterscheiden lassen. Das Modell *3dim_z* ist ein technisches Kontrollmodell, das drei zufällig eingeteilte und damit nicht inhaltlich begründete Komponenten unterscheidet.

Wie in Abschnitt 5.3.3.1 dargelegt, werden zur Überprüfung der durch die Hypothesen 1a und 1b vermuteten Passung der *Experimentalmodelle* vier Kriterien herangezogen. Die Ergebnisse der einzelnen Berechnungen zu diesen Kriterien werden im Folgenden dargestellt.

6.1.1.1 Modellpassung

Als Kennwert für die Passung eines Testmodells auf die empirischen Daten werden zwei Indices herangezogen. Für beide Indices (AIC_c und BIC) gilt dabei: je geringer der Wert, desto höher die Wahrscheinlichkeit, dass das jeweilige Modell zur Beschreibung der Daten geeignet ist (vgl. 5.3.3.1, Komponenten: Modellpassung). In den nachfolgend beschriebenen Analysen verhalten sich beide Maße gleich, d. h. ein Modell, welches nach dem AIC_c -Index besser passt, weist auch nach BIC-Index eine bessere Passung auf. Dementsprechend wird nachfolgend nur der AIC_c betrachtet. In der Tabelle 6.1 sind dennoch beide Kennwerte sowie der Likelihood-Wert für alle berechneten Modelle aufgeführt.

Über alle Modelle hinweg weist das Modell *3dim_z* mit 16659 den höchsten AIC_c -Wert und damit die schlechteste Modellpassung auf. Insbesondere passt es schlechter als das Modell *1dim* (AIC_c : 16639). Die geringe Passung dieses technischen Kontrollansatzes *3dim_z* belegt, dass eine beliebige, nicht inhaltlich begründete Zuteilung der Items zu drei Komponenten und damit eine erhöhte Zahl an Modellparametern im Vergleich zum Modell *1dim* nicht automatisch zu einer besseren Modellpassung führt.

Die beste Passung von allen berechneten Modellen konnte mit einem AIC_c -Wert von 16164 mit dem Modell *5dim* erzielt werden. Die Differenz zum Modell *1dim* ist mit 475 erheblich. Das zweitbeste Modell ist das Modell *3dim*, welches mit einem AIC_c von 16219 und damit einer vergleichsweise geringen Differenz von 55 zum Modell *5dim* nur wenig schlechter passt als das letztere. Dass das *3dim*-Modell eine erhebliche AIC_c -Differenz von 460 zum Modell *3dim_z* mit der zufälligen Itemaufteilung aufweist, unterstützt insbesondere die Unterteilung in die drei Komponenten Informationsentnahme, Konstruktion und Integration.

Tabelle 6.1: Kennwerte der Modellpassung. AIC_c = besondere Form des Akaike's Information Criterion, BIC: Best Information Criterion, L = Likelihoodwert.

Modell	Zahl der geschätzten Parameter	AIC_c	Differenz AIC_c	BIC	Differenz BIC	-2 log L
3dim_z	62	16659	-	16852	-	16500.75199
1dim	57	16639	20	16819	33	16496.21378
3dim_q	62	16469	170	16662	157	16310.75304
2dim_within	59	16423	46	16609	54	16274.18798
2dim	59	16387	36	16572	36	16237.92543
2dim_asym	59	16260	126	16446	126	16111.58142
3dim	62	16219	42	16412	34	16060.28086
5dim	71	16164	55	16377	34	15974.96497

Hinweis: Die Modelle sind von höchstem zu niedrigstem AIC_c geordnet. Die Differenzen beziehen sich jeweils auf zwei aufeinanderfolgende Modelle. AIC_c und BIC sind gerundet.

Des Weiteren zeigt der Vergleich des Modells *3dim* zum Modell *3dim_q* mit einer AIC_c -Differenz von 250 eine bedeutend bessere Passung des *3dim*-Modells. Die Differenzierung in die drei Komponenten Rahmen, Daten und Integration scheint demnach zur Modellierung der Diagrammkompetenz nicht geeignet. Auch dieses Ergebnis spricht für die drei theoretisch begründeten Komponenten.

Der Vergleich der zweidimensionalen Modelle (*2dim_within*, *2dim* und *2dim_asym*) sowohl zum Modell *3dim* als auch zum Modell *5dim* weist auf eine Relevanz der Komponente Integration hin: Alle drei Modelle, die die Integration nicht als eigenständige Komponente postulieren, passen schlechter als die beiden Experimentalmodelle mit Integration. Innerhalb der drei zweidimensionalen Modelle konnte dabei mit dem Modell *2dim_asym* und damit durch die Zusammenfassung von Informationsentnahme und Integration die beste Passung erreicht werden. Mit einer AIC_c -Differenz von 42 ist die bessere Passung des *3dim*-Modells zwar bedeutsam und unterstützt die Trennung der zwei genannten Komponenten. Dennoch ist die bessere Passung des Modells *3dim* vergleichsweise gering, was gleichfalls auf einen starken Zusammenhang zwischen Informationsentnahme und Integration hinweisen kann.

Wie in Abschnitt 5.3.3.1 ausgeführt, unterscheidet sich das *5dim*-Modell vom *3dim*-Modell nicht in der grundsätzlichen Struktur, sondern es postuliert die gleichen Hauptkomponenten wie das Modell *3dim*, nämlich Informationsentnahme, Konstruktion und Integration. Die ersten beiden dieser Komponenten werden im Modell *5dim* ihrerseits feiner unterteilt. Beide Modelle weisen im Vergleich zu den Kontrollmodellen nach den Indices AIC_c und BIC die besten Modellpassungen auf. Damit erfüllen sie beide das erste

der vier in Abschnitt 5.3.3.1 dargelegten Kriterien, so dass im Folgenden die weiteren Kriterien für die Überprüfung der Hypothesen auf beide Modelle angewandt werden können.

6.1.1.2 Reliabilitäten

Die Reliabilitäten geben Aufschluss darüber, wie zuverlässig eine Komponente durch ihre Items abgebildet wird. Eine zu geringe Reliabilität einer Komponente wäre als Hinweis darauf zu werten, dass die Komponente nicht als Ganzes interpretiert werden dürfte. Dies spräche gegen die Passung des Modells (vgl. 5.3.3.1, Komponenten: Reliabilitäten). Wie in Abschnitt 5.3.3.1 angemerkt, werden in der vorliegenden Arbeit zwei Reliabilitätsmaße herangezogen, die WLE-Reliabilitäten und die EAP/PV-Reliabilitäten. Sie werden in der folgenden Tabelle 6.2 für die einzelnen Komponenten des Modells *3dim* und des Modells *5dim* aufgeführt. In Klammern hinter dem Komponentennamen ist die Itemzahl der jeweiligen Komponente vermerkt.

Tabelle 6.2: Reliabilitäten der Komponenten des Modells *3dim* (a) und der Komponenten des Modells *5dim* (b). N=289.

a)		b)	
Informations-entnahme (22)	Konstruktion (18)	Identifizierung (6)	Aufbau (12)
WLE: .76 EAP/PV: .78	WLE: .84 EAP/PV: .84	WLE: .22 EAP/PV: .68	WLE: .81 EAP/PV: .83
Integration (7)		AbleSEN (16)	Daten (6)
WLE: .61 EAP/PV: .74		WLE: .74 EAP/PV: .82	WLE: .32 EAP/PV: .71
		Integration (7)	
		WLE: .61 EAP/PV: .79	

Im Modell *3dim* liegen die WLE-Reliabilitäten der Komponenten Informationsentnahme und Konstruktion über $Rel = .7$ und sind damit zufrieden stellend. Der geringere WLE-Reliabilitätswert von $Rel = .61$ bei der Integration ist im Hinblick auf die geringere Anzahl von sieben Items hinreichend groß. Die EAP/PV-Reliabilitäten sind durchweg höher als die WLE-Reliabilitäten, insbesondere für die Integration liegt der Wert bei $Rel = .74$ und damit im guten Bereich. Damit weisen die Reliabilitäten die drei Komponenten insgesamt als interpretierbare Skalen aus und unterstützen damit die Passung des *3dim*-Modells.

Im Modell *5dim* sind die WLE-Reliabilitäten nur für die Komponenten AbleSEN, Aufbau und Integration mit Werten zwischen $Rel = .61$ und $Rel = .81$ annehmbar. Für die

Unterkomponenten Dateneintrag und Identifizierung zeigen sich bezüglich der WLE-Reliabilitäten geringere Werte, die zwar im Hinblick auf die geringe Itemzahl zu erklären, jedoch nicht vertretbar sind. Da für diese Komponenten die EAP/PV-Reliabilitäten deutlich höher und insbesondere im zufrieden stellenden Bereich liegen, kann davon ausgegangen werden, dass die WLE-Reliabilitäten aufgrund des Vorhandenseins von Versuchspersonen mit extrem hohen Fähigkeitswerten bezüglich der jeweiligen Unterkomponente unterschätzt wurden (vgl. 5.3.3.1, Komponenten: Reliabilitäten). Wird in diesen Fällen auf die EAP/PV-Reliabilität zurückgegriffen, so liegen die Reliabilitäten der fünf Komponenten durchweg über $Rel = .68$ und sind damit zufrieden stellend. Nach diesem Kriterium ist somit auch die Passung des *5dim*-Modells gegeben.

6.1.1.3 Personenhomogenität

Wie in Abschnitt 5.3.3.1 (vgl. Komponenten: Personenhomogenität) beschrieben, wurden die Tests zur Personenhomogenität durch Korrelationsberechnungen durchgeführt, deren Ergebnisse in der nachfolgenden Tabelle 6.3 dargestellt sind. Dabei sind jeweils die Korrelationen dreier Zufallsteilungen Z_1 bis Z_3 der Stichprobe sowie die der Teilung der Stichprobe nach Klassenstufe und nach Geschlecht übereinander aufgelistet.

Tabelle 6.3: Korrelationen (Pearsons r) der Testungen auf Personenhomogenität bezüglich den (Unter-)Komponenten der Modelle *3dim* (a) und *5dim* (b). Z_1 bis Z_3 = zufällige Teilungen der Stichprobe (bei den Berechnungen der Modelle *3dim* und *5dim* wurden verschiedene Teilungen vorgenommen), K = Teilung nach Klassenstufe, G = Teilung nach Geschlecht. ** $p < .01$

a)		b)	
Informations-entnahme	Konstruktion	Identifizierung	Aufbau
$Z_1: .976^{**}$	$Z_1: .977^{**}$	$Z_1: .990^{**}$	$Z_1: .987^{**}$
$Z_2: .970^{**}$	$Z_2: .971^{**}$	$Z_2: .962^{**}$	$Z_2: .984^{**}$
$Z_3: .962^{**}$	$Z_3: .959^{**}$	$Z_3: .996^{**}$	$Z_3: .939^{**}$
K: .992**	K: .904**	K: .966**	K: .895**
G: .926**	G: .948**	G: .926**	G: .952**
Integration		AbleSEN	Daten
$Z_1: .978^{**}$ $Z_2: .979^{**}$ $Z_3: .984^{**}$		$Z_1: .980^{**}$	$Z_1: .933^{**}$
K: .970**		$Z_2: .973^{**}$	$Z_2: .988^{**}$
G: .981**		$Z_3: .970^{**}$	$Z_3: .956^{**}$
		K: .969**	K: .969**
		G: .931**	G: .986*
		Integration	
		$Z_1: .978^{**}$ $Z_2: .998^{**}$ $Z_3: .970^{**}$	
		K: .970**	
		G: .981**	

Die Korrelationskoeffizienten zeigen sowohl im Modell *3dim* als auch im Modell *5dim* hoch signifikante Korrelationen über $r = .9$ an. Damit sind die Komponenten ausnahmslos als personenhomogen zu bezeichnen, d. h. alle Personen, insbesondere die Personen der verschiedenen Klassenstufen respektive des verschiedenen Geschlechts, bearbeiteten die

Aufgaben aufgrund jeweils einer Fähigkeit. Nach diesem Kriterium ist somit die Passung sowohl des *3dim*-Modells als auch des *5dim*-Modells unterstützt.

6.1.1.4 Vorhersagewerte der Personenparameter

Es wurde für beide Modelle überprüft, ob die Personenparameter der einzelnen Komponenten unabhängige Vorhersagewerte für die Personenparameter des eindimensionalen Modells sind. Zu diesem Zwecke wurden schrittweise Regressionen berechnet, deren Ergebnisse in Tabelle 6.4 und Tabelle 6.5 dargestellt sind.

Tabelle 6.4: Kennwerte der schrittweisen Regression der Personenparameter des *3dim*-Modells auf den Personenparameter des *1dim*-Modells. *** $p < .001$.

Prädiktor	B	SE _B	β
Informationsentnahme	.406	.010	.456***
Konstruktion	.397	.006	.585***
Integration	.096	.006	.177***
$R^2_{\text{korr.}} = .980, F(3,289) = 4772.946, p < .001$			

Tabelle 6.5: Kennwerte der schrittweisen Regression der Personenparameter des *5dim*-Modells auf den Personenparameter des *1dim*-Modells. ** $p < .001$.

Prädiktor	B	SE _B	β
Identifizierung	.117	.007	.147***
Ablesen	.278	.008	.368***
Aufbau	.263	.006	.450***
Dateneintrag	.155	.008	.189***
Integration	.099	.005	.182***
$R^2_{\text{korr.}} = .984, F(5,289) = 3525.291, p < .001$			

Aus beiden Regressionen geht hervor, dass sowohl im Modell *3dim* als auch im Modell *5dim* sämtliche Komponenten unabhängige Vorhersagewerte für die Gesamtfähigkeit sind. Dabei klären beide Modelle über 98 % der Varianz der Personenparameter der eindimensionalen Modellierung auf. In beiden Regressionen variieren die Beta-Koeffizienten dabei in Bereichen von $\beta = .177$ und $\beta = .585$. Damit stellen sich alle betrachteten Unterkomponenten als relevant für die Gesamtfähigkeit heraus. Für beide Modelle ist nach diesem Kriterium eine Passung gewährleistet.

6.1.1.5 Zusammenfassung

Hinsichtlich der Modellpassung erweisen sich im Vergleich mit den Kontrollmodellen die Modelle *3dim* und *5dim* als beste Modelle zur Modellierung der Diagrammkompetenz. Die Analysen zur Reliabilität, zur Personenhomogenität sowie zu den Vorhersagewerten der Personenparameter unterstützen diese Passung, so dass beide Hypothesen 1a und 1b (vgl. 4.1.1) und damit die theoretisch abgeleiteten Strukturen der Diagrammkompetenz bestätigt werden können. Die Differenzierung in die drei Komponenten Informationsentnahme, Konstruktion und Integration ermöglicht eine adäquate Beschreibung der Diagrammkompetenz. Darüber hinaus erweist sich eine feinere Modellierung, bei der die beiden Komponenten Informationsentnahme und Konstruktion jeweils in zwei Unterkomponenten aufgefächert werden, ebenfalls als geeignet.

6.1.2 Diagrammtypen

Um zu analysieren, ob in einem Strukturmodell der Diagrammkompetenz die beiden Diagrammtypen (DT) Linien- und Säulendiagramm zusammengefasst werden können, wie es in den Modellen *3dim* und *5dim* der Fall ist, werden letztere Modelle mit weiteren Modellen verglichen, in denen die Diagrammtypen unterschieden werden. Die übrigen Strukturen, beispielsweise die Differenzierung der drei Komponenten Informationsentnahme, Konstruktion und Integration im Modell *3dim*, wird beibehalten. Wie in Abschnitt 4.1.2 beschrieben, ergeben sich daraus drei relevante Testmodelle: Das Modell *2dimDT*, welches Diagrammkompetenz als einheitliches Konstrukt auffasst und nur die Unterscheidung der Diagrammtypen berücksichtigt, das Modell *2dimDTx3dim*, welches die Struktur des *3dim*-Modells aufgreift und um die Differenzierung nach Diagrammtypen erweitert, sowie das Modell *2dimDTx5dim*, welches die Einteilung des *5dim*-Modells um die Unterscheidung der Diagrammtypen erweitert.

Wie in Abschnitt 6.1.1 werden nachfolgend die Kriterien zur Annahme eines Modells, die in Abschnitt 5.3.3.1 methodisch beschrieben wurden, herangezogen.

6.1.2.1 Modellpassung

In der nachfolgenden Tabelle werden die genannten Testmodelle von schlechtester zu bester Modellpassung aufgeführt. Analog zu Abschnitt 6.1.1 wird in den weiteren Beschreibungen lediglich der AIC_c -Index aufgegriffen.

Tabelle 6.6: Kennwerte der Modellpassung. AIC_c = besondere Form des Akaike's Information Criterion, BIC: Best Information Criterion, L = Likelihood-Wert.

Modell	Zahl der geschätzten Parameter	AIC_c	Differenz AIC_c	BIC	Differenz BIC	-2 log L
<i>1dim</i>	57	16639	-	16819	-	16496.21378
<i>2dimDT</i>	57	16563	76	16743	76	16420.21681
<i>3dim</i>	62	16219	344	16412	331	16060.28086
<i>5dim</i>	71	16164	55	16377	35	15974.96497
<i>2dimDTx3dim</i>	71	16139	25	16353	24	15950.36960
<i>2dimDTx5dim</i>	84	16085	54	16323	30	15847.46496

Hinweis: Die Modelle sind von höchstem zu niedrigstem AIC_c -Index geordnet. Die Differenzen beziehen sich jeweils auf zwei aufeinander folgende Modelle. AIC_c und BIC sind gerundet.

Aus der obigen Tabelle geht hervor, dass ein Modell, das lediglich nach den beiden Diagrammtypen differenziert, keine bessere Modellpassung aufweist als die Modelle, die hinsichtlich der Komponentenstruktur als beste Modelle identifiziert wurden (vgl. 6.1.1.1). So zeigt das Modell *2dimDT* mit einem AIC_c von 16563 bereits eine deutlich schlechtere Modellpassung als das *3dim*-Modell (Differenz zwischen beiden Modellen: 344) und entsprechend eine erheblich schlechtere Passung als das *5dim*-Modell (Differenz: 399). Für die beiden Modelle *2dimDTx3dim* und *2dimDTx5dim*, die eine Differenzierung sowohl nach Diagrammtyp als auch nach Komponenten vornehmen, zeigen beide Modellpassungsindices eine Überlegenheit gegenüber den Modellen *3dim* und *5dim* an. Insbesondere die AIC_c -Differenz von 79 (=25+54) des Modells *2dimDTx5dim* zum *5dim*-Modell weist darauf hin, dass dieses Modell die Diagrammkompetenz am besten abzubilden vermag. Aus diesem Grund werden, wie in Abschnitt 6.1.1, im Folgenden zuerst die Reliabilitäten der Einzelskalen als Gütekriterium der Passung herangezogen, um die bessere Passung des Modells *2dimDTx5dim* gegenüber dem *5dim*-Modell (und ebenso dem *3dim*-Modell) zu bestätigen oder zu verwerfen.

6.1.2.2 Reliabilitäten

In der nachfolgenden Tabelle 6.7 sind für beide Modelle die Reliabilitäten (WLE-Reliabilität und EAP/PV-Reliabilität, vgl. 5.3.3.1; Komponenten) der jeweiligen Skalen dargestellt. Es wird deutlich, dass für beide Modelle die WLE-Reliabilitäten inakzeptabel sind und dagegen sprechen, die Skalen zu interpretieren. Dies wird auch durch die EAP/PV-Reliabilitäten nicht revidiert: Im Modell *2dimDTx3dim* liegt nur die Reliabilität der Konstruktion in einem akzeptablen Bereich, im Modell *2dimDTx5dim* lediglich die Integration und die Rasch-Dimensionen Linien- und Säulendiagramm. Insgesamt ist mit

jeweils vier Skalen, die eine ungenügende Reliabilität aufweisen, keines der beiden Modelle geeignet, die Diagrammkompetenz adäquat zu beschreiben.

Tabelle 6.7: Reliabilitäten der Rasch-Dimensionen in den Modellen *2dimDTx3dim* und *2dimDTx5dim* (N = 289).

<i>2dimDTx3dim</i>	Informationsentnahme		Konstruktion		Integration	Linien-diagramm	Säulen-diagramm
Rel. WLE	.000		.000		.000	.000	.000
Rel. EAP/PV	.338		.666		.596	.463	.481
<i>2dimDTx5dim</i>	Identifizierung	Ablesen	Aufbau	Dateneintrag	Integration	Linien-diagramm	Säulen-diagramm
Rel. WLE	.000	.000	.000	.000	.000	.000	.000
Rel. EAP/PV	.384	.542	.556	.503	.640	.826	.783

Ausgehend von diesem Ergebnis sind die beiden Modelle *2dimDTx3dim* und *2dimDTx5dim* zur Modellierung der Diagrammkompetenz abzulehnen. Die Modelle *3dim* und *5dim*, die keine Unterscheidung der Komponenten nach Diagrammtypen vornehmen, bilden die Diagrammkompetenz besser ab. Somit kann die Hypothese 2 (vgl. 4.1.2) bestätigt werden, wonach Linien- und Säulendiagramme bei der Modellierung der Diagrammkompetenz zusammengefasst werden können.

6.1.3 Zusammenfassung

Wie die vorangegangenen Analysen hinsichtlich der Komponenten der Diagrammkompetenz belegen, sind beide Modelle *3dim* und *5dim*, die die Komponenten des Strukturmodells der Diagrammkompetenz respektive deren Unterkomponenten repräsentieren, zur Modellierung der Diagrammkompetenz geeignet. Sie erfüllen die vier in Abschnitt 5.3.3.1 angeführten Kriterien, die sie als zur Modellierung der Diagrammkompetenz geeignete Modelle ausweisen. Damit können beide Hypothesen 1a und 1b bestätigt werden. Die Komponenten Informationsentnahme, Konstruktion und Integration sowie die Unterkomponenten der Informationsentnahme (Identifizierung und Ablesen) sowie die Unterkomponenten der Konstruktion (Aufbau und Dateneintrag) lassen sich als Strukturen der Diagrammkompetenz empirisch nachweisen.

Darüber hinaus zeigt sich, dass eine Differenzierung nach den Diagrammtypen Linien- und Säulendiagramm für ein Strukturmodell der Diagrammkompetenz ungeeignet ist. Ein Modell, das lediglich nach Diagrammtypen unterteilt und keine weiteren Strukturierungen vornimmt, passt deutlich schlechter als das Strukturmodell mit den drei Komponenten Informationsentnahme, Konstruktion und Integration. Die beiden Modelle, die die Modelle *3dim* und *5dim* jeweils um die Unterteilung nach Diagrammtypen erweitern, sind abzulehnen, weil sie überwiegend zu nicht interpretierbaren Skalen führen. Somit kann auch die Hypothese 2 bestätigt werden.

Zusammenfassend bleibt festzuhalten, dass sich das in Abschnitt 3.6 beschriebene Strukturmodell mit drei respektive fünf Fähigkeitskomponenten als geeignet zur Modellierung der Diagrammkompetenz erweist. Insbesondere ist zu konstatieren, dass dieses Modell damit übergreifend für beide Diagrammtypen als gültig erachtet werden kann.

6.2 Differenzierte Beschreibung der Modellstruktur

Entsprechend den in Kapitel 4 dargelegten Forschungsfragen ist dieser Abschnitt in vier Teile gegliedert. Im ersten Teil wird untersucht, welche Relationen zwischen den einzelnen Komponenten und Unterkomponenten der Diagrammkompetenz bestehen (6.2.1). Im Abschnitt 6.2.2 werden die (Unter-)Komponenten hinsichtlich ihrer Schwierigkeiten miteinander verglichen. Der Abschnitt 6.2.3 ist der Analyse der Binnenstrukturen einzelner Komponenten gewidmet. Im Abschnitt 6.2.4 werden die Schwierigkeiten für die verschiedenen Diagrammtypen analysiert.

6.2.1 Relationen

Die folgende Darstellung von Ergebnissen zu Zusammenhangsanalysen gliedert sich in drei Teile. Zuerst werden überblicksartig die Zusammenhänge zwischen jeweils zwei Komponenten dargestellt (6.2.1.1). Im zweiten Teil werden zum einen die genannten Zusammenhänge miteinander verglichen, zum anderen die komplexeren Zusammenhänge zwischen mehreren (Unter-)Komponenten betrachtet, wobei die Integration vorerst ausgeschlossen wird (6.2.1.2). Im dritten Teil werden die komplexen Zusammenhänge zwischen der Integration und den weiteren zwei bzw. vier (Unter-)Komponenten analysiert (6.2.1.3).

6.2.1.1 Zusammenhänge zwischen jeweils zwei Komponenten

Aufschluss über die Zusammenhänge zwischen den einzelnen Komponenten der Diagrammkompetenz geben die nachfolgend dargestellten Korrelationen.

Tabelle 6.8: Korrelationen (Pearsons r) der Komponenten innerhalb des Modells *3dim*. ** $p < .01$.

	Konstruktion	Integration
Informationsentnahme	.435**	.608**
Konstruktion		.334**

Tabelle 6.9: Korrelationen (Pearsons r) der Komponenten innerhalb des Modells *5dim*. ** $p < .01$.

	Ablesen	Aufbau	Dateneintrag	Integration
Identifizierung	.406**	.312**	.190**	.329**
Ablesen		.418**	.287**	.601**
Aufbau			.651**	.349**
Dateneintrag				.214**

Es zeigen sich für alle (Unter-)Komponenten hoch signifikante Zusammenhänge untereinander. Die Stärke des Zusammenhangs variiert dabei zwischen $r = .334$ und $r = .608$ im Modell *3dim* und zwischen $r = .190$ und $r = .601$ im *5dim*-Modell. In den nachfolgenden Analysen wird detailliert Bezug auf die Korrelationen genommen.

6.2.1.2 Zusammenhänge zwischen Informationsentnahme und Konstruktion bzw. deren Unterkomponenten

Laut Korrelationsanalysen (vgl. Tabelle 6.8) zeigt sich ein Zusammenhang von $r = .435$ der Komponenten Informationsentnahme und Konstruktion. In einer einfachen bivariaten Regression wird dazu als Effektstärke (vgl. 5.3.3.2, Relationen der Komponenten zueinander) $R^2_{\text{kor.}} = .186$ ermittelt⁶². Mit einer aufgeklärten Varianz von 18,6 % liegt ein mittlerer Effekt vor.

Im Vergleich zum genannten Zusammenhang zwischen den Komponenten Informationsentnahme und Konstruktion sind sämtliche Korrelationen einzelner Unterkomponenten der Informationsentnahme mit den zwei Unterkomponenten der anderen Komponente (Konstruktion) geringer, variieren dabei allerdings stark. Hinsichtlich der Identifizierung besteht ein höherer Zusammenhang zum Aufbau ($r = .312$) als zum Dateneintrag ($r = .190$). Hinsichtlich des Ablesens sind die Zusammenhänge stärker, weisen aber ein ähnliches Verhältnis auf: Auch hier besteht ein stärkerer Zusammenhang zum Aufbau ($r = .418$) als zum Dateneintrag ($r = .287$). Der geringere Zusammenhang der beiden Unterkomponenten Identifizierung und Ablesen zum Dateneintrag wird zudem weiter verringert, wenn in einer schrittweisen Regression der Dateneintrag als abhängige Variable und die übrigen drei Unterkomponenten als Prädiktoren eingesetzt werden: In diesem Fall zeigt sich nur der Aufbau als signifikanter Vorhersagewert (vgl. Tabelle 6.10).

⁶² Die vollständigen Kennwerte der Regression befinden sich im Anhang (s. Tabelle 12.7).

Tabelle 6.10: Kennwerte der schrittweisen Regression mit Dateneintrag als abhängiger Variable.
*** $p < .001$.

Prädiktor	b	SE _b	β
Aufbau	.464	.032	.651***
$R^2_{\text{korr.}} = .424, F(2,289) = 210.839, p < .001$			

Es wird damit deutlich, dass der Dateneintrag einen erheblich stärkeren Bezug zum Aufbau hat als zu den Unterkomponenten der Informationsentnahme. Mit einer Varianzaufklärung von 42,4% liegt dabei ein starker Effekt vor.

Insbesondere hängen Aufbau und Dateneintrag mit $r = .651$ vergleichsweise stark zusammen, während der totale Zusammenhang zwischen Identifizierung und Ablesen mit $r = .406$ geringer ist. Dies weist darauf hin, dass sich die Unterkomponenten innerhalb der Informationsentnahme besser separieren lassen als in der Konstruktion.

Des Weiteren sind die eben genannten Zusammenhänge *innerhalb* der beiden Komponenten Informationsentnahme und Konstruktion höher als die Zusammenhänge *zwischen* den beiden Unterkomponenten, die sich auf den Diagrammrahmen beziehen ($r = .312$) bzw. zwischen den beiden, die sich auf den Werteintrag beziehen ($r = .287$). Dies unterstützt das Ergebnis aus den Modellberechnungen (vgl. 6.1.1.1), nach denen das Modell *3dim_g*, das die Komponenten Rahmen und Werteintrag differenziert, im Vergleich zum *3dim*-Modell nicht zur Modellierung der Diagrammkompetenz geeignet ist.

Hervorzuheben ist darüber hinaus, dass nach den Korrelationen die Unterkomponenten Ablesen und Aufbau, die verschiedenen Hauptkomponenten zugeteilt sind, mit $r = .418$ stark zusammenhängen. Eine schrittweise Regression mit Identifizierung und Ablesen als Prädiktoren für den Aufbau (vgl. Tabelle 6.11) bestätigt diesen vergleichsweise starken Zusammenhang.

Tabelle 6.11: Kennwerte der schrittweisen Regression mit Aufbau als abhängiger Variable. ** $p < .01$,
*** $p < .001$.

Prädiktor	B	SE _B	β
Ablesen	.451	.075	.349***
Identifizierung	.232	.079	.171**
$R^2_{\text{korr.}} = .192, F(2,289) = 35.552, p < .001$			

So zeigen sich sowohl Ablesen als auch Identifizierung als signifikante Prädiktoren für den Aufbau, wobei Ablesen der stärkere Prädiktor ist ($\beta = .349$). Mit einer Varianzaufklärung von 19,2 % liegt ein mittlerer Effekt und damit eine vergleichbare Effektstärke wie beim Zusammenhang zwischen Informationsentnahme und Konstruktion vor.

Im Vergleich dazu wird in einer Regression mit der abhängigen Variable Dateneintrag, in die die Identifizierung und das Ablesen als unabhängige Variablen eingegeben werden, nur das Ablesen signifikanter Prädiktor mit $\beta = .287$. Die Varianzaufklärung dieses Regressionsmodells ist mit 7,9 % gering.

Tabelle 6.12: Kennwerte der schrittweisen Regression mit Dateneintrag als abhängiger Variable.
*** $p < .001$.

Prädiktor	B	SE _B	β
Ablesen	.265	.0521	.287***
$R^2_{\text{kor.}} = .079, F(1,289) = 32.619, p < .001$			

Wenngleich sich in einer der beiden letztgenannten Regressionen die Identifizierung als signifikanter Prädiktor zeigt, ist festzuhalten, dass das Ablesen die Unterkomponente mit den stärkeren Bezügen zur Konstruktion ist. Da sich diese Bezüge allerdings auf den Aufbau beschränken, kann gefolgert werden, dass der Zusammenhang zwischen den Komponenten Informationsentnahme und Konstruktion vordergründig auf den Zusammenhang zwischen Ablesen und Aufbau zurückgeführt werden kann.

6.2.1.3 Zusammenhänge zwischen Integration und den anderen Komponenten bzw. Unterkomponenten

In den Korrelationsanalysen (vgl. 6.2.1.1) zeigen sich sowohl zwischen Informationsentnahme und Integration als auch zwischen Konstruktion und Integration hoch signifikante Zusammenhänge, wobei die Korrelation hinsichtlich der Informationsentnahme stärker ist ($r = .608$). Der Zusammenhang zwischen Konstruktion und Integration bleibt in einer schrittweisen Regression mit Informationsentnahme und Konstruktion als Prädiktoren für Integration indes nicht bestehen, es wird nur noch die Informationsentnahme signifikant mit $\beta = .608$ (vgl. Tabelle 6.13). Mit der resultierenden Modellgleichung werden 36,7 % der Varianz aufgeklärt, was auf einen starken Effekt hinweist.

Tabelle 6.13: Kennwerte der Regression mit Integration als abhängiger Variable. *** $p < .001$.

Prädiktor	B	SE _B	β
Informationsentnahme	.998	.077	.608***
$R^2_{\text{kor.}} = .367, F(1,289) = 167.924.841, p < .001$			

Dieses Ergebnis legt insbesondere vor dem Hintergrund, dass sich eine Korrelation zwischen Informationsentnahme und Konstruktion ($r = .435$) gezeigt hat, nahe zu überprüfen, ob der durch die Korrelationsanalysen (s.o.) festgestellte Zusammenhang

zwischen Konstruktion und Integration auf die Informationsentnahme zurückzuführen ist. Für die Darstellung der entsprechenden Mediationsanalyse wird in Anlehnung an die methodische Beschreibung (vgl. 5.3.3.2, Relationen der Komponenten zueinander) im Folgenden auf das in Abbildung 6.1 dargestellte Schaubild Bezug genommen.

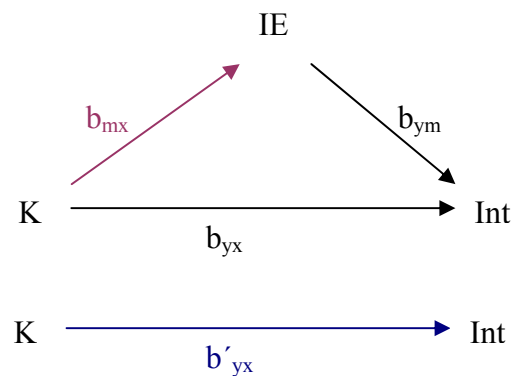
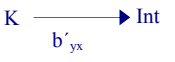
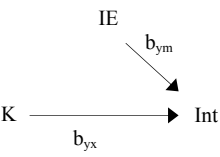
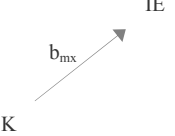


Abbildung 6.1: Schaubild zu den berechneten Zusammenhängen zwischen den Komponenten des *3dim*-Modells. IE = Informationsentnahme, K = Konstruktion, Int = Integration. Zum Abgleich mit der methodischen Beschreibung (vgl. 5.3.3.2, Relationen der Komponenten) wurden die Benennungen der Regressionskoeffizienten beibehalten.

Die notwendigen Berechnungen 1 bis 3 zur Überprüfung der vier in 5.3.3.2 (s. Mediation) genannten Voraussetzungen (1) bis (4) sind in Kurzdarstellung in Tabelle 6.14 zusammengefasst. Die ausführlichen Darstellungen der Berechnungen befinden sich im Anhang (s. Tabelle 12.8).

Tabelle 6.14: Übersicht über die Berechnungen und Prüfungen der Voraussetzungen für eine Mediation des Zusammenhangs zwischen Konstruktion und Integration durch die Informationsentnahme. K = Konstruktion, IE = Informationsentnahme, Int = Integration. Zum Abgleich mit der methodischen Beschreibung (vgl. 5.3.3.2, Relationen der Komponenten) wurden die Benennungen der Regressionskoeffizienten beibehalten. ** $p < .01$, *** $p < .001$.

Ergebnisse der Regressionen 1 bis 3			erfüllte Vorausss.
1 	K ist höchst signifikanter Prädiktor für Int, $R^2_{\text{kor.}} = .108^{***}$.	$b'_{yx} = .399^{***}$	(1)
2 	K ist kein signifikanter Prädiktor für Int, IE ist hoch signifikanter Prädiktor für Int, $R^2_{\text{kor.}} = .371^{***}$.	$b_{yx} = .103$, insbesondere gilt $b_{yx} < b'_{yx}$	(2)
		$b_{ym} = .936^{**}$, $SE(b_{ym}) = .085$	(3)
3 	K ist höchst signifikanter Prädiktor für IE, $R^2_{\text{kor.}} = .186^{***}$.	$b_{mx} = .316^{***}$, $SE(b_{mx}) = .039$	(4)

Da alle vier Bedingungen erfüllt sind, ist die Berechnung des indirekten Zusammenhangs zwischen Konstruktion und Integration gemäß der in Abschnitt 5.3.3.2 (s. Relationen der Komponenten zueinander) dargestellten Formeln möglich:

$$b_{mx \cdot ym} = b_{mx} \cdot b_{ym} = 0.316 \cdot 0.936 = 0.295776 \text{ und}$$

$$\begin{aligned} SE_{b_{mx \cdot ym}} &= \sqrt{b_{mx}^2 \times SE_{b_{ym}}^2 + b_{ym}^2 \times SE_{b_{mx}}^2} \\ &= \sqrt{0.316^2 \times 0.085^2 + 0.936^2 \times 0.039^2} \\ &= \sqrt{0.0020540016} \\ &= 0.045321095 \end{aligned}$$

Damit ergibt sich insgesamt $t(288) = 6,5262$, womit sich der Sobel-Test mit $p < .001$ als signifikant erweist. Folglich mediiert die Informationsentnahme den Zusammenhang von Konstruktion und Integration. Da sich Konstruktion in der Regression 2 nicht als signifikanter Prädiktor für Integration herausstellt, fungiert die Informationsentnahme sogar als totaler Mediator des Zusammenhangs. Mit einem indirekten Zusammenhang von $b_{mx \cdot ym} = .296$ und einem totalen Zusammenhang von $b'_{yx} = .399$ werden dabei rund 74 % des Zusammenhangs von Konstruktion und Integration über die Informationsentnahme vermittelt.

Bezüglich des *5dim* zeigen in den Korrelationsanalysen (s.o.) alle vier Unterkomponenten des *5dim* einen signifikanten Zusammenhang zur Integration, der allerdings stark variiert. Am stärksten korreliert das Ablesen ($r = .601$), die zweithöchste Korrelation zeigt sich zwischen Aufbau und Integration ($r = .349$). Auch in einer schrittweisen Regression, in die als Prädiktoren für die Integration alle vier Unterkomponenten aufgenommen wurden, werden lediglich Ablesen und Aufbau als Vorhersagewerte signifikant, wobei hier ebenfalls ein deutlicher Unterschied zwischen beiden erkennbar ist: Der Zusammenhang zwischen Ablesen und Integration ist mit $\beta = .552$ erheblich stärker als der zwischen Aufbau und Integration mit $\beta = .118$ (vgl. Tabelle 6.15).

Tabelle 6.15: Kennwerte der schrittweisen Regression für die abhängige Variable Integration. * $p < .05$
*** $p < .001$.

Prädiktor	b	SE _b	β
Ablesen	.770	.072	.552***
Aufbau	.128	.056	.118*

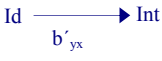
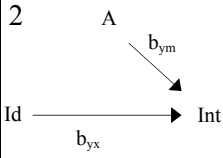
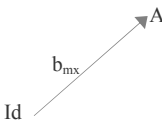
$R^2_{\text{kor.}} = .369, F(2,289) = 85.191, p < .001$

In den Korrelationsanalysen zeigen sich damit signifikante Zusammenhänge zwischen allen vier Komponenten und der Integration (vgl. Tabelle 6.9), in der Regression werden die Komponenten Identifizierung und Dateneintrag als Vorhersagewerte signifikant, wobei das Ablesen den weitaus stärkeren Prädiktor darstellt. Aufgrund dessen wurden Mediationsanalysen durchgeführt, um aufzudecken, ob die eingangs berichteten Zusammenhänge über das Ablesen mediiert werden.

Mediatoreffekte für den Zusammenhang zwischen Identifizierung und Integration

In den folgenden Berechnungen wird überprüft, ob der Zusammenhang zwischen Identifizierung und Integration über das Ablesen vermittelt wird. Wie oben wird die folgende tabellarische Übersicht (Tabelle 6.16) zur Darstellung der Ergebnisse der drei Regressionsanalysen genutzt, die ausführliche Darstellung befindet sich im Anhang (s. Tabelle 12.9).

Tabelle 6.16: Übersicht über die Berechnungen und Prüfungen der Voraussetzungen für eine Mediation des Zusammenhangs zwischen Identifizierung und Integration durch das Ablesen. Id = Identifizierung, A = Ablesen, Int = Integration. Zum Abgleich mit der methodischen Beschreibung (vgl. 5.3.3.2, Relationen der Komponenten) wurden die Benennungen der Regressionskoeffizienten beibehalten. * $p < .05$, *** $p < .001$.

Ergebnisse der Regressionen 1 bis 3			erfüllte Vorausss.
1 	Id ist höchst signifikanter Prädiktor für Int, $R^2_{\text{korrr.}} = .105^{***}$.	$b'_{yx} = .482^{***}$	(1)
2 	Id ist signifikanter Prädiktor für Int, A ist höchst signifikanter Prädiktor für Int, $R^2_{\text{korrr.}} = .366^{***}$.	$b_{yx} = .148^*$, insbesondere gilt $b_{yx} < b'_{yx}$	(2)
		$b_{ym} = .782^{***}$, $SE(b_{ym}) = .072$	(3)
3 	Id ist höchst signifikanter Prädiktor für A, $R^2_{\text{korrr.}} = .162^{***}$.	$b_{mx} = .427^{***}$, $SE(b_{mx}) = .057$	(4)

Es ergibt sich ein indirekter Zusammenhang zwischen Identifizierung und Integration von $b_{mx \cdot ym} = .334$. Die Prüfung mit dem Sobeltest zeigt, dass ein mit $p < .001$ höchst signifikanter Mediatoreffekt vorliegt. Da in Regression 2 die Identifizierung ein zwar schwacher, aber dennoch signifikanter Prädiktor ist, wird das Ablesen lediglich als partieller Mediator des Zusammenhangs zwischen Identifizierung und Integration klassifiziert. Der Vergleich des indirekten Zusammenhangs ($b_{mx \cdot ym} = .334$) mit dem totalen Zusammenhang ($b'_{yx} = .482$) zeigt, dass etwa 69 % des Zusammenhangs durch das Ablesen vermittelt werden.

Mediatoreffekte für den Zusammenhang zwischen Aufbau und Integration

Für den Zusammenhang zwischen Aufbau und Integration wird eine Mediationswirkung durch das Ablesen überprüft. Die übersichtliche Darstellung der Ergebnisse aus den relevanten Regressionsanalysen erfolgt in Tabelle 6.17, die ausführlichen Ergebnisse sind im Anhang (s. Tabelle 12.10) aufgeführt.

Tabelle 6.17: Übersicht über die Berechnungen und Prüfungen der Voraussetzungen für eine Mediation des Zusammenhang zwischen Aufbau und Integration durch das Ablesen. Auf = Aufbau, A = Ablesen, Int = Integration. Zum Abgleich mit der methodischen Beschreibung (vgl. 5.3.3.2, Relationen der Komponenten) wurden die Benennungen der Regressionskoeffizienten beibehalten. * $p < .05$, *** $p < .001$.

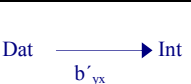
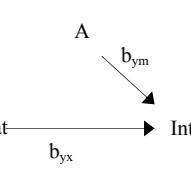
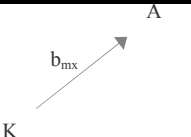
Ergebnisse der Regressionen 1 bis 3			erfüllte Vorausss.
1 Auf $\xrightarrow{b'_{yx}}$ Int	Auf ist höchst signifikanter Prädiktor für Int, $R^2_{\text{kor.}} = .119^{***}$.	$b'_{yx} = .377^{***}$	(1)
2 A \searrow Int Auf $\xrightarrow{b_{yx}}$ Int	Auf ist signifikanter Prädiktor für Int, A ist höchst signifikanter Prädiktor für Int, $R^2_{\text{kor.}} = .369^{***}$.	$b_{yx} = .128^*$, insbesondere gilt $b_{yx} < b'_{yx}$	(2)
		$b_{ym} = .770^{***}$, $SE(b_{ym}) = .072$	(3)
3 Auf \nearrow A	Auf ist höchst signifikanter Prädiktor für A, $R^2_{\text{kor.}} = .172^{***}$.	$b_{mx} = .323^{***}$, $SE(b_{mx}) = .041$	(4)

Wie aus der Tabelle ersichtlich, sind alle vier Voraussetzung zur Prüfung der Mediatorwirkung des Ablesens erfüllt. Der Sobel-Test zeigt, dass der indirekte Zusammenhang $b_{mx \cdot ym} = .249$ zwischen dem Aufbau und der Integration mit $p < .001$ höchst signifikant wird. Wie im vorangegangenen Fall liegt, da der Aufbau in Regression 2 signifikanter Vorhersagewert ist, eine partielle Mediation vor. Dabei werden rund 66 % des totalen Zusammenhangs $b'_{yx} = .377$ zwischen Aufbau und Integration durch das Ablesen vermittelt.

Mediatoreffekte für den Zusammenhang zwischen Dateneintrag und Integration

Im Folgenden wird überprüft, ob der Zusammenhang zwischen Dateneintrag und Integration durch das Ablesen vermittelt wird.

Tabelle 6.18: Übersicht über die Berechnungen und Prüfungen der Voraussetzungen für eine Mediation des Zusammenhangs zwischen Dateneintrag und Integration durch das Ablesen. Dat = Dateneintrag, A = Ablesen, Int = Integration. Zum Abgleich mit der methodischen Beschreibung (vgl. 5.3.3.2, Relationen der Komponenten) wurden die Benennungen der Regressionskoeffizienten beibehalten. *** $p < .001$.

Ergebnisse der Regressionen 1 bis 3			erfüllte Vorausss.
1		Dat ist höchst signifikanter Prädiktor für Int, $R^2_{\text{kor.}} = .043^{***}$.	$b'_{yx} = .325^{***}$ (1)
2		Dat ist kein signifikanter Prädiktor für Int,	$b_{yx} = .045$, insbesondere gilt $b_{yx} < b'_{yx}$ (2)
		A ist hoch signifikanter Prädiktor für Int, $R^2_{\text{kor.}} = .359^{***}$.	$b_{ym} = .821^{***}$, $SE(b_{ym}) = .069$ (3)
3		Dat ist höchst signifikanter Prädiktor für A, $R^2_{\text{kor.}} = .079^{***}$.	$b_{mx} = .312^{***}$, $SE(b_{mx}) = .061$ (4)

Alle Voraussetzungen für eine Mediation sind erfüllt. Erneut wurde der indirekte Zusammenhang $b_{mx \cdot ym} = .256$ mithilfe des Sobel-Tests auf Signifikanz überprüft. Es ergab sich eine Signifikanz mit $p = .00000236$, die das Ablesen als Mediator des Zusammenhangs zwischen Aufbau und Integration ausweist. Da der Dateneintrag in gemeinsamer Regression 2 mit dem Ablesen als Vorhersagewert nicht signifikant wurde, liegt ein totaler Mediatoreffekt vor. Hinsichtlich des totalen Zusammenhangs $b'_{yx} = .325$ zwischen Dateneintrag und Integration werden mit $b_{mx \cdot ym} = .256$ rund 79 % durch das Ablesen mediiert.

6.2.1.4 Zusammenfassung

Insgesamt zeigen sich Zusammenhänge zwischen allen (Unter-)Komponenten, die sich dabei in ihrer Stärke unterscheiden. Erwartungsgemäß hängen die jeweils zwei Unterkomponenten von Informationsentnahme und Konstruktion stark zusammen. Schüler, die die Identifizierung beherrschen, besitzen somit auch gewisse Fähigkeiten zum Ablesen und umgekehrt. Ebenso können Schüler, die einen Diagrammrahmen konstruieren können, auch in gewissem Maße Daten in ein Diagramm eintragen und umgekehrt. Dies unterstützt die Modellstruktur des 3dim-Modells und hebt die Möglichkeit, die Diagrammkompetenz grundsätzlich in drei Komponenten zu unterteilen, hervor.

Zwischen den Komponenten Informationsentnahme und Konstruktion konnte ein Zusammenhang von $r = .435$ festgestellt werden. Wie differenzierte Analysen anhand des 5dim-Modells zeigen, ist dabei das Ablesen die Unterkomponente der

Informationsentnahme mit den stärkeren Bezügen zur Konstruktion. Die Zusammenhänge beschränken sich allerdings auf den Aufbau, so dass zusammenfassend festgestellt werden kann, dass der Zusammenhang zwischen Informationsentnahme und Konstruktion vor allem auf dem starken Zusammenhang zwischen Ablesen und Aufbau beruht.

Hinsichtlich der Integration kann die Informationsentnahme als bedeutsamer als die Konstruktion identifiziert werden. Zudem lässt sich der Zusammenhang von Konstruktion und Integration zu 74 % auf das Ablesen zurückführen. In den Analysen zum *5dim*-Modell wird deutlich, dass der starke Zusammenhang zwischen Informationsentnahme und Integration hauptsächlich auf dem Ablesen beruht. Es korreliert von den vier Unterkomponenten mit Abstand am höchsten mit der Integration und fungiert darüber hinaus als Mediator der Zusammenhänge zwischen den übrigen drei Komponenten und der Integration. Dabei können zwischen 66 % und 79 % der Zusammenhänge über das Ablesen erklärt werden.

6.2.2 Schwierigkeiten der Komponenten

Anhand der Itemparameter, die bei den Modellierungen geschätzt werden, werden die Anforderungen der einzelnen Komponenten ermittelt. Nachfolgend ist die deskriptive Statistik bezüglich der Itemparameter der Komponenten des *3dim*-Modells dargestellt.

Tabelle 6.19: Deskriptive Statistik der Itemschwierigkeiten für die Komponenten des *3dim*-Modells.

3dim-Modell	N	M	SD	Median	Min	Max	Spannweite
Informationsentnahme	22	-0.29	1.17	0.09	-2.42	2.06	4.47
Konstruktion	18	0.07	1.04	-0.02	-2.15	1.69	3.84
Integration	7	0.48	1.58	0.74	-1.73	2.72	4.45
Insgesamt	47	-0.04	1.20	0.05	-2.42	2.72	5.14

In der Tabelle 6.19 sowie der Abbildung 6.2 wird deutlich, dass die Informationsentnahme im Mittel die geringste Itemschwierigkeit aufweist ($M = -0.29$, $SD = 1.17$), die Spannweite der Schwierigkeiten dieser Komponente allerdings mit 4,47 sehr hoch ist und damit die Spannweite der Konstruktion überragt. Der Mittelwert der Itemschwierigkeiten der Konstruktion liegt dabei um 0.36 Logit höher, der der Integration um weitere 0.41 Logits. Wie die Kruskal-Wallis-Varianzanalyse zeigt, sind diese Unterschiede jedoch nicht signifikant ($\chi^2(2) = 1.525$; $p = .466$ n.s.)

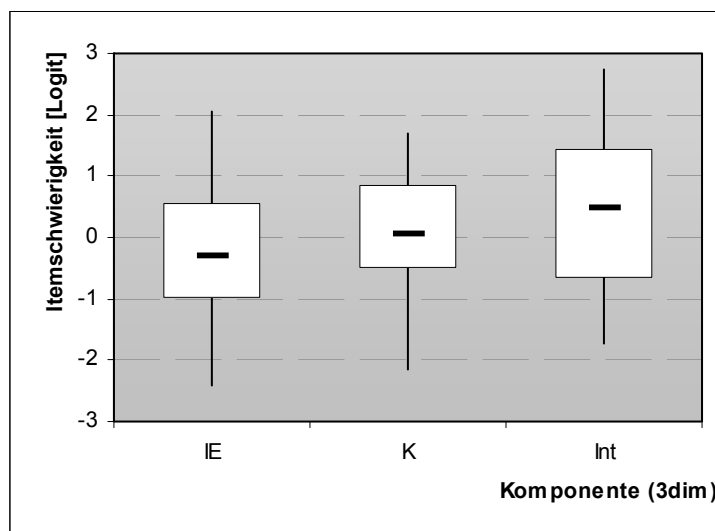


Abbildung 6.2: Mittlere Itemschwierigkeiten der Komponenten im 3dim-Modell.

Im Folgenden wird die deskriptive Statistik der Itemschwierigkeiten für die Komponenten des Modells *5dim* dargestellt.

Tabelle 6.20: Deskriptive Statistik über die Itemschwierigkeiten für die Komponenten des *5dim*-Modells.

<i>5dim</i> -Modell	N	M	SD	Median	Min	Max	Spannweite
Identifizierung	6	-0.55	1.37	-0.20	-2.27	0.74	3.01
Ablesen	16	-0.12	1.15	0.15	-2.47	2.22	4.69
Aufbau	12	-0.15	1.08	-0.15	-2.15	1.85	3.99
Dateneintrag	6	0.70	0.77	0.80	-0.32	1.59	1.91
Integration	7	0.52	1.58	0.78	-1.68	2.78	4.46
Insgesamt	47	0.02	1.21	0.10	-2.47	2.78	5.25

Insgesamt sind die beiden Komponenten, die sich auf den Rahmen beziehen, leichter als die anderen drei Komponenten. Die Identifizierung zeigt mit $M = -0.55$ ($SD = 1.37$) im Mittel die geringste Schwierigkeit. Als zweitleichteste Komponente erweist sich der Aufbau, der im Mittel um 0.4 Logit schwerer als die Identifizierung ist. Die Schwierigkeitsdifferenz vom Aufbau zum Ablesen ist mit 0.03 Logit sehr gering, zudem überragt die Spannweite der Schwierigkeiten der Ableseitems sowohl die der Identifizierung als auch die des Aufbaus. Die zweitschwierigste Komponente ist die Integration, die mit einer mittleren Itemschwierigkeit von $M = 0.52$ Logit ($SD = 1.58$) deutlich über der des Ablesens liegt. Die Items des Dateneintrags weisen im Mittel eine Schwierigkeit von $M = 0.7$ Logit ($SD = 0.77$) auf.

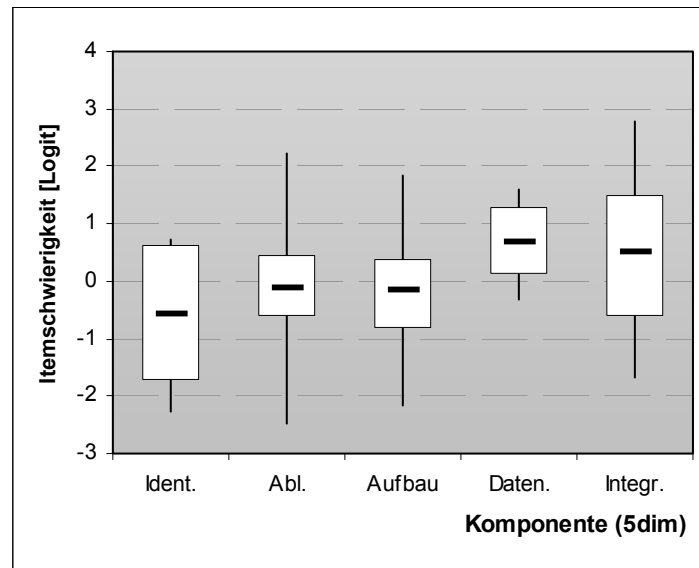


Abbildung 6.3: Mittlere Itemschwierigkeiten der Komponenten im 5dim-Modell.

Wie im Boxplot (Abbildung 6.3) anhand der großen Streuungen der Itemschwierigkeiten innerhalb der Komponenten sichtbar, überlappen die Komponenten einander stark. Entsprechend ergeben sich auch in einer Kruskal-Wallis-Varianzanalyse keine signifikanten Unterschiede zwischen den Komponenten ($\chi^2(4) = 4.552$; $p = .336$ n.s.).

6.2.3 Analyse der Binnenstrukturen der Komponenten Informationsentnahme und Konstruktion

Im theoretisch abgeleiteten Modell der Diagrammkompetenz sind innerhalb der Komponenten Informationsentnahme und Konstruktion jeweils mehrere Aspekte aufgeführt, deren Relationen zueinander in den folgenden Analysen beleuchtet werden.

Innerhalb der Informationsentnahme wird dabei die Identifizierung als ein Aspekt und die vier Ableseordnungen als weitere Aspekte formuliert. Innerhalb der Konstruktion werden die sechs Aspekte Auswahl des Diagrammtyps, Achsenbelegung, Achsenbeschriftung, Skalierung, Werteintrag und Skizzierung unterschieden. Innerhalb der Integration werden die zwei Aspekte *Resultat Text* und *Resultat Diagramm* differenziert. Eine Modellierung der Diagrammkompetenz in die sich insgesamt ergebenden 13 Aspekte bringt im Vergleich zur Modellierung in drei bzw. fünf Hauptkomponenten sowohl nach dem AIC_c -Index als auch nach dem BIC-Index keine bessere Passung, sondern erweist sich als deutlich ungeeigneter: So beträgt beispielsweise die AIC_c -Differenz zum 3dim-Modell 239 (vgl. Tabelle 6.21). Auch Modellierungen, die jeweils nur eine Komponente nach ihren Aspekten differenzieren und die anderen zwei Komponenten wie im 3dim-Modell zusammengefasst formuliert auffassen, sind nach den Modellkennwerten keinesfalls besser zu einer Modellierung der Diagrammkompetenz geeignet (vgl. Tabelle 12.12). Die

nachfolgend berichteten Vergleiche verschiedener Modelle beziehen sich deshalb jeweils auf die einzelnen Komponenten Informationsentnahme, Konstruktion und Integration.

6.2.3.1 Informationsentnahme

Wie in Abschnitt 5.3.3.2 (s. Analyse der Binnenstrukturen) beschrieben, wird ein eindimensionales Modell für die Unterkomponente Informationsentnahme (*1dim IE*), mit einem zweidimensionalen Modell (*2dim IE*), welches die beiden Unterkomponenten der Informationsentnahme unterscheidet, und mit einem fünfdimensionalen Modell (*5dim IE*), das die Identifizierung und die vier Ableseordnungen differenziert, verglichen. Die jeweiligen Modellkennwerte zeigt die nachfolgende Tabelle 6.21.

Tabelle 6.21: Kennwerte der Modellpassung. AIC_c = besondere Form des Akaike's Information Criterion, BIC: Best Information Criterion, L = Likelihoodwert.

Modell	Zahl der geschätzten Parameter	AIC_c	Differenz AIC_c	BIC	Differenz BIC	-2 log L
3dim Gesamt	62	16219	-	16412	-	16060.28086
13dim Gesamt	147	16458	-239	16688	-277	15855.30020
<i>1dim IE</i>	23	7181	-	7262	-	7131.18518
<i>5dim IE</i>	37	7151	31	7275	-14	7065.56628
<i>2dim IE</i>	25	7145	5	7232	43	7090.35684
<i>1dim Abl</i>	17	5256	-	5316	-	5220.08864
<i>4dim Abl</i>	26	5257	-1	5347	-30	5199.49298

Wie in Tabelle 6.21 ersichtlich, weist das *5dim IE* zwar nach dem AIC_c -Index eine bessere Passung gegenüber dem eindimensionalen Modell der Informationsentnahme (*1dim IE*) auf, nach dem BIC-Index kehrt sich dies um, so dass der BIC-Index stärker gewichtet und somit das sparsamere *1dim IE*-Modell vorgezogen wird. Entsprechend der Passung des *5dim*-Modells für die gesamte Diagrammkompetenz zeigt sich allerdings auch hier die beste Modellpassung für das *2dim IE*-Modell, das die beiden Unterkomponenten aufgreift.

Ausgehend von diesem Befund werden die nachfolgenden Betrachtungen weiter eingeschränkt auf die Unterkomponente Ablesen. Zur Absicherung wird auch für diese Unterkomponente ein nach den vier Ableseordnungen differenzierendes Modell (*4dim Abl*) mit einem eindimensionalen Modell zum Ablesen (*1dim Abl*) verglichen. Wie oben ist das eindimensionale Modell zur Modellierung des Ablesens besser geeignet (vgl. Tabelle 6.21), dessen WLE-Reliabilität von $Rel. = .741$ unterstützt die Eignung⁶³. Die bessere Passung des eindimensionalen Modells ist nach den Simulationsstudien von Einhaus

⁶³ Auf die Überprüfung der Personenhomogenität kann an dieser Stelle verzichtet werden, da sie bereits im Abschnitt 6.1.1.3 für die einzelnen Komponenten und Unterkomponenten durchgeführt wurde.

(2007) ein Kriterium dafür, dass es sich bei den Ableseordnungen um die Graduierung *einer* Fähigkeit handelt. Damit sollten sich die Itemschwierigkeiten über die vier Ordnungen hinweg anordnen lassen. Entsprechend sollte sich ein Ablesen höherer Ordnung jeweils als schwieriger erweisen als ein Ablesen niedrigerer Ordnung. In der folgenden Tabelle 6.22 sind die deskriptiven Statistiken der Itemparameter hinsichtlich der vier Ordnungen des Ablesens aufgeführt. In der darauf folgenden Abbildung 6.4 werden diese graphisch veranschaulicht.

Tabelle 6.22: Deskriptive Statistiken zu den Itemschwierigkeiten der vier Ordnungen des Ablesens.

Ordnung des Ablesens	N	M	SD	Median	Min	Max	Spannweite
1.Ordnung	3	0.29	1.01	0.73	-0.87	1.02	1.88
2.Ordnung	5	-0.36	0.99	0.17	-2.07	0.26	2.32
3.Ordnung	6	-0.67	1.03	-0.44	-2.48	0.38	2.86
4.Ordnung	2	1.38	1.16	1.38	0.56	2.20	1.64
Insgesamt	16	-0.14	1.15	0.13	-2.48	2.20	4.68

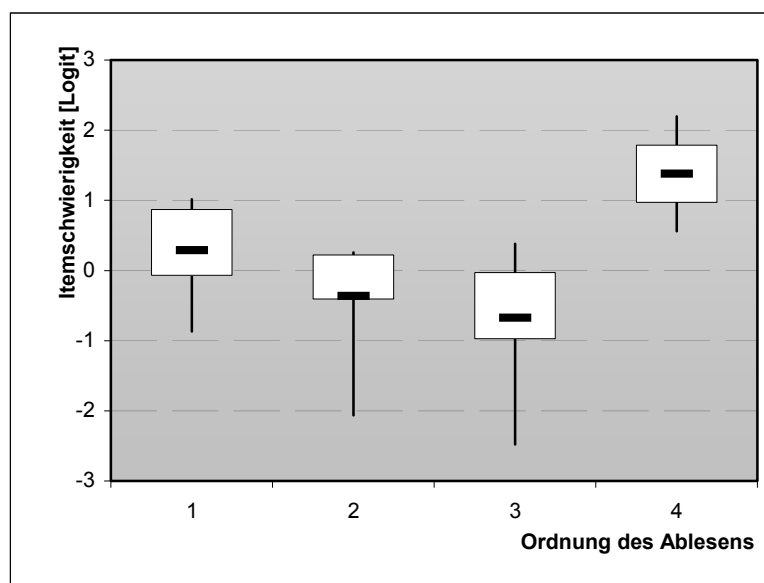


Abbildung 6.4: Boxplot zu den Schwierigkeiten der vier Ordnungen des Ablesens.

Ein Ablesen erster Ordnung erweist sich im Mittel als schwerer als das Ablesen zweiter sowie dritter Ordnung. Ebenso liegt die Schwierigkeit des Ablesens zweiter Ordnung im Mittel über der des Ablesens dritter Ordnung. Die jeweiligen Differenzen liegen im Bereich von 0.31 und 0.75 Logit, die Standardabweichungen übersteigen in allen drei Fällen diese Differenzen. Die Items zum Ablesen vierter Ordnung hingegen zeigen im Mittel eine deutlich höhere Schwierigkeit als die Items der drei anderen Ableseordnungen. Dennoch belegt ein H-Test, dass insgesamt die Differenzen der Mittelwerte zwischen allen Ordnungen nicht signifikant sind ($\chi^2 = 5.837$, $df = 3$, $p = .120$). Daher können die

berichteten Unterschiede nur als Hinweise gedeutet werden. Es kann keine Stufung der Ableseanforderungen der ersten drei Ordnungen festgestellt werden.

6.2.3.2 Konstruktion

Zur Analyse der Binnenstruktur der Konstruktion wurde, wie in Abschnitt 5.3.3.2 (s. Analyse der Binnenstrukturen) aufgeführt, ein Modell berechnet, das sechs der im Kompetenzstrukturmodell formulierten Konstruktionsaspekte aufgreift (*6dim K*). Dieses wurde sowohl einem eindimensionalen Modell zur Konstruktion (*1dim K*) als auch einem Modell, das die zwei Unterkomponenten der Konstruktion Aufbau und Dateneintrag differenziert (*2dim K*), gegenübergestellt. Die entsprechenden Modellkennwerte sind in der Tabelle 6.23 aufgeführt.

Tabelle 6.23: Kennwerte der Modellpassung der drei Modelle zur Komponente Konstruktion. AIC_c = besondere Form des Akaike's Information Criterion, BIC: Best Information Criterion, L = Likelihoodwert.

Modell	Zahl der geschätzten Parameter	AIC _c	Differenz AIC _c	BIC	Differenz BIC	-2 log L
<i>1dim K</i>	27	7099	-	7192	-	7039,46293
<i>2dim K</i>	29	7053	46	7153	39	6988,55426
<i>6dim K</i>	47	7049	4	7203	-50	6936,62573

Der AIC_c-Index zeigt die beste Modellpassung für das Modell *6dim K* an. Die Differenz vom Modell *6dim K* zum Modell *2dim K* ist mit vier allerdings nicht bedeutsam, womit das Modell *2dim K* mit einer um 18 geringeren Parameterzahl die Konstruktion vergleichbar gut abbildet. Entsprechend weist das Modell *2dim K* nach dem BIC-Index, der die Zahl der geschätzten Parameter stärker gewichtet, eine bessere Passung als das Modell *6dim K* auf. Mit einer AIC_c-Differenz von 46 zum Modell *1dim K* ist daher das Modell *2dim K*, das die beiden Unterkomponenten der Konstruktion unterscheidet, das passendste. Wie bei der Informationsentnahme entspricht dieses Ergebnis den Befunden hinsichtlich des Gesamtmodells der Diagrammkompetenz, nach denen das Modell *5dim*, das im Vergleich zum *3dim*-Modell die Unterkomponenten differenziert, eine noch bessere Passung als das *3dim*-Modell aufweist. Demzufolge werden die weiteren Analysen zur Absicherung der Modellpassung an dieser Stelle nicht differenziert berichtet, da die Struktur der Konstruktion mit zwei Unterkomponenten bereits vorher bestätigt wird.

Da das Modell *6dim K* nicht zur Modellierung der Binnenstruktur der Konstruktion geeignet ist, könnten die sechs untersuchten Aspekte der Konstruktion Stufungen innerhalb

der Komponente Konstruktion bzw. innerhalb der beiden Unterkomponenten Aufbau und Dateneintrag darstellen. Somit werden im Folgenden die mittleren Itemschwierigkeiten der einzelnen Aspekte betrachtet. Sie werden in der Abbildung 6.5 graphisch dargestellt, die entsprechenden deskriptiven Statistiken sind im Anhang (s. Tabelle 12.13) abgedruckt.

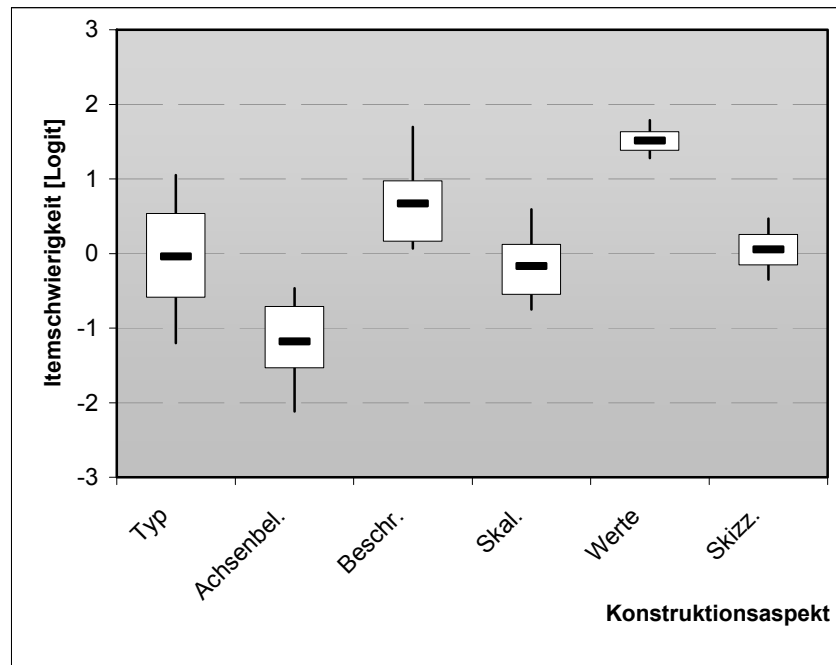


Abbildung 6.5: Boxplot zu den Schwierigkeiten der einzelnen Konstruktionsaspekte. Typ: Auswahl des Diagrammtyps, Achsenbel.: Achsenbelegung, Beschr.: Beschriftung, Skal.: Skalierung, Werte: Werteintrag, Skizz.: Skizzierung.

Insgesamt unterscheiden sich die einzelnen Aspekte der Konstruktion ebenso wie die der Informationsentnahme nicht signifikant in ihren Schwierigkeiten ($\chi^2 = 11.012$, $df = 5$, $p = .051$). Die Items der drei Aspekte Auswahl des Diagrammtyps, Skalierung und Skizzierung weisen im Mittel Schwierigkeiten nahe 0 auf. Für die Items der Achsenbelegung hingegen zeigt sich ein geringerer Mittelwert von $M = -1.18$ ($SD = 0.85$). Die richtige Belegung der Variablen stellt somit einen eher leichten Konstruktionsaspekt dar. Im Gegensatz dazu erweist sich das Eintragen von Werten in ein Diagramm als schwieriger Konstruktionsaspekt ($M = 1.52$, $SD = 0.26$). Ebenso ist die Achsenbeschriftung als eher schwierig einzustufen ($M = 0.67$, $SD = 0.89$). Die qualitative Fehleranalyse (vgl. 6.4) wird insbesondere hinsichtlich dieses Aspekts herangezogen, um Hinweise darauf zu erhalten, wodurch diese höhere Schwierigkeit bedingt wird.

6.2.3.3 Integration

Wie in 5.3.3.2 (s. Analyse der Binnenstrukturen) beschrieben, wurde ein eindimensionales Modell zur Integration (*1dim Int*) mit einem zweidimensionalen Modell, welches die zwei

postulierten Formen der Integration unterscheidet, verglichen. Wie den in der Tabelle 6.24 aufgeführten Modellkennwerten zu entnehmen ist, weist das Modell *2dim Int* sowohl nach dem AIC_c -Index als auch nach dem BIC-Index eine bessere Modellpassung auf als das *1dim Int*-Modell. Diese ist nach dem BIC-Index allerdings nicht bedeutsam (vgl. 6.1.1.1), so dass das *1dim Int*-Modell als das sparsamere Modell vorzuziehen ist. Es besitzt eine WLE-Reliabilität von $Rel. = .615$.

Tabelle 6.24: Kennwerte der Modellpassung der zwei Modelle zur Komponente Konstruktion. AIC_c = besondere Form des Akaike's Information Criterion, BIC: Best Information Criterion, L = Likelihoodwert.

Modell	Zahl der geschätzten Parameter	AIC_c	Differenz AIC_c	BIC	Differenz BIC	-2 log L
<i>1dim Int</i>	9	2122	-	2154	-	2103.39286
<i>2dim Int</i>	11	2108	14	2148	7	2085.46637

Analog zu den obigen Ausführungen werden die Itemschwierigkeiten der Einzelaspekte zur Analyse herangezogen. Aufgrund der geringeren Itemzahl der Komponente Integration werden in der folgenden Abbildung die Itemschwierigkeiten der einzelnen Items dargestellt.

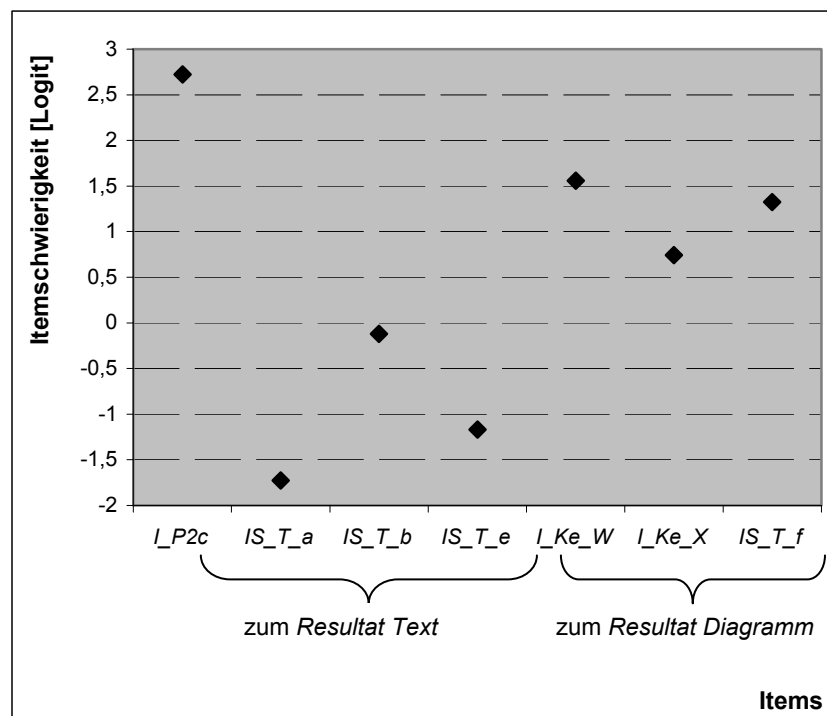


Abbildung 6.6: Schwierigkeiten der einzelnen Items zur Integration. Die Items *I_P2c*, *IS_T_a*, *IS_T_b* und *IS_T_e* bezeichnen Items, die dem Integrationsaspekt *Resultat Text* zugeordnet sind. Die Items *I_Ke_W*, *I_Ke_x* und *IS_T_f* sind Items zum Aspekt *Resultat Diagramm*.

Von links nach rechts stellen die ersten vier Items in der Abbildung 6.6 Items dar, die die Integration in Form von Text abprüften, die letzten drei Items hingegen sind Items, die als Integrationsresultat die Konstruktion von Teilen eines Diagramm forderten. Analog zu den einzelnen Ableseordnungen streuen die Itemschwierigkeiten der beiden Integrationsaspekte, vor allem für das *Resultat Diagramm*, stark (Standardabweichungen $SD = 2.0$ für den Aspekt *Resultat Text* und $SD = 0.42$ für den Aspekt *Resultat Diagramm*). Die mittlere Itemschwierigkeit der Items zum *Resultat Text* liegt dabei mit $M = -0.07$ Logit unter der mittleren Schwierigkeit der Items zum *Resultat Diagramm* mit $M = 1.21$, jedoch ist dieser Unterschied statistisch nicht signifikant, wie ein U-Test belegt ($Z(4,3) = -1.061$, $p = .289$ n.s.). Die zugehörige deskriptive Statistik der Itemschwierigkeiten getrennt nach Aspekten der Integration ist im Anhang (s. Tabelle 12.14) aufgeführt.

6.2.4 Schwierigkeiten der Diagrammtypen

Mittels eines t-Tests wurden die Itemschwierigkeiten der Gruppen „Liniendiagramm“ und „Säulendiagramm“ verglichen. Die Itemschwierigkeiten der Items, die anhand eines Liniendiagramms bearbeitet wurden, lagen mit $M = 0.04$ ($SD = 1.13$) im Mittel höher als die der Säulendiagramm-Items mit $M = -0.05$ ($SD = 1.04$). Die Differenz dieser Mittelwerte ist allerdings sehr gering und zudem nicht signifikant ($t(47) = .284$, $p = .777$).

Aufgeschlüsselt nach den Komponenten des 3dim-Modells zeigt sich beim Vergleich der Schwierigkeiten der beiden Diagrammtypen im Hinblick auf die Integration ein größerer Unterschied (vgl. Abbildung 6.7). So sind die Integrationsitems, die sich auf ein Liniendiagramm beziehen, im Mittel mit $M = 1.33$ ($SD = 0.73$) um 1,66 Logits schwerer als die Items zur Integration, die anhand eines Säulendiagramms bearbeitet wurden ($M = -0.30$, $SD = 1.10$). Bezogen auf die Komponenten Informationsentnahme und Konstruktion wurden keine nennenswerten Differenzen zwischen den Items der verschiedenen Diagrammtypen festgestellt. Die entsprechenden deskriptiven Statistiken sind im Anhang (s. Tabelle 12.15) aufgeführt.

Die eben genannten Unterschiede wurden aufgrund der geringen Itemzahlen in den einzelnen Untergruppen nicht auf Signifikanz geprüft. Die Darstellungen sind deshalb als Hinweise zu bewerten. Insbesondere ist hinsichtlich der Integration anzumerken, dass sich die Items der beiden Diagrammtypgruppen nicht direkt vergleichen lassen. Sie unterscheiden sich in den Kombinationen der Anforderungen in Bezug auf die Aspekte der Informationsentnahme und der Konstruktion (vgl. 5.2.4.1). In Folge dessen kann der Diagrammtyp nicht als einziges schwierigkeitsbestimmendes Merkmal angenommen werden. Auch aus diesem Grund ist die Aussagekraft dieses Vergleichs der Diagrammtypen eingeschränkt.

Ein exemplarischer Vergleich zweier Items (*I_Ke_W* und *IS_P_f*), die beide als Integrationsleistung das Eintragen einzelner Werte in ein Diagramm abfragten, kann jedoch als Hinweis darauf gewertet werden, dass sich, wie die Analysen über alle Items hinweg zeigten, die Itemschwierigkeiten nicht in Abhängigkeit vom Diagrammtyp unterscheiden: So ergaben sich für das Liniendiagramm mit einer Schwierigkeit von 1.560 Logit und für das Säulendiagramm mit einer Schwierigkeit von 1.324 Logit sehr ähnliche Schwierigkeitswerte.

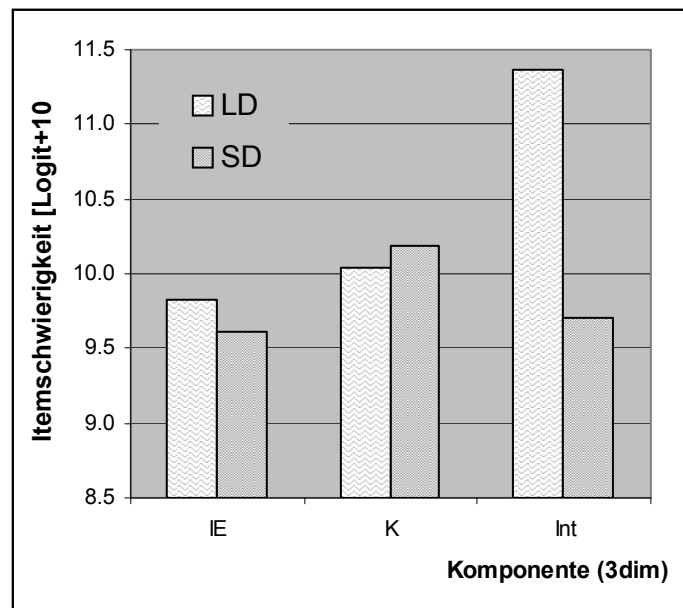


Abbildung 6.7: Schwierigkeiten der Items zu verschiedenen Diagrammtypen aufgeschlüsselt nach Komponenten des *3dim*-Modells. IE: Informationsentnahme, K: Konstruktion, Int: Integration. Zur einfacheren Darstellung sind die Personenparameter um 10 nach oben verschoben.

6.3 Lernerhintergrund und Lernermerkmale

Den Forschungsfragen entsprechend werden im folgenden Abschnitt zunächst die Analysen zum Einfluss des Lernerhintergrunds dargestellt (6.3.1). Im Anschluss daran wird der Einfluss der verschiedenen Lernermerkmale analysiert (6.3.2).

6.3.1 Einfluss des Lernerhintergrunds

6.3.1.1 Unterschiede in Bezug auf die Personenfähigkeiten

Der Nachweis der Personenhomogenität sowohl für das *3dim*- als auch das *5dim*-Modell in Bezug auf die jeweiligen Teilstichproben bedeutet insbesondere, dass für Schüler verschiedener Klassenstufen respektive Schülerinnen und Schüler die gleichen Strukturen der Diagrammkompetenz gefunden wurden. Die Kompetenzstruktur ist somit für jede der einzelnen Gruppen gültig. Ob sich die Gruppen bezüglich ihrer Fähigkeiten unterscheiden,

wird nachfolgend analysiert. Zuerst sind die deskriptiven Statistiken der zwei Klassenstufen 9 und 10 angegeben. Abhängige Variablen sind dabei die Personenfähigkeiten (WLEs) in den Komponenten des *3dim*-Modells sowie des *5dim*-Modells.

Tabelle 6.25: Deskriptive Statistik der Personenfähigkeiten in den Komponenten des Modells *3dim* getrennt nach Klassenstufen.

Komponente (<i>3dim</i>)	Klassenstufe	N	M	SD	Median	Min	Max	Spannweite
Informationsentnahme	9	134	-0.04	1.06	0.05	-3.29	3.18	6.47
	10	155	0.50	1.03	0.49	-2.65	3.18	5.83
Konstruktion	9	134	-0.35	1.52	0.15	-4.26	1.96	6.22
	10	155	-0.11	1.45	0.15	-4.26	2.65	6.91
Integration	9	134	-1.01	1.83	-0.33	-3.71	3.35	7.06
	10	155	-0.05	1.60	0.50	-3.71	2.54	6.25

Tabelle 6.26: Deskriptive Statistik der Personenfähigkeiten in den Komponenten des Modells *5dim* getrennt nach Klassenstufen.

Komponente (<i>5dim</i>)	Klassenstufe	N	M	SD	Median	Min	Max	Spannweite
Identifizierung	9	134	0.28	1.22	0.05	-3.32	3.20	6.51
	10	155	0.72	1.15	0.91	-1.85	3.20	5.04
Ablesen	9	134	-0.17	1.21	-0.26	-4.21	4.10	8.31
	10	155	0.40	1.27	0.63	-4.21	2.83	7.04
Aufbau	9	134	-0.09	1.62	0.36	-3.91	3.02	6.93
	10	155	0.26	1.65	0.59	-3.91	4.35	8.26
Dateneintrag	9	134	-0.59	1.19	-0.69	-2.80	2.68	5.48
	10	155	-0.65	1.14	-0.69	-2.80	1.68	4.47
Integration	9	134	-1.02	1.83	-0.34	-3.72	3.38	7.10
	10	155	-0.05	1.61	0.49	-3.72	2.57	6.28

Wie in Tabelle 6.25 aufgetragen, zeigen die Zehntklässler in allen drei Komponenten des *3dim*-Modells im Mittel höhere Personenfähigkeiten als die Neuntklässler. Die größten Unterschiede bestehen dabei in Bezug auf die Integration, bei der die Gruppenmittelwerte um beinahe ein Logit differieren. Einzelne t-Tests zwischen den Gruppen belegen, dass die Unterschiede im Modell *3dim* zum einen in Bezug auf die Integration signifikant sind ($t(289) = -4.72$, $p = .000$), zum anderen in Bezug auf die Informationsentnahme ($t(289) = -4.34$, $p = .000$). Damit können Zehntklässler besser Informationen aus Diagrammen entnehmen und deutlich besser Informationen aus Text und Diagramm integrieren.

Dieses Ergebnis wird durch die Aufteilung der Informationsentnahme im Modell *5dim* bestätigt (s. Tabelle 6.26): Entsprechend erweisen sich die Unterschiede zwischen den Gruppen in Bezug auf beide Unterkomponenten der Informationsentnahme (Identifizierung: $t(289) = -3.102$, $p = .002$, Ablesen: $t(289) = -3.85$, $p = .000$) als signifikant. Die Überlegenheit der Zehntklässler im Hinblick auf die Informationsentnahme bezieht sich folglich auf beide Unterkomponenten Identifizierung und Ablesen. Ebenso zeigt sich in Hinblick auf die Integration eine Überlegenheit der Zehntklässler: Ihre Personenfähigkeiten lagen signifikant über denen der Neuntklässler ($t(289) = -4.723$, $p = .000$).

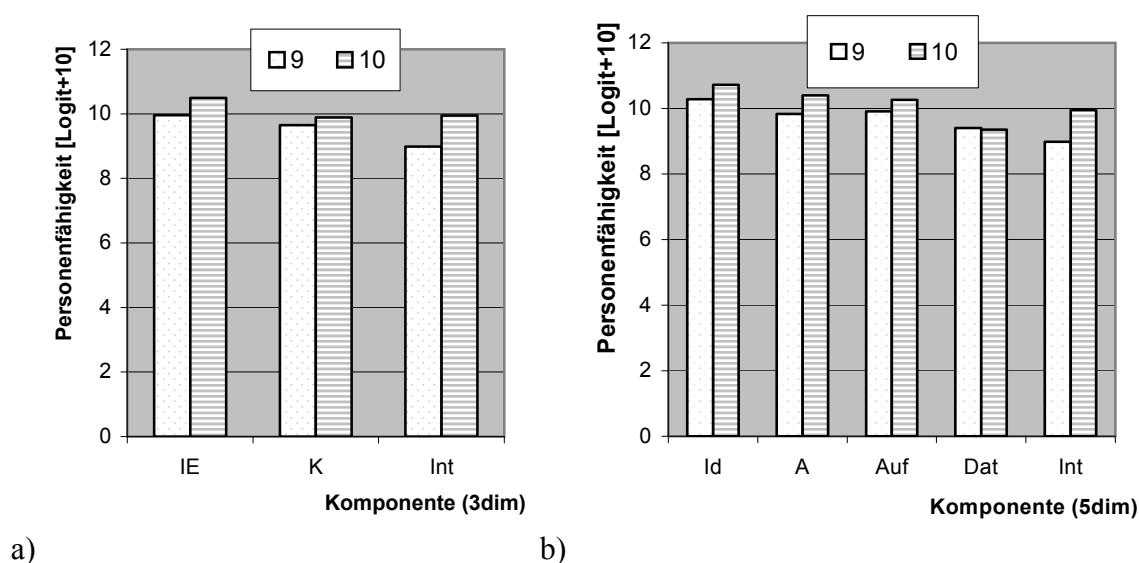


Abbildung 6.8: Vergleich der Personenfähigkeiten nach Klassenstufen für die Komponenten des *3dim*-Modells (a) und die Komponenten des *5dim*-Modells (b). Zur einfacheren Darstellung sind die Personenparameter um 10 nach oben verschoben.

Hinsichtlich des Geschlechts zeigen sich weder in Bezug auf die Komponenten des *3dim*-Modells noch auf die feineren Komponenten des *5dim*-Modells signifikante Unterschiede zwischen den Schülerinnen und Schülern. Die Differenzen der Mittelwerte der beiden Gruppen liegen in Bereichen zwischen 0.14 und 0.28 Logit. Die deskriptiven Statistiken zu diesem Gruppenvergleich sind im Anhang (s. Tabelle 12.16) aufgeführt.

6.3.1.2 Unterschiede in Bezug auf die Relationen der Komponenten

In den folgenden Analysen wird überprüft, ob sich zentrale der bisher ermittelten Zusammenhänge zwischen den Komponenten in bestimmten Teilstichproben wiederfinden oder ob sie sich innerhalb der Teilstichproben anders darstellen. Die dazu notwendige Prüfung der Personenhomogenität in Bezug auf die Teilstichproben wurde bereits in Abschnitt 6.1.1.3 berichtet. Die Ergebnisse zeigen, dass sowohl für die Klassenstufe als auch das Geschlecht Personenhomogenität gewährleistet ist.

Der Zusammenhang zwischen den Komponenten Informationsentnahme und Konstruktion liegt für die gesamte Stichprobe bei $r = .435$. Für die Teilstichprobe der neunten Klassenstufe zeigt sich eine Korrelation von $r = .427$, für die der zehnten Klassenstufe eine Korrelation von $r = .432$. Für die Neuntklässler zeigt sich somit ein ähnlicher Zusammenhang zwischen Informationsentnahme und Konstruktion wie für die Zehntklässler. Hinsichtlich der Teilung der Stichprobe nach Geschlecht ergaben sich größere Unterschiede. So wird für die Schüler ein Zusammenhang von $r = .408$ berechnet. Bei den Schülerinnen korrelieren Informationsentnahme und Konstruktion mit $r = .483$. Bei der Signifikanzprüfung über die Fisher-transformierten Z-Werte (vgl. 5.3.3.3, s. Unterschiede in Bezug auf die Relationen der Komponenten) erwiesen sich allerdings die Gruppenunterschiede hinsichtlich der genannten Korrelationen für beide Stichprobenteilungen als nicht signifikant. Bezogen auf die verschiedenen Klassenstufen ergab sich $p = .226$, bezogen auf das Geschlecht wurde $p = .479$ ermittelt. Insgesamt bedeutet dies, dass der für das Kompetenzstrukturmodell beschriebene Zusammenhang zwischen Informationsentnahme und Konstruktion sowohl klassenstufen- als auch geschlechtsübergreifend gilt. Auch bezüglich der weiteren in Abschnitt 6.2.1 beschriebenen Zusammenhänge konnten in Abhängigkeit von Klassenstufe und Geschlecht keine signifikanten Unterschiede festgestellt werden. Insgesamt belegen diese Ergebnisse, dass sich die beschriebenen Relationen zwischen den Komponenten und Unterkomponenten sowohl für beide Klassenstufen als auch für beide Geschlechter gleich darstellen.

6.3.2 Einfluss der Lernermerkmale

Zur Analyse der Zusammenhänge zwischen den in Abschnitt 3.7.2 vorgestellten Lernermerkmalen und den einzelnen (Unter-)Komponenten wurden in zwei Ansätzen mehrere schrittweise Regressionen berechnet, die zum Überblick in Tabelle 6.27 sowie in Tabelle 6.28 vorgestellt werden. Dabei werden jeweils die Betakoeffizienten der signifikanten Prädiktoren sowie das korrigierte Bestimmtheitsmaß der Regressionsgleichungen aufgeführt⁶⁴.

Die beiden Regressionsansätze **S** und **E** unterscheiden sich dahingehend, ob die **Summenscores** für das Vorwissen bezüglich biologischer Inhalte (VW_Bio) und die Fähigkeiten zum Umgang mit abstrakten Diagrammen (UAD) oder die jeweiligen **Einzelitems** als Variablen verwendet wurden. Aufgrund geringer Reliabilitäten für die beiden Skalen (VW_Bio: Cronbachs $\alpha = .231$, UAD: Cronbachs $\alpha = .114$) ist die Betrachtung der Einzelitems sinnvoll. Insbesondere zeigt das korrigierte Bestimmtheitsmaß im Ansatz E mit einer Ausnahme stärkere Effekte an, so dass an

⁶⁴ Für die Berichterstattung an dieser Stelle wird den Ergebnissen bezüglich der Unterkomponenten, die im *5dim*-Modell formuliert sind, der Vorzug gegeben, da sie nicht von den Ergebnissen bezüglich des *3dim*-Modells abweichen, sondern diese im Gegenteil aufzuschlüsseln vermögen.

geeigneten Stellen auf Ansatz E mit der Unterscheidung der Einzelitems zurückgegriffen wird. Werden Beta-Koeffizienten ohne weitere Erläuterung aufgegriffen, beziehen sie sich auf den Ansatz S.

Tabelle 6.27: Ansatz S (mit Summenscores für das biologische Vorwissen (VW_Bio) und die Fähigkeiten zum Umgang mit abstrakten Diagrammen (UAD)) der Regressionen zum Einfluss der Lernermerkmale. * $p < .05$, ** $p < .01$, *** $p < .001$.

Abhängige Variable	Identifizierung	AbleSEN	Aufbau	Dateneintrag	Integration
Prädiktor	β	β	β	β	β
Anstrengung			.267***	.266***	
Lesekompetenz	.164**		.136*	.184**	
Rechnen		.122*			
Biologienote					
Mathematiknote			.149*	.175**	
VW_Bio	.167**				
UAD	.420***	.313***	.145*		.297***
	$R^2_{\text{korr.}} = .273$	$R^2_{\text{korr.}} = .115$	$R^2_{\text{korr.}} = .164$	$R^2_{\text{korr.}} = .155$	$R^2_{\text{korr.}} = .074$

Tabelle 6.28: Ansatz E (mit Einzelitems für das biologische Vorwissen (VW_Bio) und die Fähigkeiten zum Umgang mit abstrakten Diagrammen (UAD)) der Regressionen zum Einfluss der Lernermerkmale. * $p < .05$, ** $p < .01$, *** $p < .001$.

Abhängige Variable	Identifizierung	AbleSEN	Aufbau	Dateneintrag	Integration
Prädiktor	β	β	β	β	β
Anstrengung		.170**	.297***	.266***	.178**
Lesekompetenz	.208***		.156**	.184**	
Rechnen		.117*			
Biologienote					
Mathematiknote			.160**	.175**	
<i>pop2</i>	.119*				
<i>rb2</i>	.125*				
<i>konk1</i>		.109*			
<i>uad10</i>	.225***				
<i>uad11</i>					
<i>uad6</i>					
<i>S_uad3</i>	.444***	.477***	.186**		.525***
	$R^2_{\text{korr.}} = .329$	$R^2_{\text{korr.}} = .278$	$R^2_{\text{korr.}} = .179$	$R^2_{\text{korr.}} = .155$	$R^2_{\text{korr.}} = .286$

Hinweis: *pop2*, *rb2*, *konk1* bezeichnen Items zum biologischen Vorwissen, *uad10*, *uad11*, *uad6* und *S_uad3* bezeichnen Items zum Umgang mit abstrakten Diagrammen.

Wie in Tabelle 6.27 ersichtlich, stellt sich die aufgewendete **Anstrengung** hinsichtlich beider Unterkomponenten der Konstruktion, Aufbau und Dateneintrag, als signifikanter Prädiktor heraus. Schüler, die ihre während der Testbearbeitung aufgebrauchte Anstrengung selbst hoch einschätzten, weisen höhere Fähigkeiten beim Konstruieren (in beiden Unterkomponenten) auf als Schüler, die ihre Anstrengung als gering einschätzten. Mit $\beta = .267$ und $\beta = .266$ ist die Anstrengung dabei der stärkste Vorhersagewert hinsichtlich dieser Unterkomponenten. Im Regressionsansatz E (s. Tabelle 6.28) zeigt sich die Anstrengung darüber hinaus als signifikanter Prädiktor für das Ablezen und die Integration. Im Vergleich zu den Zusammenhängen zwischen der Anstrengung und den Unterkomponenten der Konstruktion sind die Zusammenhänge zu Ablezen und Integration allerdings geringer ($\beta = .170$ bzw. $\beta = .178$).

Ein Einfluss der **Lesekompetenz** wurde in beiden Ansätzen für die Identifizierung, den Aufbau und den Dateneintrag ermittelt. Insbesondere wurde damit kein Zusammenhang zwischen der Lesekompetenz (auf Texte bezogen) und dem Ablezen von Diagrammen festgestellt. Ebenso ist die Lesekompetenz der Schüler kein Prädiktor für die Integration. Schüler mit höherer Lesekompetenz zeigten somit bessere Leistungen im Hinblick auf die Identifizierung und Konstruktion von Diagrammen als Schüler mit geringer Lesekompetenz, jedoch sind sie letzteren nicht im Hinblick auf das Ablezen und die Integration überlegen.

Die **Rechenfähigkeit** der Schüler zeigt sich als signifikanter Prädiktor für das Ablezen ($\beta = .122$). Darüber hinaus wurde in keinem der beiden Regressionsansätze für eine weitere der (Unter-)Komponenten ein Effekt ermittelt. Bezogen auf die letzte **Zeugnisnote in Mathematik** zeigt sich ein Zusammenhang zu den beiden Konstruktionsunterkomponenten Aufbau und Dateneintrag. Die entsprechenden Betakoeffizienten liegen dabei vom Betrag her in der Größenordnung derjenigen bezüglich der Lesekompetenz.

Die **Biologienote** des letzten Zeugnisses erweist sich für keine Komponente als signifikanter Vorhersagewert, auch bezüglich des **biologischen Vorwissens** zeigt sich im Regressionsansatz S (s. Tabelle 6.27) lediglich in Bezug auf die Identifizierung ein Zusammenhang. Im Ansatz E (s. Tabelle 6.28) ist erkennbar, dass sich dieser Zusammenhang auf das Item *pop2*, welches die Definition der Populationsgröße abfragt, und auf das Item *rb2*, welches die erste Lotka-Volterra-Gesetzmäßigkeit erfragt, stützt; beide Items sind signifikante Prädiktoren für die Identifizierung. Somit erzielten die Schüler, die über Kenntnisse über Populationen und Populationsschwankungen verfügten, bessere Fähigkeitswerte im Bereich der Identifizierung als Schüler ohne diese Kenntnis. Damit scheint das biologische Vorwissen, wie hier die Kenntnis des Begriffs der Populationsgröße, vor allem dabei nützlich zu sein, mithilfe der Identifizierung zu erfassen, welcher Zusammenhang in einem Diagramm dargestellt wird. Darüber zeigt sich auf der

Ebene der Einzelitems das Item *konkl*, welches auf ein Verständnis von Konkurrenz unter Tieren abzielt, als Vorhersagewert für das Ablesen.

Als starker Prädiktor für vier der fünf (Unter-)Komponenten wurden die Fähigkeiten zum **Umgang mit abstrakten Diagrammen** (UAD, s. Tabelle 6.27) ermittelt. Schüler mit höheren Fähigkeiten zum Umgang mit abstrakten Diagrammen erhielten somit höhere Fähigkeitswerte in Bezug auf die Identifizierung, das Ablesen, den Aufbau und die Integration als Schüler mit geringen Fähigkeiten zum Umgang mit abstrakten Diagrammen. Der Regressionsansatz E (s. Tabelle 6.28) zeigt dabei, dass das Item *S_uad* (s. 12.3.2.4), welches ein Ablesen höherer Ordnung im abstrakten Diagramm erfragt, das ausschlaggebende Item für diese Wirkung ist. Hier liegen die Betakoeffizienten hinsichtlich Identifizierung, Ablesen und Integration über einem Wert von $\beta = .444$, im Vergleich dazu zeigt sich hinsichtlich des Aufbaus ein geringerer Zusammenhang von $\beta = .186$. Für die Identifizierung zeigt sich darüber hinaus das Item *uad10* ($\beta = .225$), welches auf abstrakte Weise Wissen zur abhängigen Variable abfragt, als signifikanter Vorhersagewert.

Weitere Ergebnisse

Neben der Selbsteinschätzung der aufgewendeten Anstrengung kann die Häufigkeit, mit der die Versuchspersonen einzelne Items ausgelassen haben, als Hinweis auf die Anstrengungsbereitschaft gewertet werden. Die folgende Tabelle zeigt für die drei Komponenten des 3dim-Modells über die entsprechenden Items hinweg die mittlere Zahl der Personen, die ein Item der Komponente ausgelassen haben.

Tabelle 6.29: Deskriptive Statistik für die Auslassungen von Items zu den Komponenten des 3dim-Modells.

Komponente (3dim)	N	M	SD	Median	Min	Max	Spannweite
Informationsentnahme	19	60,84	55,53	34	5	146	141
Konstruktion	24	74,58	37,19	70	2	137	135
Integration	4	51,25	10,37	48,5	42	66	24
Insgesamt	47	67,04	44,41	65	2	146	144

Die Tabelle zeigt, dass die Items zur Konstruktion im Mittel am häufigsten ausgelassen wurden. Die Items zur Integration wurden im Mittel von 20 Personen mehr bearbeitet als die Konstruktionsitems.

Insgesamt ging die Häufigkeit der Bearbeitungen zum Testende hin deutlich zurück. Es zeigt sich eine hoch signifikante Korrelation zwischen der Zahl der Bearbeitungen und der Position der Items im Test ($r = .757^{**}$): Je weiter hinten das Item im Test angeordnet war, desto weniger Schüler haben es bearbeitet.

6.4 Qualitative Fehleranalyse

In diesem Abschnitt werden einzelne markante Fehler, die bei der Beantwortung der offenen Items der Konstruktion aufgetreten sind, vorgestellt und jeweils den im Kompetenzstrukturmodell formulierten Aspekten der Konstruktion zugeordnet. Überwiegend werden die Schülerantworten der Hauptstudie herangezogen. Davon abweichende Fälle sind entsprechend gekennzeichnet.

Zur Berechnung der Häufigkeit, mit der die Fehler aufgetreten sind, wird die Zahl der Schüler, die die entsprechende Aufgabe nicht ausgelassen haben, als Referenzwert ermittelt. Diese Zahlen sind in Tabelle 12.19 im Anhang aufgeführt.

6.4.1 Auswahl des Diagrammtyps

Wie in Abschnitt 3.2.2 erläutert wurde, können Daten, denen eine metrische unabhängige Variable zugrunde liegt, auch in Form eines Säulendiagramms visualisiert werden. Solche Daten waren in der Aufgabenstellung zum Diagramm *Kiefernspinner* (vgl. 5.2.4.1) gegeben. Allerdings beschrieb der Aufgabentext kontinuierliche Änderungen der Schlupfrate, so dass in diesem Fall ein Liniendiagramm der geeignetere Diagrammtyp war als das Säulendiagramm. Es zeigte sich, dass die Angemessenheit einer Interpolationslinie zur Darstellung des im Text beschriebenen Zusammenhangs zwischen Umgebungstemperatur und Schlupfrate des Kiefernspinners von 96 % der Schüler richtig erkannt wurde. Nur 4 % der Schüler, die diese Aufgabe bearbeitet haben, haben ein Säulendiagramm oder andere Diagrammtypen als ein Liniendiagramm gewählt (ein Beispiel dafür zeigt die Abbildung 6.9 zum Aspekt der Achsenbeschriftung).

Im umgekehrten Fall trat eine *fehlerhafte* Auswahl eines Liniendiagramms, wie in Abbildung 6.9 beispielhaft dargestellt, in 8 bis 13% der Diagrammkonstruktionen auf. In diesen Fällen gab die Aufgabenstellung eine nominale unabhängige Variable vor, so dass das Zeichnen einer Interpolationslinie und somit die Auswahl eines Liniendiagramms nicht angemessen war (vgl. 3.2.2).

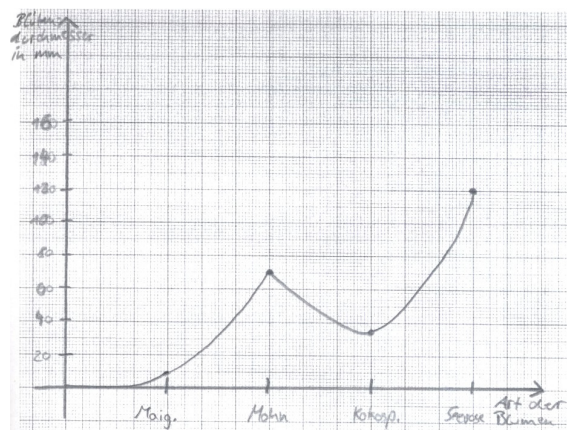


Abbildung 6.9: Beispiel für die fälschliche Auswahl eines Liniendiagramms.

6.4.2 Achsenbelegung

Das folgende Beispiel illustriert Schwierigkeiten bei der Konstruktion des Diagramms *Geschlechtsreife*. Aufgabe der Schüler war, ein Diagramm zu erstellen, das die Zeit bis zur Geschlechtsreife dreier verschiedener Tierarten darstellt. Für die Achsenbelegung waren somit die zwei Variablen „Tierart“ und „Zeit bis zur Geschlechtsreife“ relevant.

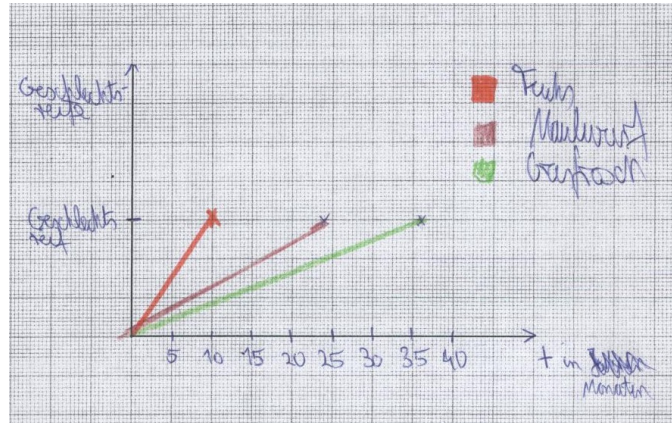


Abbildung 6.10: Beispiel für eine fehlerhafte Achsenbelegung.

Wie die Abbildung 6.10 zeigt, scheint die Belegung der Achsen mit den korrekten Variablen den Schülern bei dieser Aufgabe Schwierigkeiten bereitet zu haben. So wurde die „Zeit“ als metrisch skalierte Variable auf der Abszisse eingetragen, der Zustand „Geschlechtsreife“ als eine scheinbar metrisch skalierte Variable auf der Ordinate. Dieser Fehler trat in der Vorstudie in 32 % der Bearbeitungen auf. Die Schüler der Hauptstudie hingegen hatten mit der Achsenbelegung zum Diagramm *Geschlechtsreife* mehrheitlich keine Probleme. Ebenso traten bei den Bearbeitungen der anderen Konstruktionsaufgaben keine falschen Achsenbelegungen im Sinne der Auswahl der falschen Variablen auf. Allerdings wurde in 10 bis 23 % der Fälle die UV auf der Ordinate und die AV auf der Abszisse eingetragen.

6.4.3 Achsenbeschriftung

Zu einer vollständigen Achsenbeschriftung wurde die Angabe beider Variablen mit zugehöriger Einheit erwartet. Es zeigte sich, dass in 12 bis 16 % der Bearbeitungen zwar Achsenbeschriftungen vorgenommen wurden, jedoch an mindestens einer der Achsen die Einheit fehlte. Andererseits wurden in 20 bis 26 % der Fälle nur die Einheit und keine weitere Variablenkennzeichnung an einer oder beiden Achsen eingetragen, wie in Abbildung 6.11 illustriert ist. Im Fall des Diagramms Geschlechtsreife waren es 52 % der Fälle. Dort haben viele Schüler die Achse der Variable „Zeit bis zur Geschlechtsreife“ beispielsweise lediglich durch die Angabe „Monate“ beschriftet.

Über die fehlenden Teile der Beschriftungen hinaus wurden inhaltliche Mängel in Bezug auf die Benennung der Variablen festgestellt. Beispiele dafür sind „Blüte“ an Stelle von „Blütendurchmesser“ oder „Eier“ an Stelle von „Schlupfrate“. Ungenaue inhaltliche Bezeichnungen wie diese traten in 19 bis 51 % der vorgenommenen Beschriftungen auf.

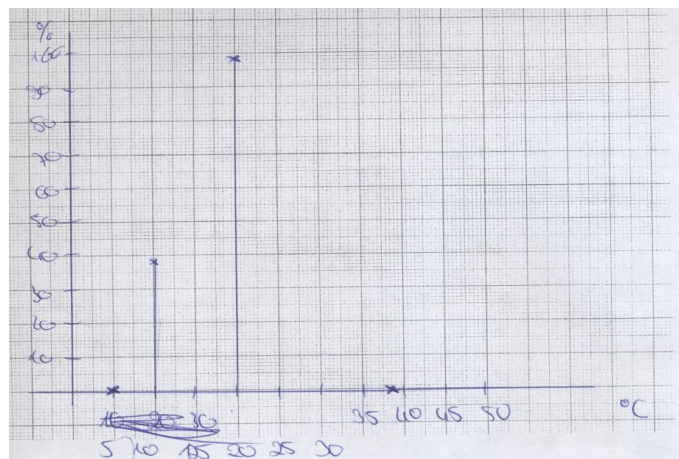


Abbildung 6.11: Beispiel für Achsenbeschriftungen nur durch die Angabe der Einheiten der Variablen.

6.4.4 Skalierung

In Abschnitt 3.5.2.2 wurden empirische Befunde hinsichtlich der Skalierung aufgeführt. In Bezug auf metrisch skalierte Variablen wurde dabei beschrieben, dass Schüler häufig Skalen erstellen, die eine Aufreihung der im Text genannten Werte darstellen. Eine solche nicht-metrische Skalierung nur aus genannten Werten, die exemplarisch in Abbildung 6.12 abgebildet wird, wurde in der vorliegenden Arbeit nur vereinzelt gefunden, in der Hauptstudie beispielsweise nur in zwei Fällen.

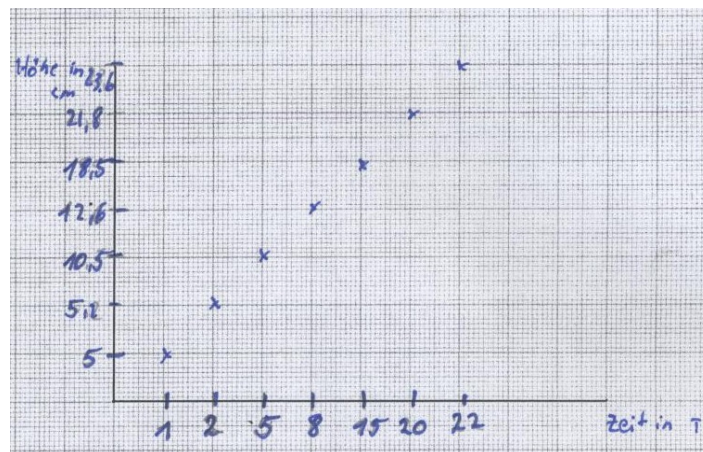


Abbildung 6.12: Beispiel für die Erstellung einer Skala, die nur die im Text gegebenen Werte auflistet. Aufgabe war es, die Höhe einer Pflanze über eine bestimmte Beobachtungszeit darzustellen.

Dem gegenüber trat im Hinblick auf die Erstellung einer metrischen Skala die Rubrik „Skalierung ungleichmäßig“ in der Hauptstudie in 36 % bis 63 % aller Schülerantworten je Item auf. Als ungleichmäßige Skalierung wird der Fall bezeichnet, dass die Abstände zwischen den Skalenunterteilungen ungleichmäßig beschriftet waren, beispielsweise als 5, 10, 20, 25.

6.4.5 Werteintrag

Die Abbildung 6.13 zeigt am Beispiel des Diagramms *Rentiere* (vgl. 5.2.4.1) eine typische Schülerantwort hinsichtlich des Werteintrags. In diesem Fall war die Aufgabe, die vier im Text genannten Punkte (1944,24), (1957,1350), (1963,6000) und (1966,42) im vorgegebenen Diagrammrahmen zu markieren. In Bezug auf die Jahreszahlen sind die Werte, wie in der Abbildung zu erkennen, bis auf eine Ausnahme einigermaßen korrekt. Die Ausnahme stellt das Jahr 1944 dar, der entsprechende Punkt wurde zwischen 1944 und 1945 eingetragen. Dieser Fehler trat insgesamt in 13 % aller Bearbeitungen dieses Konstruktionsitems auf. Hinsichtlich der Werte Rentierzahlen zeigen sich größere, jedoch nicht gehäuft auftretende oder systematische Abweichungen. Wenngleich insgesamt ein großer Teil der Schüler (65 bis 91%) keine volle Punktzahl bei den Items zum Werteintrag erhielt, war in bis zu 42 % dieser Fälle lediglich einer von drei respektive vier Punkten nicht exakt eingetragen.

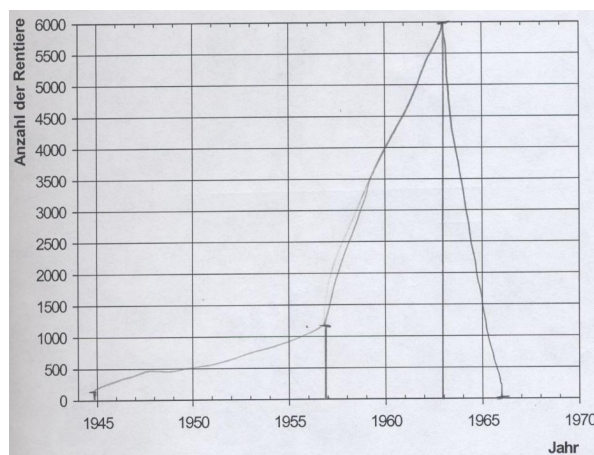


Abbildung 6.13: Beispiel für eine Antwort zum Werteintrag im Diagramm *Rentiere*.

7 Diskussion

Die Ergebnisse der vorliegenden Arbeit sind in Kapitel 6 in Bezug auf die detaillierten Forschungsfragen dargestellt worden und werden im Folgenden vernetzt diskutiert. Dabei wird zunächst die Struktur der Diagrammkompetenz in Bezug auf die Komponenten und ihre Relationen erörtert (7.1). Weitere Teile der Diskussion sind den Binnenstrukturen der Komponenten (7.2) sowie der Unterscheidung der beiden Diagrammtypen Linien- und Säulendiagramm (7.3) gewidmet. Im Anschluss daran werden die Ergebnisse zu den verschiedenen Einflussfaktoren, Lernerhintergrund und Lernermerkmalen, kritisch beleuchtet (7.4). Schließlich wird das für die vorliegende Arbeit entworfene Testinstrument diskutiert (7.5), bevor sich aus den Ergebnissen dieser Arbeit ergebende Implikationen für die Unterrichtspraxis ausgeführt werden (7.6).

7.1 Komponenten der Diagrammkompetenz und deren Relationen

In dem in der vorliegenden Arbeit vorgeschlagenen Strukturmodell werden Informationsentnahme, Konstruktion und Integration als Komponenten der Diagrammkompetenz formuliert (vgl. 3.6). Darüber hinaus wird eine Unterteilung der Informationsentnahme und der Konstruktion in jeweils zwei Unterkomponenten postuliert. Sowohl Komponenten als auch Unterkomponenten konnten in der empirischen Untersuchung nachgewiesen werden (vgl. 6.1.1.5), womit beide in Abschnitt 4.1.1 genannten Hypothesen bestätigt werden konnten. Dies bedeutet, dass ein Modell aus drei respektive fünf (Unter-)Komponenten zur Erfassung der Diagrammkompetenz von Schülern geeignet ist.

Neben der empirischen Überprüfung der Modellstruktur wurden detaillierte Analysen zu den Relationen der einzelnen (Unter-)Komponenten vorgenommen, deren Ergebnisse im Folgenden diskutiert werden. Zuerst werden Informationsentnahme und Konstruktion einander gegenüber gestellt (7.1.1). Im Anschluss daran wird die Integration als Komponente der Diagrammkompetenz sowie ihre Bezüge zu Informationsentnahme und Konstruktion diskutiert (7.1.2). Dabei werden zum einen die in den Korrelations- und Mediationsanalysen aufgedeckten Relationen zwischen den einzelnen Komponenten und Unterkomponenten aufgegriffen. Zum anderen werden an geeigneten Stellen die Passungen einzelner Kontrollmodelle (vgl. 6.1.1.1) herangezogen, da in diesen Kontrollmodellen durch das Zusammenfassen bestimmter (Unter-)Komponenten bereits Zusammenhänge zwischen diesen postuliert werden. Ein Beispiel für einen so postulierten Zusammenhang ist das Zusammenfassen der theoretisch differenzierten Komponenten Informationsentnahme und Integration zu einer Komponente im Modell *2dim_asy*.

7.1.1 Gegenüberstellung von Informationsentnahme und Konstruktion

Mit der Passung der Experimentalmodelle (vgl. 6.1.1.5) konnten die zwei Komponenten Informationsentnahme und Konstruktion als Strukturen der Diagrammkompetenz bestätigt werden. Grundsätzlich stellen damit Informationsentnahme und Konstruktion zwei unterscheidbare Kompetenzstrukturen dar, die bei der Diagnose und Förderung von Diagrammkompetenz im Biologieunterricht berücksichtigt werden sollten. Inwiefern diese Komponenten zusammenhängen, wurde durch verschiedene Modellberechnungen (vgl. 6.1.1) sowie Korrelationsanalysen (vgl. 6.2.1) untersucht. Dabei konnte ein Zusammenhang zwischen Informationsentnahme und Konstruktion festgestellt werden, der über die Effektstärke als mittelstark eingestuft wurde. Da es bisher in der Literatur an systematischen Untersuchungen des Zusammenhangs von Informationsentnahme und Konstruktion fehlt (Friel et al., 2001), kann dieser Befund der vorliegenden Arbeit nicht direkt mit denen anderer Untersuchungen verglichen werden.

Eine differenziertere Analyse ist allerdings über die Zusammenhänge der jeweiligen Unterkomponenten von Informationsentnahme und Konstruktion zueinander möglich. So wird im Modell *3dim_q* ein möglicher Zusammenhang dieser beiden Komponenten formuliert, indem Identifizierung und Aufbau zur Komponente *Rahmen* sowie Ablesen und Dateneintrag zur Komponente *Werte* zusammengefasst werden. Die Struktur dieses Modells spiegelt die Unterscheidung von globalen, auf den Rahmen bezogenen Diagrammschemata und lokalen, auf die einzelnen Werte bezogenen Diagrammschemata (vgl. Schnotz, 1994) wider. Die Theorie der Diagrammschemata bezieht sich allerdings lediglich auf die Entnahme von Informationen aus Diagrammen. Sie wurde in Anlehnung an die Annahme, dass Informationsentnahme und Konstruktion gewisse Parallelen aufweisen (Leinhardt et al., 1990), auf die Seite der Konstruktion übertragen. Das entsprechende Modell *3dim_q* zeigt eine deutlich geringere Modellpassung als die Modelle *2dim*, *2dim_{within}* und *3dim*, in denen die beiden Komponenten Informationsentnahme und Konstruktion gegenübergestellt werden (vgl. 6.1.1.1). Dieser geringeren Passung entsprechend, zeigten sich in den Korrelationsanalysen zwischen den Unterkomponenten der Komponenten *Rahmen* und *Werte* geringere Zusammenhänge als zwischen den jeweiligen Unterkomponenten innerhalb der Informationsentnahme respektive der Konstruktion (vgl. 6.2.1.1). Diese Befunde unterstreichen die Abgrenzung der Informationsentnahme von der Konstruktion und erhärten das Ergebnis, die beiden Komponenten als qualitativ unterschiedliche Teile der Diagrammkompetenz anzusehen.

Innerhalb der Komponenten Informationsentnahme und Konstruktion ist die Abgrenzung einer Unterkomponente, die sich auf den Rahmen bezieht, von einer Unterkomponente, die sich auf die Werte bezieht, sinnvoll. Dies lässt sich damit begründen, dass trotz der genannten hohen Korrelationen zwischen den jeweiligen Unterkomponenten von Informationsentnahme und Konstruktion mit Passung des *5dim*-Modells eine Separation

der jeweiligen Unterkomponenten möglich ist. Dies kann als Indiz für die Existenz getrennter globaler und lokaler Diagrammschemata zu Rahmen und Werten eines Diagramms (vgl. Schnotz, 1994) gewertet werden. Diese scheinen sich jedoch, wie oben ausgeführt, nicht übergreifend auf Informationsentnahme und Konstruktion zu beziehen. Bei der Identifizierung wird demnach ein anderes Diagrammschema als bei der Konstruktion eines Diagrammrahmens und entsprechend beim Ablesen ein anderes Diagrammschema als beim Eintragen von Werten aktiviert.

Das Modell *3dim_q*, das starke Zusammenhänge zwischen Identifizierung und Aufbau sowie zwischen Ablesen und Dateneintrag postuliert, ist, wie bereits oben erwähnt, nicht zur Modellierung der Diagrammkompetenz geeignet. Damit geht der Befund einher, dass sich der Zusammenhang zwischen Informationsentnahme und Konstruktion hauptsächlich auf einen starken Zusammenhang zwischen Ablesen und Aufbau zurückführen lässt (vgl. 6.2.1.2). Dieser stellte sich von den vier möglichen Zusammenhängen zwischen Unterkomponenten *verschiedener* Hauptkomponenten in Korrelations- und Regressionsanalysen als der stärkste heraus. Eine mögliche Erklärung kann darin gesehen werden, dass auch für das Ablesen auf der Seite der Informationsentnahme eine gewisse aktive Konstruktion erforderlich ist. So wird das Ablesen insbesondere höherer Ordnungen als Inferenzprozess aufgefasst, bei dem ein Diagrammleser eigenständig Bedeutungen konstruiert (Shah & Carpenter, 1998; Freedman & Shah, 2002).

Besonders vor dem Hintergrund der ermittelten Zusammenhänge zur Integration wird dieser Zusammenhang zwischen Informationsentnahme und Konstruktion respektive den Unterkomponenten Ablesen und Aufbau bedeutsam, weshalb er an späterer Stelle erneut aufgegriffen wird.

7.1.2 Relation der Komponente Integration zu Informationsentnahme und Konstruktion

Im vorgeschlagenen Strukturmodell der Diagrammkompetenz wurde die Integration als dritte Hauptkomponente neben Informationsentnahme und Konstruktion angenommen. In den nachfolgenden Betrachtungen werden die Ergebnisse hinsichtlich der Stellung der Komponente Integration in Bezug zu Informationsentnahme und Konstruktion zusammenfassend diskutiert.

In den beiden Experimentalmodellen wurde die Integration, wie im Strukturmodell postuliert, als dritte Komponente neben Informationsentnahme und Konstruktion unterschieden. Im Vergleich dazu wurde sie in keinem der drei Kontrollmodelle *2dim*, *2dim_within* und *2dim_asym* als eigenständige Komponente formuliert (vgl. 5.3.3.1). Durch die bessere Passung der beiden Experimentalmodelle im Vergleich zu den genannten Kontrollmodellen wird die Integration als bedeutsame Komponente der

Diagrammkompetenz ausgewiesen. Sie kann damit grundsätzlich als differenzierbare Fähigkeit interpretiert werden, die keiner der beiden anderen Komponenten zuzuordnen ist. Zu berücksichtigen ist dabei allerdings ein starker Bezug der Integration zur Informationsentnahme, der sich in den Korrelationsanalysen gezeigt hat (vgl. 6.2.1.1.) Dieser wird dadurch untermauert, dass das Modell *2dim_asym* nur eine etwas geringere Modellpassung aufweist als das Experimentalmodell *3dim* (vgl. 6.1.1.1): Ein Kompetenzstrukturmodell, das Informationsentnahme und Integration zusammengefasst der Konstruktion gegenüberstellt (*2dim_asym*), passt somit nur wenig schlechter als das Strukturmodell, das Integration sowohl von Informationsentnahme als auch Konstruktion abgrenzt (*3dim*).

Der starke Zusammenhang zwischen Integration und Informationsentnahme basiert, wie Korrelations- und Mediationsanalysen zeigen, vor allem auf einem Zusammenhang zwischen Integration und Ablesen. Letzterer Befund lässt sich darauf zurückführen, dass die für die Studien entwickelten Testaufgaben unter der Maßgabe einer hohen „inhaltlichen Relevanz“ (Reulecke & Rollett, 1976, S. 183) konstruiert wurden und damit nahe an den wirklichen, in der Schule real auftretenden Aufgabenstellungen orientiert waren. Die Informationen, die dabei aus einem Diagramm entnommen werden müssen, sind meist die in den Datenpunkten gegebenen und weniger die Informationen darüber, welcher Zusammenhang grundsätzlich in einem vorliegenden Diagramm dargestellt ist. Daher ergaben sich in den Aufgaben zur Integration häufiger konkrete Ableseanforderungen als Anforderungen an die Identifizierung. Der stärkere Zusammenhang zwischen Integration und Ablesen im Vergleich zum Zusammenhang zwischen Integration und Identifizierung ist somit nur einschränkend generalisierbar, und zwar vorerst ausschließlich für die genannten im Biologieunterricht auftretenden Integrationsanforderungen.

Der genannte starke Bezug der Integration zur Informationsentnahme übertrifft den wesentlich geringeren Zusammenhang zwischen Integration und Konstruktion (vgl. 6.2.1.1). Insbesondere ist anzumerken, dass sich die Informationsentnahme in den Mediationsanalysen als totaler Mediator des Zusammenhangs zwischen Integration und Konstruktion herausgestellt hat. Damit kann letztgenannter Zusammenhang zu einem hohen Anteil auf die Informationsentnahme zurückgeführt werden (vgl. 6.2.1.3).

An dieser Stelle wird der Zusammenhang zwischen Informationsentnahme und Konstruktion erneut aufgegriffen. Es wurde schon berichtet, dass das Ablesen im Vergleich zur Identifizierung der Faktor mit den stärkeren Bezügen zur Konstruktion ist. Die eben beschriebene Mediationswirkung der Informationsentnahme für den Zusammenhang zwischen Integration und Konstruktion geht also vor allem auf das Ablesen zurück. Entsprechend belegen weitere Mediationsanalysen, dass die Zusammenhänge der Integration zu beiden Konstruktionskomponenten zumindest partiell durch das Ablesen vermittelt werden (vgl. 6.2.1.3).

Zusammenfassend ist festzuhalten, dass das Ablesen als Unterkomponente der Informationsentnahme alle Zusammenhänge zwischen den anderen Unterkomponenten (Identifizierung, Aufbau und Dateneintrag) und der Integration vermittelt. Dies unterstreicht den starken Bezug zwischen Ablesen und Integration, der zur differenzierten Beschreibung der Modellstruktur der Diagrammkompetenz hervorzuheben ist. Insbesondere weist dieses Ergebnis darauf hin, dass für die Beantwortung der Items, die als resultierende externe Repräsentation ein Diagramm und damit eine Konstruktionsleistung forderten (*Resultat Diagramm*, vgl. 3.6.3), die Fähigkeiten zur Informationsentnahme eine größere Rolle spielen als die Fähigkeiten zur Konstruktion selbst. Damit einhergehend konnte eine Teilung der Komponente Integration in zwei Unterkomponenten nicht nachgewiesen werden, was ausführlicher im Abschnitt zu den Binnenstrukturen der Komponenten (vgl. 7.2.3) diskutiert wird.

7.2 Binnenstrukturen

Die Analysen der Binnenstrukturen (vgl. 6.2.3) der drei Komponenten Informationsentnahme, Konstruktion und Integration bestätigten die Struktur des postulierten Gesamtmodells mit drei respektive fünf (Unter-)Komponenten. Die jeweiligen Unterkomponenten von Informationsentnahme und Konstruktion wurden auch in der Modellierung der einzelnen Komponenten nachgewiesen. Darüber hinaus erwies sich für keine der (Unter-)Komponenten eine feinere Modellierung in die einzelnen Aspekte, bei der Konstruktion beispielsweise Auswahl des Diagrammtyps oder Achsenbelegung, als besser passend. Damit können die Einzelaspekte der (Unter-)Komponenten nicht als qualitativ unterschiedliche Strukturen aufgefasst werden (vgl. Einhaus, 2007). Für weitergehende Analysen, die Hinweise auf die Binnenstrukturen der (Unter-)Komponenten geben, wurden die ermittelten Itemschwierigkeiten herangezogen.

7.2.1 Ablesen

Für das Ablesen wurden im Strukturmodell der Diagrammkompetenz vier Ordnungen beschrieben, die Ableseanforderungen verschiedener Stufen repräsentieren. Die Mittelwerte der empirisch ermittelten Itemschwierigkeiten der jeweiligen Itemgruppen waren allerdings nicht entsprechend den vier Ablesordnungen angeordnet. Beispielsweise zeigte sich, dass die Items zum Ablesen zweiter und dritter Ordnung im Mittel leichter von den Probanden zu lösen waren als die zum Ablesen erster und vierter Ordnung. Die postulierte Anordnung der Ableseordnungen kann somit in der vorliegenden Arbeit nicht bestätigt werden. Allerdings kann aus den Ergebnissen aus folgenden Gründen ebenfalls nicht abgeleitet werden, dass sich eine gänzlich andere Anordnung der Ableseordnungen zeigt: erstens sind die Unterschiede zwischen den Mittelwerten nicht statistisch bedeutsam,

zweitens treten innerhalb der einzelnen Ordnungen große Streuungen der Itemschwierigkeiten auf. Es sind also nicht alle Items zum Ablesen zweiter Ordnung leichter als alle Items zum Ablesen erster Ordnung.

Für die von der Binnenstruktur des Ablesens abweichenden Itemschwierigkeiten können zwei mögliche Erklärungen herangezogen werden: Zum ersten könnte in Betracht gezogen werden, dass die zur Formulierung der Ableseordnungen herangezogene Literatur die verschiedenen Ableseordnungen nicht präzise genug beschreibt, so dass die Items anderen als in der Originalliteratur intendierten Ableseordnungen zugeordnet wurden. Beispielsweise wurde, entsprechend der Beschreibung des Ablesens dritter Ordnung als Vergleich mehrerer Trends (Schnotz, 1994), das Erkennen eines Minimums im Linienverlauf, wozu mindestens zwei Trends miteinander verglichen werden müssen, dem Ablesen dritter Ordnung zugeordnet. Es wäre durchaus denkbar, dass es nach Schnotz als Ablesen eines Punktwertes aufgefasst werden sollte und demnach dem Ablesen erster Ordnung hätte zugeordnet werden müssen. Dafür spräche die sehr geringe Itemschwierigkeit eines entsprechenden Items in der vorliegenden Arbeit. Demzufolge müssten die Beschreibungen der Ablesestufen präzisiert und ergänzt werden. Eine Möglichkeit, Hinweise auf eine solche präzisierte Beschreibung zu erhalten, wäre die Anordnung aller Items entsprechend ihrer Schwierigkeiten auf einem Kontinuum, um davon ausgehend Abschnitte festzulegen, die verschiedene Stufen der Anforderungen und entsprechend benötigte Kompetenzaspekte charakterisieren. Dieses Verfahren wird aktuell in der Entwicklung von deskriptiven Kompetenzstufenmodellen angewandt (Hartig, 2007).

Zum zweiten könnte die beschriebene Stufung des Ablesens zwar mit entsprechenden Kompetenzstufen des Ablesens einhergehen, die Itemschwierigkeit jedoch durch andere Aufgabenmerkmale beeinflusst worden sein (vgl. Prenzel et al., 2002). So wird von mehreren Autoren (Shah et al., 1999; Katz et al., 2002; Friel et al., 2001; Rinck & Glowalla, 1994) die Gruppierung bestimmter graphischer Elemente in Diagrammen, beispielsweise von Säulen mehrerer Datenreihen, und generell die Komplexität eines Diagramms, beispielsweise die Anzahl der Datenreihen oder abhängigen Variablen (Shah & Carpenter, 1995) als möglicher Einflussfaktor auf die Schwierigkeit von Ableseprozessen angeführt. Dies wird im Modell zum Diagrammverstehen von Carpenter und Shah (1998) dadurch erklärt, dass eine Information umso leichter aus einem Diagramm abgelesen werden kann, je klarer sie der einfachen Wahrnehmung aus den graphischen Strukturen zugänglich ist. Entsprechend könnte ein und derselben Person das Erkennen eines Minimums in einem Diagramm mit einer Datenreihe sehr leicht, in einem anderen Diagramm mit mehreren, sich überschneidenden Linien oder nebeneinanderstehenden Säulen sehr schwer fallen. In diesem Fall wäre die graphische Komplexität des Diagramms bei der Formulierung der einzelnen Ableseordnungen einzubeziehen. Dem vergleichbar zeigt sich in der Beschreibung der

Kompetenzstufen zur Lesekompetenz der PISA 2003-Erhebung (Schaffner, Schiefele, Drechsel & Artelt, 2004) eine Zunahme der Textkomplexität auf höheren Stufen.

7.2.2 Konstruktion

Die einzelnen untersuchten Aspekte der Konstruktion erwiesen sich als unterschiedlich schwer, wenngleich sie sich insgesamt nicht statistisch bedeutsam unterscheiden. Im Folgenden werden die Schwierigkeiten dieser Einzelaspekte diskutiert.

Besonders hinsichtlich der Schwierigkeiten von Achsenbelegung und Werteintrag weichen die Ergebnisse der vorliegenden Arbeit von Befunden von Padilla und Kollegen (1986) ab, nach denen die Achsenbelegung eher einen schwierigen, der Werteintrag eher einen leichten Konstruktionsaspekt darstellt. Für diese entgegengesetzten Befunde sind folgende Erklärungsansätze möglich: In der Studie von Padilla und Kollegen (1986) wurden Single-Choice-Aufgaben eingesetzt, die vier Antwortmöglichkeiten enthielten. Bei der Achsenbelegung umfassten diese Antwortmöglichkeiten erstens die richtige Belegung, zweitens eine in Bezug auf Ordinate und Abszisse vertauschte Belegung. Darüber hinaus waren zwei Antwortmöglichkeiten gegeben, die Belegungen mit ähnlichen, aber zur Darstellung eines Sachverhalts nicht relevanten Variablen zeigten (vgl. Abbildung 3.5). In der vorliegenden Arbeit kann das Fehlen konkurrierender Variablen in der Aufgabenstellung, wie sie beispielsweise von Baker und Kollegen (2002) eingesetzt wurden, somit zu der geringeren Schwierigkeit der Achsenbelegung geführt haben. Um dem Konstruktionsaspekt der Achsenbelegung stärkere Anteile in Bezug auf die Auswahl der Variablen, die in einem Diagramm aufgegriffen werden sollen, zu geben, wäre eine Aufnahme solcher konkurrierender Variablen in die Aufgabenstellung sinnvoll.

Auch hinsichtlich des Werteintrags kann vermutet werden, dass die Unterschiede in den Aufgabenstellungen zwischen der Studie von Padilla und Kollegen (1986) und der vorliegenden Arbeit zu den unterschiedlichen Befunden hinsichtlich der Itemschwierigkeiten beigetragen haben. So umfassten bei Padilla und Kollegen die Alternativen bei der Auswahl eines Punktes ausschließlich ganzzahlige Koordinaten in einem groben Koordinatensystem. In der vorliegenden Arbeit wurde das Eintragen der Werte strenger kodiert, so dass bereits kleine Abweichungen als „ungenaueres Eintragen“ angesehen wurden. Die qualitative Analyse der Fehler beim Eintragen der Werte (vgl. 6.4.5) ergab, dass bis zu einem guten Drittel der Schüler mit Fehlern beim Werteintrag nur aufgrund des fehlerhaften Eintragens *eines* von etwa drei respektive vier Datenpunkten keine volle Punktzahl erhielten.

Die Schüler scheinen somit grundsätzlich über die benötigten Fähigkeiten zum Eintragen der Werte zu verfügen. Ähnliche Hinweise liefert die qualitative Analyse der Fehler zur

Skalierung. In etwa einem Drittel der Antworten wies die tendenziell metrische Skalierung kleinere Ungenauigkeiten wie eine unregelmäßige numerische Beschriftung der Skala auf.

Die Interpretation liegt nahe, dass die aufgetretenen Fehler in vielen Fällen auf ein oberflächliches oder flüchtiges Arbeiten zurückgehen. Die Schüler scheinen grundsätzlich zu wissen, wie man Punkte oder Säulen einträgt. Ebenso scheinen sie grundsätzlich in der Lage zu sein, eine metrische Skala zu erstellen. Diese Ergebnisse werfen die Frage danach auf, ob die Schüler mit anderer Intention der Diagrammkonstruktion, wie beispielsweise der Überprüfung einer Hypothese, zu besseren Konstruktionsleistungen fähig wären (Shah & Hoeffner, 2002). Diese Überlegungen heben indes die Notwendigkeit zur Erfassung der volitionalen Faktoren bei der Kompetenzmessung (KMK, 2005a) hervor.

Einen Hinweis auf Schwierigkeiten bei der Diagrammkonstruktion, die sich nicht auf oberflächliches Arbeiten zurückführen lassen, geben die Analysen zur Auswahl des Diagrammtyps. Dieser Aspekt der Konstruktion hat sich mit einer Itemschwierigkeit im mittleren Bereich als gut zu bewältigen erwiesen. Dennoch ist anzumerken, dass in etwa zehn Prozent der Fälle je entsprechender Aufgabe anstelle eines Säulen- ein Liniendiagramm gezeichnet wurde. Einem Teil der Schüler scheint somit der Interpolationscharakter einer Linie nicht bewusst zu sein (Kerslake, 1981). Eine mögliche Erklärung dafür kann darin gesehen werden, dass Liniendiagramme sowohl in der Schule als auch in der Wissenschaft bei der Darstellung kategorialer Daten zur besseren Lesbarkeit verwendet werden (vgl. bspw. typische Darstellungen der Profile bei einer latenten Klassenanalyse; Rost, 2004, S. 158).

7.2.3 Integration

In Bezug auf die Integration wurden die zwei Aspekte *Resultat Text* und *Resultat Diagramm* basierend auf den verschiedenen, aus einem Integrationsprozess resultierenden externen Repräsentationsformen unterschieden (vgl. 3.6.3). Eine entsprechende Unterscheidung dieser zwei Integrationsaspekte wurde in den Modellen *2dim* und *2dim_within* für die gesamte Diagrammkompetenz vorgenommen. Dabei wurden die entsprechenden Testitems entweder je einer der zwei Komponenten Informationsentnahme und Konstruktion zugeordnet (*2dim*) oder beiden dieser Komponenten (*2dim_within*). Beide Modelle wurden jedoch aufgrund zu geringer Modellpassungen abgelehnt, wie in den Analysen zu den Komponenten der Modellstruktur aufgeführt wird (vgl. 6.1.1). Diese Ergebnisse weisen bereits darauf hin, dass sich die theoretisch formulierten Aspekte der Integration, *Resultat Text* und *Resultat Diagramm*, nicht empirisch nachweisen lassen. Dies wurde, wie bereits in der Diskussion der Komponenten und ihren Relationen angesprochen (vgl. 7.1.2), durch die Analysen zur Binnenstruktur der Integration (vgl. 6.2.3.3) bestätigt.

Letztere Analysen lassen zunächst den Schluss zu, dass sich die beiden Aspekte zwar nicht als qualitativ unterschiedliche Strukturen ansehen lassen, sich jedoch hinsichtlich der Itemschwierigkeiten voneinander unterscheiden. Allerdings sind erstens die beobachteten Unterschiede statistisch nicht signifikant. Zweitens liefern die Analysen der Itemschwierigkeiten Hinweise darauf, dass die Unterschiede in den Schwierigkeiten nicht auf den Integrationsaspekt zurückzuführen sind, sondern auf das Itemformat. So ist insbesondere ein offenes Item mit dem *Resultat Text* deutlich schwieriger als die übrigen Items mit dem *Resultat Text*, die im Single-Choice-Format gestaltet sind. Aufgrund einer möglichen Konfundierung von Integrationsaspekt und Itemformat ist somit – obwohl über alle Items hinweg das Aufgabenformat als schwierigkeitsbestimmendes Merkmal ausgeschlossen werden konnte – hinsichtlich der Binnenstruktur der Integration keine eindeutige Aussage zu treffen.

7.3 Betrachtung der verschiedenen Diagrammtypen

Die im Kompetenzstrukturmodell formulierten Komponenten und Unterkomponenten der Diagrammkompetenz beziehen sich mit Ausnahme der Skizzierung auf die beiden Diagrammtypen Linien- und Säulendiagramm. Im Folgenden werden die Analysen, die die beiden Diagrammtypen einzeln in den Blick nehmen, diskutiert.

Es wurden Modellberechnungen durchgeführt, die Aufschluss darüber geben sollten, ob zur Modellierung der Diagrammkompetenz eine Differenzierung nach Diagrammtypen nötig und sinnvoll ist. Dazu wurden jeweils für die Kompetenzstruktur aus drei sowie fünf Komponenten entsprechende Modelle berechnet, die sämtliche Testitems nach dem Diagrammtyp aufteilen (within-Modelle, vgl. 5.1.2.4). Diese Modelle wiesen gegenüber den Modellen *3dim* und *5dim*, die keine Differenzierung der Diagrammtypen vornehmen, zwar hinsichtlich der verglichenen Indices für die Modellpassungen bessere Werte auf, wurden aber aufgrund sehr geringer Reliabilitäten der einzelnen Komponenten verworfen (vgl. 6.1.2.2). Damit wurde die Hypothese 2 bestätigt, wonach ein Linien- und Säulendiagramme umfassendes Modell zur Modellierung der Diagrammkompetenz geeignet ist.

An dieser Stelle muss methodenkritisch die geringe Itemzahl in den einzelnen sich bei Differenzierung nach Diagrammtypen ergebenden Komponenten berücksichtigt werden. Die empfohlene Itemzahl von fünf pro Komponente ist in einigen Fällen unterschritten, was zudem ein Grund für die geringen Reliabilitäten der Einzelskalen sein kann. Um aussagekräftigere Ergebnisse zu gewinnen, wären Testungen mit einer höheren Itemzahl sinnvoll.

Mit den gegebenen Einschränkungen sprechen die bisherigen Ergebnisse dafür, für beide Diagrammtypen ein gemeinsames Modell für die Diagrammkompetenz zu etablieren.

Diese Ergebnisse werden von aktuellen Befunden von Ratwani und Trafton (in press) gestützt. In ihren Untersuchungen fanden sie Belege für die Existenz eines übergreifenden, sowohl Linien- als auch Säulendiagrammen unterliegenden Diagrammschemas (vgl. Schnotz, 1994; s. 3.5.1). Danach sind die genannten Diagrammtypen nicht nur vom Aufbau her ähnlich, sondern werden auf ähnliche Weise verarbeitet, was für ein übergreifendes Modell der Diagrammkompetenz spricht. Die postulierten Strukturen der Diagrammkompetenz gelten damit in Bezug auf Linien- als auch auf Säulendiagramme.

Über die Struktur hinaus wurde untersucht, ob sich für Linien- und Säulendiagramme unterschiedliche Itemschwierigkeiten ergeben, die zur differenzierteren Beschreibung des Modells der Diagrammkompetenz herangezogen werden könnten (vgl. 6.2.4). Über die Komponenten hinweg zeigte sich kein signifikanter Unterschied in den Schwierigkeiten in Abhängigkeit vom Diagrammtyp. Auch in Bezug auf die Itemschwierigkeiten innerhalb der Komponenten Informationsentnahme und Konstruktion traten nur minimale Unterschiede zwischen den Diagrammtypen auf. Hinsichtlich der Informationsentnahme decken sich diese Ergebnisse mit Befunden zur Schnelligkeit von Ableseprozessen aus Diagrammen (Jacobs, 1994, 1995b), die zwischen Linien- und Säulendiagrammen keine bedeutsamen Unterschiede offen legten (vgl. 3.5.3).

Im Gegensatz dazu zeigte sich eine vergleichsweise große Schwierigkeitsdifferenz zwischen den Linien- und Säulendiagramm-Items in Bezug auf die Integration (vgl. 6.2.4.). Diese muss allerdings in Hinblick auf die Heterogenität der Items in beiden Gruppen relativiert werden. Anders als innerhalb der Komponenten Informationsentnahme und Konstruktion sind die Aufgabenstellungen der Integrationsitems zu Linien- und Säulendiagrammen direkt nicht vergleichbar, so dass jeweils andere Kombinationen von Informationsentnahme- und gegebenenfalls Konstruktionsprozessen erforderlich waren. Um zu untersuchen, ob sich die Schwierigkeitsunterschiede auf die Diagrammtypen zurückführen lassen, müssten die Schülerantworten jeweils im Hinblick auf eine festgelegte Aufgabenstellung der Integration verglichen werden (vgl. Ratwani & Trafton, in press). Es kann somit aus den vorliegenden Ergebnissen trotz einer Differenz der mittleren Itemschwierigkeiten für die Diagrammtypen nicht gefolgert werden, dass die Integration anhand von Liniendiagrammen schwieriger sei als anhand von Säulendiagrammen. Vielmehr liefert die Gegenüberstellung zweier hinsichtlich ihrer Aufgabenstellung vergleichbarer Items, bei denen als Integrationsleistung das Eintragen einzelner Werte in ein Diagramm gefordert war, ähnliche Schwierigkeitswerte für Linien- und Säulendiagramm (vgl. 6.2.4).

Insgesamt scheint der Diagrammtyp demnach keinen zu berücksichtigenden Einflussfaktor bei der Diagnose der Diagrammkompetenz darzustellen. Es ist jedoch anzumerken, dass in den Diagrammen, die in der vorliegend beschriebenen Arbeit verwendet wurden, jeweils lediglich ein oder zwei Datenreihen dargestellt werden respektive von den Schülern

darzustellen waren. Sie sind damit als wenig komplex einzustufen (vgl. Carpenter & Shah, 1998). Im Hinblick auf komplexere Diagramme, in denen eine höhere Zahl an Datenreihen dargestellt wird, kann der Diagrammtyp wiederum als Einflussfaktor relevant werden. Hinweise darauf geben die Studien von Jacobs (1994), nach denen Probanden beim Vorliegen von acht Datenreihen Trends in Liniendiagrammen deutlich schneller und genauer erkannten als in Säulendiagrammen. Die genannten Untersuchungen von Jacobs beziehen sich allerdings nur auf die Informationsentnahme. Es ist offen, ob sich in komplexeren Diagrammen auch im Hinblick auf Konstruktion und Integration Unterschiede in Abhängigkeit vom Diagrammtyp zeigen und damit, ob das beschriebene Kompetenzstrukturmodell an dieser Stelle erweitert werden sollte.

7.4 Lernerhintergrund und Lernermerkmale

7.4.1 Lernerhintergrund

Der Lernerhintergrund wird in zahlreichen Schulleistungsstudien als Erklärungsansatz für divergierende Testergebnisse herangezogen. Er wurde in der vorliegenden Arbeit dahingehend beleuchtet, ob sich Neunt- und Zehntklässler sowie Schülerinnen und Schüler hinsichtlich ihrer Fähigkeiten in den einzelnen (Unter-)Komponenten der Diagrammkompetenz unterscheiden.

Im Hinblick auf die Klassenstufe zeigten sich dabei signifikante Unterschiede in Identifizierung, Ablesen und Integration. Innerhalb dieser (Unter-)Komponenten waren die Zehntklässler den Neuntklässlern überlegen (vgl. 6.3.1.1). Für die restlichen Unterkomponenten Aufbau und Dateneintrag hingegen wurde eine solche deutliche Überlegenheit der Zehntklässler nicht festgestellt. Zusammengenommen stellen diese Befunde eine Differenzierung der Ergebnisse von Padilla und Kollegen (1986) dar, die eine Überlegenheit der Zehntklässler für einen Gesamtwert, der sowohl Informationsentnahme als auch Konstruktion umfasst, ermittelten. Nach den vorliegenden Ergebnissen scheint sich diese Überlegenheit dabei vorrangig auf die Informationsentnahme und die Integration zu beziehen, während sich die Schüler von der neunten zur zehnten Klassenstufe im Hinblick auf die Diagrammkonstruktion kaum verbessern konnten. Letzteres entspricht den Ergebnissen zum Kompetenzzuwachs zur Naturwissenschaftlichen Kompetenz (die den Umgang mit Graphiken umfasst), der in der PISA 2003-Studie von der neunten zur zehnten Klasse ermittelt wurde. Hier wurde berichtet, dass sich lediglich 44 % der Schüler bedeutsam gesteigert haben, während bei 19 % der Schüler in der zehnten Klasse sogar eine geringere Kompetenz als in der neunten Klasse festgestellt wurde (PISA-Konsortium Deutschland, 2006b).

In Bezug auf das Geschlecht zeigten sich in der vorliegenden Arbeit für keine der (Unter-)Komponenten Unterschiede (vgl. 6.3.1.1). Dieser Befund entspricht im Hinblick auf die Informationsentnahme den Ergebnissen von Curcio (1986). In ihren Studien hatte das Geschlecht der untersuchten Siebtklässler keinen Einfluss auf deren Ableseleistungen. Abweichend davon wurden in den PISA-Studien 2000 und 2003 im Hinblick auf das Lesen nicht-kontinuierlicher Texte (zu denen sowohl Diagramme als auch Karten, Tabellen und andere Formen von Graphiken zählen) deutliche geschlechtsspezifische Unterschiede ermittelt. Nach diesen zeigte sich in beiden Studien eine Überlegenheit der Mädchen (Stanat und Kunter, 2002). Als Erklärung für diese unterschiedlichen Befunde kann das Zusammenfassen der verschiedenen Graphiktypen in den PISA-Studien herangezogen werden. Die besseren Leseleistungen der Mädchen gehen möglicherweise nicht auf die Diagramme, sondern auf die anderen Graphiktypen zurück.

Über die Fähigkeiten der Schüler hinaus wurde in der vorliegenden Arbeit untersucht, ob sich die durch das Kompetenzmodell beschriebenen Komponenten in ihrer Strukturierung und ihren Relationen zueinander in Abhängigkeit von der Klassenstufe und dem Geschlecht der Schüler unterscheiden. Es konnte gezeigt werden, dass die Items innerhalb der beschriebenen Komponenten sowohl für die Neunt- und Zehntklässler als auch für die Jungen und Mädchen personenhomogen sind (vgl. 6.1.2.1, Komponenten: Personenhomogenität). Dies bedeutet, dass die durch die jeweiligen Items getesteten Fähigkeiten sich zwischen den Gruppen nicht unterscheiden (Rost, 1996). Daraus folgt, dass die über alle Personen hinweg nachgewiesenen Strukturen der Diagrammkompetenz auch für die vier Untergruppen bestätigt werden konnten und dass das beschriebene Kompetenzstrukturmodell über die Gruppen hinweg Gültigkeit hat. Darüber hinaus zeigten sich auch in den zentralen Relationen der einzelnen Komponenten zueinander (vgl. 7.1) keine Unterschiede für einzelne Gruppen (vgl. 6.3.1.2). Daraus folgt, dass sowohl in der neunten und zehnten Klassenstufe als auch für Schüler und Schülerinnen ähnliche Zusammenhänge zwischen den Schülerfähigkeiten in den einzelnen Komponenten bestehen. Insgesamt konnten damit die über alle Personen hinweg empirisch nachgewiesenen Strukturen der Diagrammkompetenz und die ermittelten Relationen zwischen diesen für die einzelnen Untergruppen ebenfalls validiert werden.

7.4.2 Lernermerkmale

Die untersuchten Lernermerkmale wurden in unterschiedlicher Stärke als Vorhersagewerte für unterschiedliche (Unter-)Komponenten identifiziert. Bevor auf die Deutung dieser Befunde eingegangen wird, ist auf die geringe Itemzahl der einzelnen Variablen kritisch hinzuweisen. Beispielsweise wurde die Lesekompetenz mit nur einem Item erhoben, die

Skala zum biologischen Vorwissen umfasste drei Items. Diese geringe Itemzahl bringt Einschränkungen im Hinblick auf die Generalisierbarkeit der Ergebnisse mit sich. Damit werden die nachfolgend diskutierten Befunde lediglich als Indizien für die Wirkung der Lernermerkmale verstanden.

7.4.2.1 Biologisches Vorwissen

Das Vorwissen zu biologischen Themen stellte sich in der vorliegenden Arbeit lediglich in Bezug auf die Identifizierung als Einflussfaktor heraus. Schüler, die beispielsweise über Kenntnisse zur Populationsgröße verfügten, erbrachten bessere Leistungen in Hinblick auf die Identifizierung. Auf die weiteren hier untersuchten (Unter-)Komponenten der Diagrammkompetenz scheint das biologische Vorwissen keinen Einfluss zu haben. Ausgehend von den in Abschnitt 3.7.2.1 dargestellten Befunden über einen starken Einfluss des inhaltlichen Vorwissens auf die Informationsentnahme (z. B. Preece & Janvier, 1992) hätte vermutet werden können, dass sich beispielsweise auch zwischen den Fähigkeiten zum Konstruieren und dem Vorwissen ein Zusammenhang zeigt. Dieser wurde in der vorliegenden Arbeit jedoch nicht festgestellt. Eine mögliche Ursache dafür wäre die geringe Zahl an Items zur Messung des biologischen Vorwissens sowie die geringe Reliabilität der Skala in der vorliegenden Arbeit. Möglich wäre jedoch ebenso, dass für die vorliegende Arbeit Items konzipiert wurden, deren Bearbeitung gerade keinem respektive sehr geringem Einfluss des Vorwissens unterliegt. Sie wären damit geeignet, die *grundlegende* Diagrammkompetenz in ihren einzelnen Komponenten zu erfassen, ohne dass zu ihrer Bearbeitung ein spezifisches Vorwissen vorausgesetzt werden muss.

7.4.2.2 Umgang mit abstrakten Diagrammen

Die Fähigkeiten der Schüler zum Umgang mit abstrakten Diagrammen haben sich im Vergleich zu den anderen Lernermerkmalen als Prädiktor für fast alle der (Unter-)Komponenten der Diagrammkompetenz herausgestellt (vgl. 6.3.2). Zur Interpretation werden die Ergebnisse in Bezug auf die einzelnen Items (s. Tabelle 6.28) herangezogen.

Bei der Analyse der Einzelitems wurde deutlich, dass der ermittelte Einfluss der Fähigkeiten zum Umgang mit abstrakten Diagrammen größtenteils auf ein Item zurückgeht, welches ein Ablesen dritter Ordnung in einem abstrakten Säulendiagramm abfragt (Item *S_uad3*, s. 12.3.2.4). Dabei mussten die Höhendifferenzen jeweils zweier Säulen miteinander verglichen werden. Es zeigte sich, dass Schüler, die dieses Item zum Ablesen in einem abstrakten Szenario richtig beantworten konnten, über höhere Fähigkeitswerte zur Informationsentnahme verfügten als Schüler, die es falsch

beantworteten. Dieses Ergebnis steht im Widerspruch zu Befunden Curcios (1987), nach denen dieses Lernermerkmal für die Schüler der siebten Klasse keinen Einfluss auf die Ableseleistungen hatte. Möglicherweise ist dies auf Unterschiede in den Items zurückzuführen, die zur Messung der Fähigkeiten zum Umgang mit abstrakten Diagrammen verwendet wurden. Die Items, die zu ihrer Studie bekannt sind (Curcio, 1987), scheinen deutlich leichter als das Item, das sich in der vorliegenden Arbeit als Prädiktor herausgestellt hat.

Das Item, das die abstrakte Definition einer abhängigen Variablen erfragt, erwies sich als bedeutsamer Vorhersagewert für die Identifizierung. Dies scheint plausibel, da die Identifizierung das Erkennen der dargestellten Relation einschließt, zu dem wiederum das Erkennen der unabhängigen und abhängigen Variable gehört. Diese Identifizierungsprozesse scheinen Schülern in den konkreten Situationen leichter zu fallen, wenn sie die abstrakte Definition einer abhängigen Variablen kennen.

Neben dem Einfluss auf die Identifizierung und das Ablesen wurde das Item zum abstrakten Ablesen (*S_uad3*, s.o.) auch als signifikanter Prädiktor für den Aufbau und die Integration ermittelt. Damit scheinen die Ablesefähigkeiten der Schüler im Abstrakten ebenfalls Bezüge zur Konstruktion und Integration zu besitzen. Zusammengefasst deuten diese Befunde darauf hin, dass die Fähigkeiten zum Umgang mit *abstrakten* Diagrammen Bezüge zu einzelnen (Unter-)Komponenten der Diagrammkompetenz und damit zur Anwendung von Diagrammen in *konkreten* Situationen aufweisen.

7.4.2.3 Lesekompetenz

Die Lesekompetenz stellte sich in den Regressionsanalysen als signifikanter Vorhersagewert für die Identifizierung, den Aufbau und den Dateneintrag heraus. Schüler, die aus einem Text zielgerichtet eine Information heraussuchen können, besaßen damit höhere Fähigkeiten in den genannten Komponenten. Im Hinblick auf die Identifizierung ist dieses Ergebnis zu den Befunden von Curcio (1987) in Bezug zu setzen. In Curcios Studie zeigte sich die Lesekompetenz ebenfalls als Einflussfaktor auf das Lesen von Diagrammen. Darunter fasst Curcio allerdings sowohl die Identifizierung als auch Ableseaspekte zusammen (vgl. 3.5.2.1). Aus den Ergebnissen der vorliegenden Arbeit kann somit geschlossen werden, dass der Einfluss der Lesekompetenz auf das Diagrammlesen bei Curcio (1987) hauptsächlich auf dessen Einfluss auf die Identifizierung zurückgeht.

Der Einfluss der Lesekompetenz auf die beiden Unterkomponenten der Konstruktion (Aufbau und Dateneintrag) kann nach dem Modell der Text- und Diagrammverarbeitung von Schnotz (2001) darauf zurückgeführt werden, dass das Auffinden bestimmter Textinformationen mit dem Ziel, ein mentales Modell aufzubauen, direkter Bestandteil des Übersetzungsprozesses von der textlichen Repräsentation zur bildhaften Repräsentation

(Diagramm) ist. Unklar ist allerdings, weshalb die Lesekompetenz in der vorliegenden Untersuchung keinen Einfluss auf die Integration hatte.

7.4.2.4 Rechenfähigkeiten

Die erhobenen Rechenfähigkeiten stellten sich in den Regressionen als signifikanter Vorhersagewert für das Ablesen heraus (vgl. 6.3.2). Dieses Ergebnis ist konform mit den in der Literatur zu findenden Befunden (Pereira-Mendoza und Mellor, 1991; Åberg-Bengtsson & Ottosson, 2006). Ferner zeigte sich die Mathematiknote der Schüler als Vorhersagewert im Hinblick auf die beiden Unterkomponenten der Konstruktion (vgl. 6.3.2). Schüler mit guten Noten in Mathematik scheinen höhere Fähigkeiten in Bezug auf die Konstruktion von Diagrammen zu biologischen Inhalten zu besitzen als Schüler mit schlechten Noten.

7.4.2.5 Anstrengung

Die von den Schülern selbst eingeschätzte Anstrengung zeigte sich als signifikanter Vorhersagewert für die beiden Unterkomponenten der Konstruktion (Aufbau und Dateneintrag). Schülern scheint damit die Konstruktion von Diagrammen eine höhere Überwindung zu kosten. Dies wird dadurch untermauert, dass abgesehen von den Aufgaben, die sich am Ende des Tests befinden, häufiger Konstruktionsitems als Items zur Informationsentnahme oder Integration ausgelassen wurden (vgl. 6.3.2).

Dieses Ergebnis ist von Bedeutung im Hinblick auf das Verständnis von Kompetenz nach Weinert (2001a), in dem motivationale Faktoren einbezogen werden. Mit aller Vorsicht kann aufgrund der Befunde zur Anstrengung respektive zum Auslassen einzelner Items geschlossen werden, dass motivationale Faktoren im Bezug auf die Konstruktion eine größere Rolle spielen als für die anderen Komponenten der Diagrammkompetenz.

Zusammenfassend ist zu konstatieren, dass sich keines der aufgeführten Lernermerkmale als Vorhersagewert für sämtliche (Unter-)Komponenten der Diagrammkompetenz erwiesen hat. Mit den anfangs genannten Einschränkungen weisen diese Befunde darauf hin, dass eine differenzierende Betrachtung der einzelnen Komponenten nötig ist, um die unterschiedlichen Merkmale der Schüler bei der Diagnose und Förderung von Diagrammkompetenz berücksichtigen zu können.

7.5 Testinstrument

Neben den diskutierten Ergebnissen der Analysen zu den Forschungsfragen wird an dieser Stelle das in der vorliegenden Arbeit eingesetzte Testinstrument kritisch beleuchtet.

Dieses Testinstrument wurde, basierend auf den Beobachtungen in der Pilotierung der Aufgaben, so konzipiert, dass erwartet wurde, Neunt- und Zehntklässler könnten es innerhalb einer 45-minütigen Schulstunde vollständig bearbeiten. Die letzten vier Items des Tests mussten jedoch aus den Analysen ausgeschlossen werden, da sie von weniger als der Hälfte der Schüler bearbeitet worden waren (vgl. 5.3.2.2). Insgesamt ging die Bearbeitungshäufigkeit der Items zum Ende des Tests deutlich zurück (vgl. 6.3.2) was darauf hindeutet, dass der Test für diese Klassenstufen etwas zu lang war.

Die sinkende Bearbeitungshäufigkeit zum Ende des Tests wirft darüber hinaus die Frage nach möglichen Positionseffekten der Aufgaben auf: So wäre es möglich, dass die Position eines Items am Ende des Tests zur Schwierigkeit des Items beigetragen hat (*end-of-test-effect*, vgl. Åberg-Bengtsson & Ottosson, 2006). Zur Kontrolle dieses möglichen Positionseffektes wäre es erstrebenswert, eine weitere Erhebung mit einer zweiten Version des Haupttestinstruments durchzuführen, in der die Reihenfolge der Aufgaben verändert ist (vgl. Rotationsdesign, OECD, 2005, S. 16).

Ferner ist das Format der Testitems zu diskutieren. Für die vorliegende Arbeit wurde eine Kombination aus Single-Choice-Items und Items mit offenem Antwortformat gewählt. Damit unterscheidet sich das eingesetzte Testinstrument von anderen, zur Erhebung von Schülerleistungen genutzten Tests wie beispielsweise dem *Test of Graphing in Science* (TOGS, McKenzie & Padilla, 1986) oder dem Diagramm-Subtest des *Swedish Scholastic Aptitude Test* (SweSAT, Subtest Diagrams, Tables & Maps, Stage & Ögren, 2004; vgl. auch Åberg-Bengtsson, 1999). In beiden letzteren werden ausschließlich Single-Choice-Aufgaben verwendet. Berg und Smith (1994) argumentieren, dass mit Aufgaben dieses Formats keine eventuellen Fehlkonzepte (vgl. Mevarech & Kramarsky, 1997) von Schülern aufgedeckt werden können, und plädieren für offene Aufgabenstellungen. Dieser Empfehlung wurde in der vorliegenden Arbeit durch die Verwendung von Items in offenem Format vor allem hinsichtlich der Integration und Konstruktion nachgekommen. Dabei wurden bereits aus der Literatur bekannte auftretende Fehler in den Rubriken des Kodierleitfadens berücksichtigt und diese darüber hinaus in der Pilotierung um Rubriken erweitert, die weitere Fehler charakterisieren. Durch die Erfassung von insgesamt sechs Aspekten der Konstruktion stellen die in der vorliegenden Arbeit berichteten auftretenden Fehler eine Ergänzung zu den aus der Literatur bekannten Befunden, sowohl aus Single-Choice-Erfassungen als auch aus der Analyse offener Antworten, dar. Beispielsweise bezieht sich die differenzierte Analyse der Fehlkonzepte in der Studie von Berg und Smith (1994) lediglich auf das freie Skizzieren einer Linie zur Darstellung einer konkreten Situation, da in der Aufgabenstellung bereits der vollständige Diagrammrahmen

vorgegeben war. Darüber hinausgehend liefert die vorliegende Arbeit differenzierte Befunde darüber, hinsichtlich *welcher* Aspekte innerhalb eines gesamten Konstruktionsprozesses Schwierigkeiten auftreten. In den qualitativen Analysen wurde zudem die Art der auftretenden Fehler (Shah & Hoeffner, 2002) beleuchtet.

Aus ökonomischen Gründen wurden zur Erfassung der Fähigkeiten zur Informationsentnahme überwiegend Single-Choice-Items eingesetzt. Diese sind nicht in dem Maße zur Aufdeckung der Ursachen von Schwierigkeiten geeignet wie offene Items (Berg & Smith, 1994), können jedoch zur Messung der Fähigkeiten von Schülern genutzt werden. Insbesondere können bekannte Fehler durch die Formulierung entsprechender Distraktoren berücksichtigt und erfasst werden. Da die Informationsentnahme im Vergleich zur Konstruktion wesentlich besser erforscht ist (Shah & Hoeffner, 2002), wird der überwiegende Einsatz der Single-Choice-Items in Bezug auf diese Komponente als vertretbar angesehen.

7.6 Implikationen für die Unterrichtspraxis

Die zentralen Befunde der vorliegenden Arbeit werden nachfolgend aufgegriffen, um einige Implikationen für die Diagnose und Förderung der Diagrammkompetenz im Biologieunterricht abzuleiten.

7.6.1 Komponenten der Diagrammkompetenz und deren Relationen

Durch die Modellierungen der vorliegenden Arbeit sind Informationsentnahme und Konstruktion als differenzierte Komponenten der Diagrammkompetenz ausgewiesen worden. Grundsätzlich bedeutet das, dass beide Komponenten nicht gleichgesetzt werden können, sondern bei der Diagnose und Förderung von Diagrammkompetenz einzeln berücksichtigt werden sollten. Die Analysen zu den Relationen der Komponenten respektive ihrer Unterkomponenten geben andererseits Hinweise darauf, dass Bezüge zwischen einzelnen Kompetenzkomponenten bestehen. Diese können dazu genutzt werden, Ansätze für eine systematische Förderung zu entwickeln.

Es zeigte sich ein Zusammenhang von Informationsentnahme und Konstruktion, der als mittelstark eingestuft werden kann. Dies weist darauf hin, dass Informationsentnahme und Konstruktion zumindest einzelne gemeinsame oder sehr ähnliche Kompetenzanteile umfassen. Insbesondere unterscheiden sich die beiden genannten Komponenten nach den Befunden der vorliegenden Arbeit in ihrer Schwierigkeit nicht signifikant voneinander. Insgesamt bereiten diese Ergebnisse die empirische Basis für die Vorschläge Kattmanns (2006) sowie Klipperts (2005) zur Förderung der Diagrammkompetenz. Beide Autoren gehen davon aus, dass das eigenständige Konstruieren im Unterricht zur Förderung der Fähigkeiten zur Informationsentnahme führt. Insbesondere die Modellierung der

Diagrammkompetenz in Form der fünf (Unter-)Komponenten erlaubt hier weitere Empfehlungen.

Die Befunde, dass zwischen den Unterkomponenten, die sich auf den Diagrammrahmen beziehen (Identifizierung und Aufbau), und den Unterkomponenten, die sich auf die Daten beziehen (AbleSEN und Dateneintrag), jeweils keine starken Bezüge bestehen (vgl. 7.1.1), sprechen dagegen, dass beispielsweise eine explizite Schulung des Werteintrags die Ablesefähigkeiten der Schüler steigert. Vielmehr legen die Ergebnisse nahe, vor allem die Fähigkeiten zum Aufbau und damit zur Konstruktion des Diagrammrahmens zu fördern. Aufgrund der starken Bezüge der Unterkomponente Aufbau zum Dateneintrag und darüber hinaus zum AbleSEN kann vermutet werden, dass eine Förderung der Fähigkeiten zum Aufbau sich ebenfalls in einer Erhöhung der Fähigkeiten in den anderen Unterkomponenten und damit übergreifend auf die Informationsentnahme niederschlägt.

Nach den vorliegenden Befunden scheint es generell möglich, auch umgekehrt durch eine Förderung der Informationsentnahme eine Förderung der Konstruktion bewirken zu können (vgl. Kramarski, 2004). Der ermittelte Bezug zwischen AbleSEN und Aufbau lässt vermuten, dass hohe Ablesefähigkeiten für die Konstruktion des Diagrammrahmens hilfreich sein könnten. Ein Befund von Adams und Shrum (1990) liefert allerdings einen Hinweis darauf, dass eine Steigerung der Ablesefähigkeiten zumindest nicht in allen Fällen mit einer Steigerung der Konstruktionsfähigkeiten einhergeht. In ihrer Studie absolvierten die Schüler ein computerbasiertes Training zum Umgang mit Diagrammen. Dieses Training hatte zwar einen positiven Effekt auf die Fähigkeiten zur Informationsentnahme, konnte aber die Konstruktionsleistungen der Schüler nicht steigern.

Neben der Informationsentnahme und der Konstruktion ist die Integration als eigenständige Komponente der Diagrammkompetenz identifiziert worden. Damit wurde empirisch belegt, dass die Fähigkeit, Text und Diagramm als verschiedene Repräsentationsformen (Schnotz, 2001) zu verarbeiten und aufeinander zu beziehen, für den Umgang mit Diagrammen bedeutsam und beachtenswert ist. Die Relevanz dieser Fähigkeit wird normativ unterstrichen durch die explizite Formulierung des Standards „Schülerinnen und Schüler werten Informationen zu biologischen Fragestellungen aus verschiedenen Quellen zielgerichtet aus [...]“ in den Bildungsstandards für die Biologie (KMK, 2005a) und entsprechenden Standards in Physik und Chemie.

Wie durch die Korrelations- und Mediationsanalysen gezeigt werden konnte, steht die Integration in starkem Zusammenhang zur Informationsentnahme. Darüber hinaus wird der festgestellte, gering ausgeprägte Zusammenhang zwischen Integration und Konstruktion überwiegend durch eine Unterkomponente der Informationsentnahme, nämlich das AbleSEN mediiert. Demzufolge ist nicht anzunehmen, dass sich die Fähigkeit zur Integration von Informationen aus Text und Diagramm direkt über die Fähigkeit zur Konstruktion eines Diagramms fördern lässt. Vielmehr ist davon auszugehen, dass eine

Förderung der Integration einerseits direkt über die Informationsentnahme sowie andererseits nur indirekt vermittelt des Ablesens durch die Konstruktion möglich ist. Die bereits diskutierte Möglichkeit, durch Schulung der Konstruktion eine Förderung der Fähigkeiten zur Informationsentnahme zu bewirken (Kattmann, 2006; Klippert, 2005), legt dies ebenfalls nahe. Beispielsweise könnte die Vermittlung und Übung der Fähigkeiten zu den einzelnen Komponenten der Diagrammkompetenz anhand der Datenauswertung zu einem naturwissenschaftlichen Experiment erfolgen. Gemäß dem naturwissenschaftlichen Erkenntnisweg werden zur Überprüfung einer Hypothese zunächst Experimente geplant und durchgeführt. Dabei werden Daten erhoben, die dann mit Hilfe eines Diagramms dokumentiert und visualisiert werden können (Köhler, 2004). Die Konstruktion eines Diagramms ist somit zentraler Bestandteil der Datenauswertung als Teil des naturwissenschaftlichen Erkenntniswegs und könnte so auch im Biologieunterricht geübt werden. Damit werden nach o.g. Überlegungen bereits auch wichtige Grundlagen für die korrekte Entnahme von Informationen aus einem Diagramm geschaffen. Im nächsten Schritt, dem deduktiven Schlussfolgern, das zur Unterstützung oder Ablehnung der Hypothese führt (Graf, 2004) müssen aus dem erstellten Diagramm die relevanten Informationen entnommen und zur Hypothese respektive der dahinterstehenden Theorie in Beziehung gesetzt werden. Dabei können die Fähigkeiten zum Ablesen und Integrieren direkt geschult werden. Es empfiehlt sich in jedem Falle, den Einsatz von Diagrammen im Unterricht in den naturwissenschaftlichen Erkenntnisweg einzubetten und für die Schüler als Teil desselben kenntlich zu machen.

7.6.2 Binnenstrukturen

Die innerhalb des Ablesens beschriebenen Ordnungen ließen sich in der vorliegenden Arbeit nicht empirisch nachweisen. Als eine Erklärung wurde der Einfluss weiterer, bisher nicht in das Modell der Diagrammkompetenz einbezogener Aufgabenmerkmale diskutiert. Daraus leitet sich für den praktischen Einsatz von Diagrammen im Unterricht der Hinweis ab, dass zur Einschätzung der Schwierigkeit bestimmter Ableseanforderungen weitere Aufgabenmerkmale möglicherweise berücksichtigt werden müssen. Dies spielt u. a. bei der Erstellung unterschiedlicher, aber vergleichbarer Versionen von Klassenarbeiten eine Rolle, da der Umgang mit Diagrammen ausdrücklich als Mittel der Leistungsmessung herangezogen werden soll (LP SA, Bio, S. 25). Es muss beachtet werden, dass sich die verwendeten Diagramme, wie oben diskutiert, in der Anordnung ihrer graphischen Komponenten und in ihrer Komplexität unterscheiden können und dadurch eventuell unterschiedlich schwer zu verarbeiten sind (Carpenter & Shah, 1998).

Hinsichtlich der Konstruktion können die in der vorliegenden Arbeit aufgedeckten Schwierigkeiten und häufig auftretenden Fehler bei der Erstellung von Lehrmaterial oder Lernaufgaben berücksichtigt werden. So weist auch Kattmann (2006) darauf hin, bezüglich einer unangemessenen Auswahl eines Liniendiagramms die durch eine Interpolationslinie

induzierte Interpretation mit den Lernenden zu diskutieren respektive das Zeichnen von Verbindungslinien zwischen Punkten zu vermeiden, wenn eine nominal oder ordinal skalierte unabhängige Variable gegeben ist.

7.6.3 Diagrammtypen

Die erörterten Ergebnisse belegen mit den genannten Einschränkungen, dass das in der vorliegenden Arbeit entwickelte und untersuchte Strukturmodell der Diagrammkompetenz für beide Diagrammtypen, Linien- und Säulendiagramme, Gültigkeit hat. Für die Unterrichtspraxis legen diese Befunde nahe, dass die Diagrammkompetenz unter Einsatz beider Diagrammtypen gleichermaßen diagnostiziert und gefördert werden kann. Zur Diagnose der Diagrammkompetenz ist es demnach für eine Lehrperson nicht nötig, Testaufgaben jeweils zu *beiden* Diagrammtypen einzusetzen. Es kann davon ausgegangen werden, dass ein Schüler, der beispielsweise über ein gewisses Maß an Ablesefähigkeiten in Liniendiagrammen verfügt, auch zur Informationsentnahme aus Säulendiagrammen in der Lage ist, wenngleich die Befunde zu den einzelnen Komponenten, wie oben diskutiert, mit Einschränkungen im Hinblick auf die Generalisierbarkeit zu betrachten sind.

8 Ausblick

Mit den in der Diskussion abgeleiteten Implikationen für die Unterrichtspraxis ergeben sich aus den Befunden der vorliegenden Arbeit Ansätze für weiterführende Forschungsarbeiten auf dem Gebiet der Diagrammkompetenz. Die detaillierten Ergebnisse liefern Ausgangspunkte für Implementationsstudien, in denen Möglichkeiten der Förderung von Diagrammkompetenz entwickelt und evaluiert werden können. Wie bereits in der Diskussion angesprochen, stellt die Einbettung des Umgangs mit Diagrammen in den naturwissenschaftlichen Erkenntnisprozess eine interessante Möglichkeit zur Schulung der Diagrammkompetenz dar. Über die Konstruktion eines Diagramms zur Datenaufbereitung könnte die Informationsentnahme anhand der Auswertung der Daten erlernt werden. Im Anschluss stellt die Interpretation der Daten unter Rückgriff auf eine Hypothese und das deduktive Schlussfolgern die Integration dar, in der die Informationen aus dem Diagramm zu den Aussagen in der Hypothese respektive der dahinter stehenden Theorie in Bezug gesetzt werden. Anhand dieses Vorgehens könnten somit ebenfalls Fähigkeiten im Hinblick auf Informationsentnahme und Integration als weitere Komponenten der Diagrammkompetenz gefördert werden. Das in der vorliegenden Arbeit entwickelte Testinstrument kann dabei genutzt werden, um die Wirksamkeit der untersuchten Förderungsmaßnahmen zu messen.

Im Bereich der Forschung zur Kompetenzmodellierung sind weitere Perspektiven zu sehen. Das in der vorliegenden Arbeit theoretisch abgeleitete Strukturmodell der Diagrammkompetenz konnte empirisch überprüft und somit validiert werden. Damit wurde gezeigt, dass die Modellierung einer Kompetenzstruktur, die sich auf konkrete Aspekte der Bildungsstandards bzw. der naturwissenschaftlichen Methodik bezieht, prinzipiell zu leisten ist, und dass sich das entwickelte Kompetenzstrukturmodell validieren und zur Leistungsmessung im Unterricht einsetzen lässt. Ein weiteres Ziel der aktuellen Bildungsforschung ist es, die Entwicklungsverläufe einer solchen Kompetenz bei Schülern aufzudecken und durch Kompetenzentwicklungsmodelle (Klieme, 2003) zu beschreiben. Die in der vorliegenden Arbeit beschriebene Kompetenzstruktur kann den Ausgangspunkt hierzu liefern. Sie wurde insbesondere für die zehnte Klassenstufe und damit für diejenige Schülerteilpopulation empirisch abgesichert, auf die sich die Zielvorgabe der Bildungsstandards für den mittleren Schulabschluss (z. B. KMK, 2005a) richtet. Als erster Hinweis auf eine Kompetenzentwicklung kann der Befund der vorliegenden Arbeit gewertet werden, dass Fähigkeiten der Schüler in allen drei Komponenten der Diagrammkompetenz bereits in der neunten Klasse vorhanden sind, wenngleich im Hinblick auf die Informationsentnahme und die Integration in geringerem Maße als bei den Zehntklässlern. Die Klärung, ob sich diese Strukturen auch in der Unterstufe nachweisen lassen und ob sich die beschriebenen Relationen zwischen den Komponenten dort unterscheiden, wäre Aufgabe der weiteren Forschung.

Weiterhin ergibt sich die Frage nach der Fachspezifität von Diagrammkompetenz. In der vorliegenden Arbeit wurde die Diagrammkompetenz am Beispiel der Biologie als einem naturwissenschaftlichen Unterrichtsfach untersucht. Dabei hat sich insbesondere herausgestellt, dass das biologische Fachwissen der Schüler keinen nennenswerten Einfluss auf deren Fähigkeiten zum Lesen und Konstruieren von Diagrammen hat. Es ist möglich, dass ein Einfluss des Fachwissens in den anderen Naturwissenschaften Chemie und Physik deutlich stärker ausgeprägt ist. In beiden Fächern werden häufig auf spezifische Unterrichtsinhalte ausgerichtete Diagramme eingesetzt, wie beispielsweise Weg-Zeit-Diagramme und Geschwindigkeits-Zeit-Diagramme in der Physik sowie Zustandsdiagramme in der Chemie, in welchen die Phasenzusammensetzung eines Materials beschrieben wird. Für den Umgang mit solchen Diagrammen scheint somit ein Verständnis der dahinter liegenden Konzepte grundlegend. Insbesondere das Verständnis von Geschwindigkeit scheint maßgeblich an der Interpretation von Kurvenverläufen in den Liniendiagrammen zum Weg-Zeit-Verhältnis beteiligt. Es zeigen sich insbesondere Schwierigkeiten von Schülern (Bell & Janvier, 1981) – und auch Lehrern (Stump, 1997) – bei der Interpretation von Steigungen in solchen Liniendiagrammen. Ausgehend von diesen Betrachtungen stellt sich die Frage, ob sich die Komponenten der Diagrammkompetenz, die in der vorliegenden Arbeit für Inhalte des Biologieunterrichts identifiziert wurden, in der gleichen Weise auch Chemie und Physik nachweisen lassen. Möglicherweise besteht in diesen Fächern aufgrund des Einflusses inhaltlichen Wissens ein noch stärkerer Zusammenhang zwischen den Fähigkeiten der Informationsentnahme und der Integration. Als Forschungsdesiderat ergibt sich daraus die fächerübergreifende Absicherung und gegebenenfalls Weiterentwicklung des in der vorliegenden Arbeit beschriebenen Strukturmodells der Diagrammkompetenz.

9 Literaturverzeichnis

- Åberg-Bengtsson, L. & Ottosson, T. (2006). What lies behind graphicacy? Relating students' results on a test of graphically represented quantitative information to formal academic achievement. *Journal of Research in Science Teaching*, *43*, 43-62.
- Adams, R. J. & Khoo, S. H. (1996). *Quest [Computer software]*. Melbourne, Australia: Australian Council for Educational Research.
- Adams, R. J., Wilson, M. & Wang, W. (1997). The multidimensional random coefficients multinomial logit model. *Applied Psychological Measurement*, *21*, 1-23.
- Ainsworth, S. E. (1999). A functional taxonomy of multiple representations. *Computers and Education*, *33*, 131-152.
- Ainsworth, S. E., Bibby, P. A. & Wood, D. J. (2002). Examining the effects of different multiple representational systems in learning primary mathematics. *Journal of the Learning Sciences*, *11*, 25-62.
- Andreß, H. J. (2001). *Glossar zur Datenerhebung und statistischen Analyse*. Verfügbar unter: <http://psydok.sulb.uni-saarland.de/volltexte/2004/268/> [10.04.2008].
- APA (American Psychological Association) (Ed.). (2001). *Publication Manual of the American Psychological Association* (5 ed.). Washington, DC: American Psychological Association.
- Apra, C. & Ebner, H. G. (2001). Elaboration der Aneignung wirtschaftsberuflichen Wissens durch Generierung von Diagrammen. In H. Reinisch, R. Bader & S. G.A. (Eds.), *Modernisierung der Berufsbildung in Europa. Neue Befunde der berufspädagogischen Forschung* (pp. 63-72): Opladen.
- Apra, C. & Ebner, H. G. (2003). Generierung von Diagrammen als Lernhandlung: Effekte eines Kurzzeittrainings zur Förderung der Text-Graphik-Transformation. In F. Achtenhagen & E. G. John (Eds.), *Meilensteine der beruflichen Bildung* (Vol 1) (pp. 117-138). Bielefeld: Bertelsmann.
- Artelt, C., Stanat, P., Schneider, W. & Schiefele, U. (2001). Lesekompetenz: Testkonzeption und Ergebnisse. In Deutsches PISA-Konsortium (Ed.), *PISA 2000: Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich* (pp. 69-137). Opladen: Leske+Budrich.
- Backhaus, K., Erichson, B., Plinke, W. & Weiber, R. (2003). *Multivariate Analysemethoden: Eine anwendungsorientierte Einführung*. Berlin: Springer.
- Baker, F. (2001). *The Basics of Item Response Theory* (2nd. ed.). University of Maryland, College Park, MD: ERIC Clearinghouse on Assessment and Evaluation.
- Baker, R. S., Corbett, A. T. & Koedinger, K. R. (2001). Toward a model of learning data representations. In J. D. Moore & K. Stenning (Eds.), *Proceedings of the Twenty-Third Annual Conference of the Cognitive Science Society* (pp. 45-50). Mahwah, NJ: Erlbaum.

- Baker, R. S., Corbett, A. T. & Koedinger, K. R. (2002). *The Resilience of Overgeneralization of Knowledge about Data Representations*. Paper presented at the American Educational Research Association Conference 2002.
- Baron, R. M. & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic and statistical considerations. *Journal of Personality and Social Psychology*, 51, 1173-1182.
- Baumert, J., Stanat, P. & Demmrich, A. (2001). PISA 2000: Untersuchungsgegenstand, theoretische Grundlagen und Durchführung der Studie. In D. PISA-Konsortium (Ed.), *PISA 2000: Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich* (pp. 15-68). Opladen: Leske+Budrich.
- Baumert, J., Lehmann, R., Lehrke, M., Clausen, M., Hosenfeld, I., Neubrand, J., et al. (Eds.). (1998a). *Testaufgaben Mathematik TIMSS 7./8. Klasse (Population 2)*. Berlin: Max-Planck-Institut für Bildungsforschung.
- Baumert, J., Lehmann, R., Lehrke, M., Clausen, M., Hosenfeld, I., Neubrand, J., et al. (Eds.). (1998b). *Testaufgaben Naturwissenschaften TIMSS 7./8. Klasse (Population 2)*. Berlin: Max-Planck-Institut für Bildungsforschung.
- Baumert, J., Bos, W., Klieme, E., Lehmann, R., Lehrke, M., Hosenfeld, I., et al. (Eds.). (1999). *Testaufgaben zu TIMSS/III Mathematisch-naturwissenschaftliche Grundbildung und voruniversitäre Mathematik und Physik der Abschlußklassen der Sekundarstufe II (Population 3)*. Berlin: Max-Planck-Institut für Bildungsforschung.
- Bayrhuber, H. & Kull, U. (Eds.). (2005). *Linder Biologie: Lehrbuch für die Oberstufe* (Vol. 22). Braunschweig: Schroedel.
- Bell, A. & Janvier, C. (1981). The interpretation of graphs representing situations. *For the Learning of Mathematics*, 2, 34-42.
- Berg, C. A. & Smith, P. (1994). Assessing students' abilities to construct and interpret line graphs: Disparities between multiple choice and free-response instruments. *Science Education*, 78, 527-554.
- Bertin, J. (1974). *Graphische Semiologie: Diagramme, Netze, Karten* (G. Jensch, D. Schade & W. Scharfe, Trans.). Berlin: de Gruyter.
- Bickel, H., Claus, R., Frank, R., Gropengießer, H., Haala, G., Knauer, B., et al. (1995). *Natura Band 3: Biologie für Gymnasien, Oberstufe*. Stuttgart: Ernst Klett Schulbuchverlag.
- Bodemer, D., Plötzner, R., Feuerlein, I. & Spada, H. (2004). The active integration of information during learning with dynamic and interactive visualisations. *Learning and Instruction*, 14, 325-341.
- Bohl, E. (2006). *Mathematik in der Biologie* (4. ed.). Berlin: Springer.
- Bortz, J. (1999). *Statistik für Sozialwissenschaftler* (5. ed.). Berlin: Springer.

- Bowen, G. M. & Roth, W.-M. (2005). Data and graph interpretation practices among preservice science teachers. *Journal of Research in Science Teaching*, 42, 1063-1088.
- Burnham, K. P. & Anderson, D. R. (2004). Multimodel inference: Understanding AIC and BIC in model selection. *Sociological Methods Research*, 33, 261-304.
- Carney, R. N. & Levin, J. R. (2002). Pictorial illustrations still improve students' learning from text. *Educational Psychology Review*, 14, 5-26.
- Carpenter, P. A. & Shah, P. (1998). A model of the perceptual and conceptual processes in graph comprehension. *Journal of Experimental Psychology: Applied*, 4, 75-100.
- Chandler, P. & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction*, 8, 293-332.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. New York: Erlbaum.
- Cox, R. (1999). Representation construction, externalised cognition and individual differences. *Learning and Instruction*, 9, 343-363.
- Curcio, F. R. (1987). Comprehension of the mathematical relationships expressed in graphs. *Journal for Research in Mathematics Education*, 18, 382-393.
- Deutsches PISA-Konsortium (Ed.). (2000). *Schülerleistungen im internationalen Vergleich: Eine Rahmenkonzeption für die Erfassung von Wissen und Fähigkeiten*. Berlin: Max-Planck-Institut für Bildungsforschung.
- Van Dijk, T. A. & Kintsch, W. (1983). *Strategies of discourse comprehension*. New York: Academic Press.
- Dreyfus, T. & Eisenberg, T. (1990). On difficulties with diagrams: Theoretical issues. In G. Booker, P. Cobb & T. N. De Mendicuti (Eds.), *Proceedings of the 14th Annual Conference of the International Group for the Psychology of Mathematics Education* (Vol 1) (pp. 27-36). Oaxtepe: PME.
- Dytham, C. (2005). *Choosing and Using Statistics: A Biologist's Guide*: Blackwell.
- Einhaus, E. (2007). *Schülerkompetenzen im Bereich Wärmelehre - Entwicklung eines Testinstruments zur Überprüfung und Weiterentwicklung eines normativen Modells fachbezogener Kompetenzen*. Berlin: Logos.
- Feeney, A., Holo, A. K. W., Liversedge, S. P., Findlay, J. M. & Metcalf, R. (2000). How People Extract Information from Graphs: Evidence from a Sentence-Graph Verification Paradigm. In *Diagrams 2000*. Berlin: Springer.
- Feeney, A. & Webber, L. (2003). Analogical Representation and Graph Comprehension. In A. Butz, A. Krüger & P. Olivier (Eds.), *Smart Graphics: Third International Symposium on Smart Graphics, SG 2003 Heidelberg, Germany, July 2-4, 2003 Proceedings* (pp. 212-221). Berlin: Springer LNCS 2733.

- Fisseni, H.-J. (1997). *Lehrbuch der psychologischen Diagnostik* (2. ed.). Göttingen: Hogrefe.
- Flindt, R. (2000). *Biologie in Zahlen* (Vol. 5). Heidelberg: Spektrum Akademischer Verlag.
- Freedman, E. & Shah, P. (2002). Toward a model of knowledge-based graph comprehension. In M. Hegarty, B. Meyer & H. Narayanan (Eds.), *Diagrammatic representation and inference* (pp. 18-31). Berlin: Springer.
- Friege, G. & Lind, G. (2003). Allgemeine und fachspezifische Problemlösekompetenz. *Zeitschrift für die Didaktik der Naturwissenschaften*, 9, 63-74.
- Friel, S. N., Curcio, F. R. & Bright, G. W. (2001). Making sense of graphs: Critical factors influencing comprehension and instructional implications. *Journal for Research in Mathematics Education*, 32, 124-158.
- Gillan, D. J. & Lewis, R. (1994). A componential model of human interaction with graphs: 1. Linear regression modeling. *Human Factors*, 36, 419-440.
- Goodman, L. A. (1960). On the exact variance of products. *Journal of the American Statistical Association*, 55, 708-713.
- Graf, E. (2004). Erkenntnismethoden und problemlösendes Vorgehen im Biologieunterricht. In E. Graf (Ed.), *Biologiedidaktik für Studium und Unterrichtspraxis* (pp. 30-37). Donauwörth: Auer.
- Hambleton, R. K. & Jones, R. W. (1993). An NCME Instructional Module on Comparison of Classical Test Theory and Item Response Theory and Their Applications to Test Development. *Educational Measurement: Issues and Practice*, 12, 38-47.
- Hammann, M. (2004). Kompetenzentwicklungsmodelle: Merkmale und ihre Bedeutung - dargestellt anhand von Kompetenzen beim Experimentieren. *MNU*, 57, 196-203.
- Hartig, J. & Klieme, E. (2006). Kompetenz und Kompetenzdiagnostik. In K. Schweizer (Ed.), *Leistung und Leistungsdiagnostik* (pp. 127-143). Heidelberg: Springer Medizin.
- Hartig, J. (2007). Skalierung und Definition von Kompetenzniveaus. In B. Beck & E. Klieme (Eds.), *Sprachliche Kompetenzen. Konzepte und Messung - DESI-Studie* (pp. 83-99). Weinheim: Beltz.
- Heiser, J. & Tversky, B. (2006). Arrows in comprehending and producing mechanical diagrams. *Cognitive Science*, 30, 581-592.
- Issing, L. J. (1983). Bilder als didaktische Medien. In L. J. Issing & J. Hannemann (Eds.), *Lernen mit Bildern* (pp. 9-39). Grünewald: Institut für Film und Bild in Wissenschaft und Unterricht.
- Jacobs, B. (1994). *Der Einfluß von Graphotyp und Graphanordnung auf das Graphverstehen bei der Analyse von Verläufen*. Arbeitsbericht Nr. 13. Medienzentrum der Philosophischen Fakultät der Universität des Saarlandes.

- Verfügbar unter: <http://psydok.sulb.uni-saarland.de/volltexte/2003/37/pdf/graphex1.pdf> [20.04.2008].
- Jacobs, B. (1995a). *Globale Vergleiche, lokale Vergleiche und Größenschätzungen in Liniendiagramm und Säulendiagramm unter Superposition und Juxtaposition*. Arbeitsbericht Nr. 16. Medienzentrum der Philosophischen Fakultät der Universität des Saarlandes. Verfügbar unter: <http://psydok.sulb.uni-saarland.de/volltexte/2003/33/pdf/graphex3.pdf> [08.05.2008].
- Jacobs, B. (1995b). *Experimentelle Analysen zur Wahrnehmung von Kurvenverläufen und Kurvenvergleichen in Säulendiagramm und Liniendiagramm unter Superposition und Juxtaposition*. Arbeitsbericht Nr. 15. Medienzentrum der Philosophischen Fakultät der Universität des Saarlandes. Verfügbar unter: <http://psydok.sulb.uni-saarland.de/volltexte/2003/32/pdf/graphex2.pdf> [08.05.2008].
- Janssen, J. & Laatz, W. (2005). *Statistische Datenanalyse mit SPSS für Windows* (5. ed.). Berlin: Springer.
- Janvier, C. (1981). Use of situations in mathematics education. *Educational Studies in Mathematics*, 12, 113-122.
- de Jong, T. & van Joolingen, W. R. (1998). Scientific discovery learning with computer simulations of conceptual domains. *Review of Educational Research*, 68, 179-201.
- Kattmann, U. (2006). Diagramme. In H. Gropengießer & U. Kattmann (Eds.), *Fachdidaktik Biologie. Die Biologiedidaktik begründet von Dieter Eschenhagen, Ulrich Kattmann und Dieter Rodi* (pp. 340-356). Köln: Aulis.
- Katz, I. R., Xi, X., Kim, H. & Cheng, P. C.-H. (2002). Graph structure supports graph description. In *Proceedings of the Twenty-fourth Annual Meeting of the Cognitive Science Society* (pp. 157-162). Hillsdale, NJ: Erlbaum.
- Kerslake, D. (1981). Graphs. In K. Hart, D. Kerslake, M. L. Brown, et al. (Eds.), *Children's Understanding of Mathematics: 11-16* (pp. 120-136). New York: John Murray.
- Klahr, D. & Dunbar, K. (1988). Dual space search during scientific reasoning. *Cognitive Science*, 12, 1-48.
- Klieme, E., Avenarius, H., Blum, W., Döbrich, P., Gruber, H., Prenzel, M., et al. (2003). *Zur Entwicklung nationaler Bildungsstandards*. Bonn: Bundesministerium für Bildung und Forschung.
- Klieme, E. & Leutner, D. (2006). *Kompetenzmodelle zur Erfassung individueller Lernergebnisse und zur Bilanzierung von Bildungsprozessen. Überarbeitete Fassung des Antrags an die DFG auf Einrichtung eines Schwerpunktprogramms*. Verfügbar unter: <http://www.kompetenzdiagnostik.de> [09.11.2007].
- Klieme, E. & Hartig, J. (2008). Kompetenzkonzepte in den Sozialwissenschaften und im erziehungswissenschaftlichen Diskurs. In M. Prenzel, I. Gogolin & H.-H. Krüger (Eds.), *Kompetenzdiagnostik: Sonderheft der Zeitschrift für Erziehungswissenschaft 8/2007*. Wiesbaden: VS Verlag für Sozialwissenschaften.

- Klippert, H. (2006). *Methoden-Training: Übungsbausteine für den Unterricht* (Vol. 16). Weinheim: Beltz.
- KMK (Ständige Konferenz der Kultusminister der Länder der Bundesrepublik Deutschland) (Ed.). (2005a). *Bildungsstandards im Fach Biologie für den Mittleren Schulabschluss: Beschluss vom 16.12.2004*. München: Wolters Kluwer.
- KMK (Ständige Konferenz der Kultusminister der Länder der Bundesrepublik Deutschland) (Ed.). (2005b). *Bildungsstandards im Fach Chemie für den Mittleren Schulabschluss: Beschluss vom 16.12.2004*. München: Wolters Kluwer.
- KMK (Ständige Konferenz der Kultusminister der Länder der Bundesrepublik Deutschland) (Ed.). (2005c). *Bildungsstandards im Fach Physik für den Mittleren Schulabschluss: Beschluss vom 16.12.2004*. München: Wolters Kluwer.
- KMK (Ständige Konferenz der Kultusminister der Länder der Bundesrepublik Deutschland) (Ed.). (2005d). *Bildungsstandards der Kultusministerkonferenz. Erläuterungen zur Konzeption und Entwicklung*. München: Wolters Kluwer.
- Köhler, K. (2004). Welche fachgemäßen Arbeitsweisen werden im Biologieunterricht eingesetzt? In U. Spörhase-Eichmann & W. Ruppert (Eds.), *Biologiedidaktik: Praxishandbuch für die Sekundarstufe I und II* (pp. 146-159). Berlin: Cornelsen Scriptor.
- Köhler, W., Schachtel, G. A. & Voleske, P. (2002). *Biostatistik: Eine Einführung für Biologen und Agrarwissenschaftler* (3. ed.). Berlin: Springer.
- Kosslyn, S. M. (1989). Understanding charts and graphs. *Applied Cognitive Psychology*, 3, 185-225.
- Kozma, R. B. & Russell, J. (1997). Multimedia and understanding: Expert and novice responses to different representations of chemical phenomena. *Journal of Research in Science Teaching*, 34, 949-968.
- Kramarski, B. (2004). Making sense of graphs: does metacognitive instruction make a difference on students' mathematical conceptions and alternative conceptions? *Learning and Instruction*, 14, 593-619.
- Kramer, B., Prechtel, H. & Bayrhuber, H. (2005). Text- und Bildverarbeitungsstrategien bei der Nutzung einer multimedialen Lernumgebung. In R. Klee, A. Sandmann & H. Vogt (Eds.), *Lehr- und Lernforschung in der Biologiedidaktik (Bd. 2)* (pp. 165-179). Innsbruck: Studienverlag.
- Krohn, R. (1991). Why are graphs so central in science? *Biology and Philosophy*, 6, 181-203.
- Larkin, J. & Simon, H. (1987). Why a diagram is (sometimes) worth ten thousands words. *Cognitive Science*, 11, 65-99.
- Leinhardt, G., Zaslavsky, O. & Stein, M. S. (1990). Functions, graphs and graphing: Tasks, learning and teaching. *Review of Educational Research*, 1, 1-64.

- Lewalter, D. (1997). *Lernen mit Bildern und Animationen: Studie zum Einfluss von Lernermerkmalen auf die Effektivität von Illustrationen*. Münster: Waxmann.
- Lind, D. & Knoche, N. (2004). Testtheoretische Modelle und Verfahren bei PISA-2000-Mathematik. In M. Neubrand (Ed.), *Mathematische Kompetenzen von Schülerinnen und Schülern in Deutschland: Vertiefende Analysen im Rahmen von PISA 2000* (pp. 51-69). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Lohse, G. L. (1993). A cognitive model for understanding graphical perception. *Human Computer Interaction*, 8, 353-388.
- Lücken, M. (2007). "Hat mein neu entwickeltes Unterrichtskonzept den gewünschten Erfolg?" - Ein Einblick in den Umgang mit statistischen Prüfverfahren in der empirischen Unterrichtsforschung. In H. Bayrhuber, D. Elster, D. Krüger, et al. (Eds.), *Kompetenzentwicklung und Assessment*. Innsbruck: Studienverlag.
- Maichle, U. (1994). Cognitive processes in understanding line graphs. In W. Schnotz & R. W. Kulhavy (Eds.), *Comprehension of Graphics* (pp. 207-226). Amsterdam: North-Holland.
- Masters, G. N. (1982). A Rasch model for partial credit scoring. *Psychometrika*, 47, 149-174.
- Mayer, R. E. & Gallini, J. K. (1990). When is an illustration worth ten thousand words? *Journal of Educational Psychology*, 82, 715-726.
- Mayer, R. E. (2003). The promise of multimedia learning: using the same instructional design methods across different media. *Learning and Instruction*, 13, 125-139.
- Mayring, P. (2000). Qualitative Inhaltsanalyse. *Forum Qualitative Sozialforschung [Online Journal]*, 1.
- McKenzie, D. L. & Padilla, M. J. (1986). The construction and validation of the test of graphing in science (togs). *Journal of Research in Science Teaching*, 23, 571-579.
- Mevarech, Z. R. & Kramarski, B. (1997). From verbal descriptions to graphic representations: Stability and change in students' alternative conceptions. *Educational Studies in Mathematics*, 32, 229-263.
- Meyer, G. (1988). Mathematische Aspekte im Biologieunterricht. *Unterricht Biologie*, 12, 4-13.
- Nerdel, C., Precht, H. & Bayrhuber, H. (2003). Interactive animations and understanding of biological processes: An empirical investigation on the effectiveness of computer-assisted learning environments in biology instruction. In J. Lewis, A. Margo & L. Simonneaux (Eds.), *Biology Education For The Real World: Student - Teacher - Citizen* (pp. 45-58). Toulouse, France: Ecole Nationale de Formation Agronomique (ENFA).
- Novick, L. R. & Hurley, S. M. (2001). To matrix, network, or hierarchy: That is the question. *Cognitive Psychology*, 42, 158-216.

- OECD (Organisation for Economic Co-operation and Development). (2000). *PISA 2000 - Beispielaufgaben aus dem Lesekompetenztest (sowie zugehörige Lösungen)*. Verfügbar unter: [http://www.mpib-berlin.mpg.de/pisa/Loesungen_Lesen.pdf, http://www.mpib-berlin.mpg.de/pisa/Beispielaufgaben_Lesen.pdf].
- OECD (Organisation for Economic Co-operation and Development) (Ed.). (2005). *PISA 2003 Technical Report*. Paris: OECD Publishing.
- Padilla, M. J., McKenzie, D. L. & Shaw, E. L. (1986). An examination of the line graphing ability of students in grades seven through twelve. *School Science and Mathematics*, 86, 20-26.
- Pereira-Mendoza, L. & Mellor, J. (1991). Students' concepts of bar graphs: Some preliminary findings. In D. Vere-Jones (Ed.), *Proceedings of the Third International Conference on Teaching Statistics* (Vol 1) (pp. 150-157). Voorburg: International Statistical Institute.
- Pinker, S. (1990). A theory of graph comprehension. In R. O. Freedle (Ed.), *Artificial Intelligence and the future of testing* (pp. 73-126). Hillsdale, NJ: Erlbaum.
- PISA-Konsortium Deutschland. (2006b). *PISA 2003: Untersuchungen zur Kompetenzentwicklung im Verlauf eines Schuljahres - Zusammenfassung*. Verfügbar unter: http://pisa.ipn.uni-kiel.de/PISA_2003_Kompetenzentwicklung_Zusfsg.pdf [20.03.2008].
- PISA-Konsortium Deutschland (Ed.). (2005). *PISA 2003: Der Bildungsstand der Jugendlichen in Deutschland - Ergebnisse des zweiten internationalen Vergleichs*. Münster: Waxmann.
- PISA-Konsortium Deutschland (Ed.). (2006a). *PISA 2003: Dokumentation der Erhebungsinstrumente*. Münster: Waxmann.
- Preacher, K. J. (2002). *Calculation for the test of the difference between two independent correlation coefficients [Computer software]*. Available from <http://www.quantpsy.org>.
- Preece, J. & Janvier, C. (1992). A study of the interpretation of trends in multiple curve graphs in ecological situations. *School Science and Mathematics*, 92, 299-363.
- Prenzel, M., Rost, J., Senkbeil, M., Häußler, P. & Klopp, A. (2001). Naturwissenschaftliche Grundbildung: Testkonzeption und Ergebnisse. In Deutsches PISA-Konsortium (Ed.), *PISA 2000: Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich* (pp. 191-248). Opladen: Leske+Budrich.
- Prenzel, M., Häußler, P., Rost, J. & Senkbeil, M. (2002). Der PISA-Naturwissenschaftstest: Lassen sich die Aufgabenschwierigkeiten vorhersagen? *Unterrichtswissenschaft*, 30, 120-135.
- Ratwani, R. M. & Trafton, J. G. (in press). Shedding Light on the Graph Schema: Perceptual features vs. invariant structure. *Psychonomics Bulletin and Review*.

- Reulecke, W. & Rollett, B. (1976). Pädagogische Diagnostik und lernzielorientierte Tests. In K. Pawlik (Ed.), *Diagnose der Diagnostik: Beiträge zur Diskussion der psychologischen Diagnostik in der Verhaltensmodifikation* (pp. 177-202). Stuttgart: Klett.
- Rinck, M. & Glowalla, U. (1994). Wissensstrukturierung durch statistische Graphen: Weitere Auswirkungen auf das Verstehen. *Zeitschrift für experimentelle und angewandte Psychologie*, 49, 132-153.
- Rost, J. (1996). *Lehrbuch Testtheorie/Testkonstruktion*. Bern: Verlag Hans Huber.
- Rost, J. (2004). *Lehrbuch Testtheorie - Testkonstruktion* (2. ed.). Bern: Hans Huber.
- Rost, J. & Walter, O. (2006). Multimethod Item Response Theory. In M. Eid & E. Diener (Eds.), *Handbook of Multimethod Measurement in Psychology* (pp. 249-268). Washington, DC: American Psychological Association.
- Roth, H. (1971). *Pädagogische Anthropologie. Band 2: Entwicklung und Erziehung*. Hannover.
- Roth, W.-M., Bowen, G. M. & McGinn, M. K. (1999). Differences in graph-related practices between high school biology textbooks and scientific ecology journals. *Journal of Research in Science Teaching*, 36, 977-1019.
- Roth, W.-M. & Bowen, G. M. (2003). When are graphs worth ten thousand words? An expert-expert study. *Cognition and Instruction*, 21, 429-473.
- Rychen, D. S. & Salganik, L. H. (Eds.). (2001). *Defining and Selecting Key Competencies*. Seattle: Hogrefe & Huber.
- Salomon, G. (1984). Television is "easy" and print is "tough": The differential investment of mental effort in learning as a function of perceptions and attributions. *Journal of Educational Psychology*, 76, 647-658.
- Sauermost, R. & Freudig, D. (Eds.). (2000). *Lexikon der Biologie*. Heidelberg: Spektrum Akademischer Verlag.
- Scaife, M. & Rogers, Y. (1996). External Cognition: how do Graphical Representations work? *International Journal of Human-Computer Studies*, 45, 185-213.
- Schaffner, E., Schiefele, U., Drechsel, B. & Artelt, C. (2004). Lesekompetenz. In P.-K. Deutschland (Ed.), *PISA 2003: Der Bildungsstand der Jugendlichen in Deutschland - Ergebnisse des zweiten internationalen Ländervergleichs* (pp. 93-110). Münster: Waxmann.
- Schecker, H. & Parchmann, I. (2006). Modellierung naturwissenschaftlicher Kompetenz. *Zeitschrift für die Pädagogik der Naturwissenschaften*, 12, 45-66.
- Schnotz, W. (1994). Wissenserwerb mit logischen Bildern. In B. Weidenmann (Ed.), *Wissenserwerb mit Bildern: Instruktionale Bilder in Printmedien, Film/Video und Computerprogrammen* (pp. 95-147). Bern: Hans Huber.

- Schnotz, W., Zink, T. & Pfeiffer, M. (1996). Visualisierungen im Lehr-Lern-Prozeß. *Zeitschrift für Pädagogik*, 42, 193-213.
- Schnotz, W. (2001). Wissenserwerb mit Multimedia. *Unterrichtswissenschaft*, 29, 292-318.
- Schnotz, W. (2002). Wissenserwerb mit Texten, Bildern und Diagrammen. In L. Issing & P. Klimsa (Eds.), *Information und Lernen mit Multimedia und Internet* (pp. 65-81). Weinheim: Beltz/PVU.
- Schnotz, W. & Bannert, M. (2003). Construction and interference in learning from multiple representation. *Learning and Instruction*, 13, 141-156.
- Senkbeil, M., Rost, J., Carstensen, C. H. & Walter, O. (2005). Der nationale Naturwissenschaftstest PISA 2003: Entwicklung und empirische Überprüfung eines zweidimensionalen Facettendesigns. *Empirische Pädagogik*, 19, 166-189.
- Seufert, T. (2003). Supporting coherence formation in learning from multiple representations. *Learning and Instruction*, 13, 227-237.
- Shah, P. & Carpenter, P. A. (1995). Conceptual limitations in comprehending line graphs. *Journal of Experimental Psychology: General*, 124, 43-61.
- Shah, P., Mayer, R. E. & Hegarty, M. (1999). Graphs as aids to knowledge construction: Signaling techniques for guiding the process of graph comprehension. *Journal of Educational Psychology*, 91, 690-702.
- Shah, P. (2001). Graph comprehension: The role of format, content, and individual differences. In M. Anderson, B. Meyer & P. Olivier (Eds.), *Diagrammatic Representation and Reasoning* (pp. 173-185). Berlin: Springer.
- Shah, P. & Hoeffner, J. (2002). Review of graph comprehension research: Implications for instruction. *Educational Psychology Review*, 14, 47-69.
- Sobel, M. E. (1982). Asymptotic confidence intervals for indirect effects in structural equation models. In S. Leinhardt (Ed.), *Sociological Methodology* (pp. 290-312). Washington DC: American Sociological Association.
- Stage, C. & Ögren, G. (2004). *The Swedish Scholastic Assessment Test (SweSAT): Development, Results and Experiences*. Artikel EM No 49: Umea University.
- Stanat, P. & Kunter, M. (2002). Geschlechterspezifische Leistungsunterschiede bei Fünfzehnjährigen im internationalen Vergleich. *Zeitschrift für Erziehungswissenschaft*, 5, 28 - 48.
- Stern, E., Aprea, C. & Ebner, H. G. (2003). Improving cross-content transfer in text processing by means of active graphical representation. *Learning and Instruction*, 13, 191-203.
- Stump, S. L. (1997). *Secondary mathematics teachers' knowledge of concept of slope*. Paper presented at the Annual Meeting of the American Educational Research Association, Chicago, IL.

- Trickett, S. B. & Trafton, J. G. (2004). Spatial transformations in graph comprehension. In A. Blackwell, K. Marriott & A. Shimojima (Eds.), *Diagrammatic Representation and Inference: Third International Conference, Diagrams 2004, Cambridge, UK, March 22-24, 2004. Proceedings* (pp. 372 - 375).
- Urban, D. & Mayerl, J. (2006). *Regressionsanalyse: Theorie, Technik und Anwendung* (2. ed.). Wiesbaden: VS Verlag.
- Vogel, G. & Angermann, H. (1992). *dtv-Atlas zur Biologie: Tafeln und Texte, Band 3*. München: Deutscher Taschenbuch Verlag.
- Wainer, H. (1992). Understanding graphs and tables. *Educational Researcher*, 21, 14-23.
- Walter, O. (2005). *Kompetenzmessung in den PISA-Studien. Simulationen zur Schätzung von Verteilungsparametern und Reliabilitäten*. Lengerich: Pabst Science Publishers.
- Warm, T. A. (1989). Weighted likelihood estimation of ability in item response theory. *Psychometrika*, 54, 427-450.
- Wavering, M. J. (1989). Logical reasoning necessary to make line graphs. *Journal of Research in Science Teaching*, 26, 373-379.
- Webber, L. & Feeney, A. (2004). *Spatial Representations for Graphical Reasoning*. Paper presented at the Spatial Cognition 2004 International Conference, 11.-13. October 2004.
- Weidenmann, B. (1988). *Psychische Prozesse beim Verstehen von Bildern*. Bern: Huber.
- Weinert, F. E. (2001a). Vergleichende Leistungsmessung in Schulen - eine umstrittene Selbstverständlichkeit. In F. E. Weinert (Ed.), *Leistungsmessung in Schulen* (pp. 17-31). Weinheim: Beltz.
- Weinert, F. E. (2001b). Concept of Competence: A Conceptual Clarification. In D. S. Rychen & L. H. Salganik (Eds.), *Defining and Selecting Key Competencies* (pp. 45-65). Seattle: Hogrefe & Huber.
- Wilson, M. (1992). The ordered partition model: An extension of the partial credit model. *Applied Psychological Measurement*, 16, 309-325.
- Wilson, M. (2005). *Constructing Measures: An Item Response Modeling Approach*. Mahwah: Lawrence Erlbaum Associates.
- Winn, B. (1987). Charts, graphs, and diagrams in educational materials. In H. A. Houghton & D. M. Willows (Eds.), *The Psychology of Illustration* (Vol 1) (pp. 152-198). New York: Springer.
- Wirtz, M. & Caspar, F. (2002). *Beurteilerübereinstimmung und Beurteilerreliabilität: Methoden zur Bestimmung und Verbesserung der Zuverlässigkeit von Einschätzungen mittels Kategoriensystemen und Ratingskalen*. Göttingen: Hogrefe.

- Wu, M. & Adams, R. (2006). Modelling mathematics problem solving item responses using a multidimensional IRT model. *Mathematics Education Research Journal*, 18, 93-113.
- Wu, M. L., Adams, R. J. & Wilson, M. R. (1998). *ACER ConQuest: Generalised Item Response Modeling Software [Computerprogramm und Manual]*. Camberwell, Victoria: Australian Council for Educational Research.
- Zacks, J. & Tversky, B. (1999). Bars and lines: A study of graphic communication. *Memory & Cognition*, 27, 1073-1079.
- Zelazny, G. (2001). *Say it with Charts: The Executive's Guide to Visual Communication* (4. ed.): McGraw-Hill.

10 Lehrplanverzeichnis

LP BW	Ministerium für Kultus, Jugend und Sport (2004). <i>Bildungsplan 2004 für allgemein bildendes Gymnasium</i> . Verfügbar unter: http://www.bildungstaerkt-Menschen.de/service/downloads/Bildungsplaene/Gymnasium/Gymnasium_Bildungsplan_Gesamt.pdf [24.05.2008]
LP Berlin	Senatsverwaltung für Bildung, Jugend und Sport (2006). Rahmenlehrplan für Sekundarstufe I, Biologie. Verfügbar unter: http://www.berlin.de/imperia/md/content/sen-bildung/schulorganisation/lehrplaene/sek1_biologie.pdf [24.05.2008]
LP Bd Bio	Land Brandenburg, Ministerium für Bildung, Jugend und Sport (2002). <i>Rahmenlehrplan Biologie, Sekundarstufe I</i> . Verfügbar unter: http://www.bildungbrandenburg.de/fileadmin/bbs/unterricht_und_pruefungen/rahmenlehrplaene/sekundarstufe_I/rahmenlehrplaene/S1-Biologie.pdf [24.05.2008]
LP HH	Behörde für Bildung und Sport (2004). <i>Rahmenplan Naturwissenschaften / Technik für das achtstufige Gymnasium Sekundarstufe I</i> . Verfügbar unter: http://lbs.hh.schule.de/bildungsplaene/Sek-I_Gy9/NWT_Gy9_SekI.pdf [23.05.2008]
LP MV Bio	Ministerium für Bildung, Wissenschaft und Kultur (2005). <i>Rahmenplan für Gymnasium und integrierte Gesamtschule (Sekundarstufe I), Biologie</i> . Verfügbar unter: http://www.bildungsserver-mv.de/download/rahmenplaene/rp_biologie_7-10-gym-2005.pdf [24.05.2008]
LP MV Phy	Ministerium für Bildung, Wissenschaft und Kultur (2002). <i>Rahmenplan für Gymnasium und integrierte Gesamtschule (Sekundarstufe I), Physik</i> . Verfügbar unter: http://www.bildungsserver-mv.de/download/rahmenplaene/rp-physik-7-10-gym-02.pdf [24.05.2008]
LP Ni	Niedersächsisches Kultusministerium (2007). <i>Kerncurriculum für das Gymnasium Schuljahrgänge 5-10, Naturwissenschaften</i> . Verfügbar unter: http://www.nibis.de/nli1/gohrgs/kerncurricula_nibis/kc_2007/kc07_gym/kc_gym_nws_07_nib.pdf [23.05.2008]
LP RP	Ministerium für Bildung, Wissenschaft, Jugend und Kultur. <i>Lehrplanentwürfe, Lernbereich Naturwissenschaften: Biologie, Physik, Chemie</i> . Verfügbar unter: http://lehrplaene.bildung-rp.de/no-ache/schulart.html?tx_abdownloads_pi1%5Baction%5D=getviewcatalog&tx_abdownloads_pi1%5Bcategory_uid%5D=118&tx_abdownloads_pi1%5Bcid%5D=5785 [24.05.2008]
LP Saar Bio9	Ministerium für Bildung, Kultur und Wissenschaft (2004). <i>Lehrplan für das Fach Biologie der Klassenstufe 9</i> . Verfügbar unter: http://www.saarland.de/dokumente/thema_bildung/biologie9.pdf [23.05.2008]
LP Saar Chem8	Ministerium für Bildung, Kultur und Wissenschaft (2004). <i>Lehrplan für das Fach Chemie der Klassenstufe 8</i> . Verfügbar unter: http://www.saarland.de/dokumente/thema_bildung/chemie8.pdf [23.05.2008]
LP Saar ChemGr	Ministerium für Bildung, Kultur und Wissenschaft (2004). <i>Lehrplan Chemie für den Grundkurs</i> . Verfügbar unter: http://www.saarland.de/dokumente/thema_bildung/CHGKJuni2007V3.pdf [23.05.2008]

LP Sach Chem	Sächsisches Staatsministerium für Kultus (2004/2007). Lehrplan Gymnasium, Chemie. Verfügbar unter: http://www.sachsen-macht-schule.de/apps/lehrplandb/downloads/lehrplaene/lp_gy_chemie_2007.pdf [24.05.2008]
LP SA Bio	Kultusministerium des Landes Sachsen-Anhalt (2003). <i>Rahmenrichtlinien Gymnasium</i> Biologie für Schuljahrgänge 5-12. Verfügbar unter: http://www.rahmenrichtlinien.bildung-lsa.de/pdf/biogyma.pdf [23.05.2008]
LP SA Mathe	Kultusministerium des Landes Sachsen-Anhalt (2003). <i>Rahmenrichtlinien Gymnasium</i> Mathematik für Schuljahrgänge 5-12. Verfügbar unter: http://www.rahmenrichtlinien.bildung-lsa.de/pdf/mathegyma.pdf [23.05.2008]
LP SH Bio	Ministerium für Bildung, Wissenschaft, Forschung und Kultur des Landes Schleswig-Holstein (1997). <i>Lehrplan für die Sekundarstufe I der weiterführenden allgemeinbildenden Schulen Hauptschule, Realschule, Gymnasium, Gesamtschule – Biologie</i> . Verfügbar unter: http://lehrplan.lernnetz.de/intranet1/links/materials/1107160960.pdf [23.05.2008]
LP SH Mathe	Ministerium für Bildung, Wissenschaft, Forschung und Kultur des Landes Schleswig-Holstein (1997). <i>Lehrplan für die Sekundarstufe I der weiterführenden allgemeinbildenden Schulen Hauptschule, Realschule, Gymnasium, Gesamtschule – Mathematik</i> . Verfügbar unter: http://lehrplan.lernnetz.de/intranet1/links/materials/1107161435.pdf [23.05.2008]
LP SH Phy	Ministerium für Bildung, Wissenschaft, Forschung und Kultur des Landes Schleswig-Holstein (1997). <i>Lehrplan für die Sekundarstufe I der weiterführenden allgemeinbildenden Schulen Hauptschule, Realschule, Gymnasium, Gesamtschule – Physik</i> . Verfügbar unter: http://lehrplan.lernnetz.de/intranet1/links/materials/1107161553.pdf [23.05.2008]
LP Th Bio	Thüringer Kultusministerium (1999). <i>Lehrplan für das Gymnasium, Biologie</i> . Verfügbar unter: http://www.thillm.de/thillm/pdf/lehrplan/gy/gy_lp_bio.pdf [24.05.2008]
LP Th Chem	Thüringer Kultusministerium (1999). <i>Lehrplan für das Gymnasium, Chemie</i> . Verfügbar unter: http://www.thillm.de/thillm/start_service.html [24.05.2008]

11 Abkürzungsverzeichnis

AV	Abhängige Variable
bspw.	beispielsweise
bzw.	beziehungsweise
d. h.	das heißt
EAP/PV	expected a posteriori/plausible values
etc.	et cetera
ICC	Itemcharakteristik-Kurve (item characteristic curve)
i. e.	id est
o. g.	oben gegeben
s.	siehe
s. o.	siehe oben
u. a.	unter anderem
UV	Unabhängige Variable
vgl.	vergleiche
WLE	weighted likelihood estimator
WMNSQ	weighted mean square
z. B.	zum Beispiel

12Anhang

12.1 Tabellenverzeichnis

Tabelle 3.1: Strukturmodell der Diagrammkompetenz. Durch * gekennzeichnete Operationen beziehen sich nur auf Liniendiagramme.....	43
Tabelle 5.1: Stichprobengröße in der Pilotierung.....	70
Tabelle 5.2: Stichprobengröße in der Haupttestung. Die durch * gekennzeichneten Werte sind jeweils um eins höher als die Zeilensummen, da eine Person der neunten Klasse keine Angabe zum Geschlecht gemacht hat.....	70
Tabelle 5.3: Kodierleitfaden für die Achsenbeschriftung.....	83
Tabelle 5.4: Kriterien der Itemselektion.....	87
Tabelle 5.5: Übersicht über die zu berechnenden Regressionen und zu prüfenden Voraussetzungen für eine Mediationsanalyse. * = signifikanter Koeffizient, $p < .05$ mindestens.....	101
Tabelle 6.1: Kennwerte der Modellpassung. AIC_c = besondere Form des Akaike's Information Criterion, BIC: Best Information Criterion, L = Likelihoodwert.....	108
Tabelle 6.2: Reliabilitäten der Komponenten des Modells <i>3dim</i> (a) und der Komponenten des Modells <i>5dim</i> (b). $N=289$	109
Tabelle 6.3: Korrelationen (Pearsons r) der Testungen auf Personenhomogenität bezüglich den (Unter-)Komponenten der Modelle <i>3dim</i> (a) und <i>5dim</i> (b). Z_1 bis Z_3 = zufällige Teilungen der Stichprobe (bei den Berechnungen der Modelle <i>3dim</i> und <i>5dim</i> wurden verschiedene Teilungen vorgenommen), K = Teilung nach Klassenstufe, G = Teilung nach Geschlecht. ** $p < .01$	110
Tabelle 6.4: Kennwerte der schrittweisen Regression der Personenparameter des <i>3dim</i> -Modells auf den Personenparameter des <i>1dim</i> -Modells. *** $p < .001$	111
Tabelle 6.5: Kennwerte der schrittweisen Regression der Personenparameter des <i>5dim</i> -Modells auf den Personenparameter des <i>1dim</i> -Modells. ** $p < .001$	111
Tabelle 6.6: Kennwerte der Modellpassung. AIC_c = besondere Form des Akaike's Information Criterion, BIC: Best Information Criterion, L = Likelihood-Wert.....	113
Tabelle 6.7: Reliabilitäten der Rasch-Dimensionen in den Modellen <i>2dimDTx3dim</i> und <i>2dimDTx5dim</i> ($N = 289$).....	114
Tabelle 6.8: Korrelationen (Pearsons r) der Komponenten innerhalb des Modells <i>3dim</i> . ** $p < .01$	115
Tabelle 6.9: Korrelationen (Pearsons r) der Komponenten innerhalb des Modells <i>5dim</i> . ** $p < .01$	116
Tabelle 6.10: Kennwerte der schrittweisen Regression mit Dateneintrag als abhängiger Variable. *** $p < .001$	117
Tabelle 6.11: Kennwerte der schrittweisen Regression mit Aufbau als abhängiger Variable. ** $p < .01$, *** $p < .001$	117
Tabelle 6.12: Kennwerte der schrittweisen Regression mit Dateneintrag als abhängiger Variable. *** $p < .001$	118
Tabelle 6.13: Kennwerte der Regression mit Integration als abhängiger Variable. *** $p < .001$	118
Tabelle 6.14: Übersicht über die Berechnungen und Prüfungen der Voraussetzungen für eine Mediation des Zusammenhangs zwischen Konstruktion und Integration durch die Informationsentnahme. K = Konstruktion, IE = Informationsentnahme, Int = Integration. Zum Abgleich mit der methodischen Beschreibung (vgl. 5.3.3.2,	

Relationen der Komponenten) wurden die Benennungen der Regressionskoeffizienten beibehalten. ** $p < .01$, *** $p < .001$.	120
Tabelle 6.15: Kennwerte der schrittweisen Regression für die abhängige Variable Integration. * $p < .05$ *** $p < .001$.	121
Tabelle 6.16: Übersicht über die Berechnungen und Prüfungen der Voraussetzungen für eine Mediation des Zusammenhangs zwischen Identifizierung und Integration durch das Ablesen. Id = Identifizierung, A = Ablesen, Int = Integration. Zum Abgleich mit der methodischen Beschreibung (vgl. 5.3.3.2, Relationen der Komponenten) wurden die Benennungen der Regressionskoeffizienten beibehalten. * $p < .05$, *** $p < .001$.	122
Tabelle 6.17: Übersicht über die Berechnungen und Prüfungen der Voraussetzungen für eine Mediation des Zusammenhangs zwischen Aufbau und Integration durch das Ablesen. Auf = Aufbau, A = Ablesen, Int = Integration. Zum Abgleich mit der methodischen Beschreibung (vgl. 5.3.3.2, Relationen der Komponenten) wurden die Benennungen der Regressionskoeffizienten beibehalten. * $p < .05$, *** $p < .001$.	123
Tabelle 6.18: Übersicht über die Berechnungen und Prüfungen der Voraussetzungen für eine Mediation des Zusammenhangs zwischen Dateneintrag und Integration durch das Ablesen. Dat = Dateneintrag, A = Ablesen, Int = Integration. Zum Abgleich mit der methodischen Beschreibung (vgl. 5.3.3.2, Relationen der Komponenten) wurden die Benennungen der Regressionskoeffizienten beibehalten. *** $p < .001$.	124
Tabelle 6.19: Deskriptive Statistik der Itemschwierigkeiten für die Komponenten des 3dim-Modells.	125
Tabelle 6.20: Deskriptive Statistik über die Itemschwierigkeiten für die Komponenten des 5dim-Modells.	126
Tabelle 6.21: Kennwerte der Modellpassung. AIC_c = besondere Form des Akaike's Information Criterion, BIC: Best Information Criterion, L = Likelihoodwert.	128
Tabelle 6.22: Deskriptive Statistiken zu den Itemschwierigkeiten der vier Ordnungen des Ablesens.	129
Tabelle 6.23: Kennwerte der Modellpassung der drei Modelle zur Komponente Konstruktion. AIC_c = besondere Form des Akaike's Information Criterion, BIC: Best Information Criterion, L = Likelihoodwert.	130
Tabelle 6.24: Kennwerte der Modellpassung der zwei Modelle zur Komponente Konstruktion. AIC_c = besondere Form des Akaike's Information Criterion, BIC: Best Information Criterion, L = Likelihoodwert.	132
Tabelle 6.25: Deskriptive Statistik der Personenfähigkeiten in den Komponenten des Modells 3dim getrennt nach Klassenstufen.	135
Tabelle 6.26: Deskriptive Statistik der Personenfähigkeiten in den Komponenten des Modells 5dim getrennt nach Klassenstufen.	135
Tabelle 6.27: Ansatz S (mit Summenscores für das biologische Vorwissen (VW_Bio) und die Fähigkeiten zum Umgang mit abstrakten Diagrammen (UAD)) der Regressionen zum Einfluss der Lernermerkmale. * $p < .05$, ** $p < .01$, *** $p < .001$.	138
Tabelle 6.28: Ansatz E (mit Einzelitems für das biologische Vorwissen (VW_Bio) und die Fähigkeiten zum Umgang mit abstrakten Diagrammen (UAD)) der Regressionen zum Einfluss der Lernermerkmale. * $p < .05$, ** $p < .01$, *** $p < .001$.	138
Tabelle 6.29: Deskriptive Statistik für die Auslassungen von Items zu den Komponenten des 3dim-Modells.	140
Tabelle 12.1: Verteilung sämtlicher Items auf die Testhefte der Pilotierung. Items zur Diagrammkompetenz: Hellgrau unterlegt die Items zur Informationsentnahme, weiß unterlegt die Items zur Konstruktion, dunkelgrau unterlegt die Items zur Integration.	188
Tabelle 12.2: Itemkennwerte der Informationsentnahme.	196

Tabelle 12.3: Itemkennwerte der Konstruktion.....	197
Tabelle 12.4: Itemkennwerte der Integration.....	198
Tabelle 12.5: Itemkennwerte der Items zu den Lernermerkmalen.....	199
Tabelle 12.6: Itemkennwerte der Items zur Diagrammkompetenz im Haupttest.....	201
Tabelle 12.7: Regressionsanalyse zur Bestimmung der Effektstärke des Zusammenhang von Informationsentnahme und Konstruktion.....	203
Tabelle 12.8: Regressionsanalysen zur Mediationswirkung der Informationsentnahme auf den Zusammenhang zwischen Konstruktion und Integration.....	203
Tabelle 12.9: Regressionsanalysen zur Mediationswirkung des Ablesens auf den Zusammenhang zwischen Identifizierung und Integration.....	204
Tabelle 12.10: Regressionsanalysen zur Mediationswirkung des Ablesens auf den Zusammenhang zwischen Aufbau und Integration.....	204
Tabelle 12.11: Regressionsanalysen zur Mediationswirkung des Ablesens auf den Zusammenhang zwischen Dateneintrag und Integration.....	205
Tabelle 12.12: Modellvergleiche der Gesamtmodelle, die jeweils die Aspekte einer Komponente differenzieren.....	205
Tabelle 12.13: Deskriptive Statistik zu den Itemschwierigkeiten in den einzelnen Aspekten der Konstruktion.....	205
Tabelle 12.14: Deskriptive Statistik der Itemschwierigkeiten hinsichtlich der zwei Aspekte der Integration.....	206
Tabelle 12.15: Deskriptive Statistik der Itemschwierigkeiten in den einzelnen Komponenten des <i>3dim</i> -Modells nach Diagrammtypen differenziert. LD: Liniendiagramm, SD: Säulendiagramm.....	206
Tabelle 12.16: Deskriptive Statistik der Itemschwierigkeiten in den einzelnen Komponenten des <i>5dim</i> -Modells nach Diagrammtypen differenziert. LD: Liniendiagramm, SD: Säulendiagramm.....	206
Tabelle 12.17: Deskriptive Statistik der Personenfähigkeiten in den Komponenten des Modells <i>3dim</i> getrennt nach Geschlecht. m = Schüler, w = Schülerinnen.....	207
Tabelle 12.18: Deskriptive Statistik der Personenfähigkeiten in den Komponenten des Modells <i>5dim</i> getrennt nach Klassenstufen. m = Schüler, w = Schülerinnen.....	207
Tabelle 12.19: Referenzzahlen für die Häufigkeiten der auftretenden Fehler bei der Konstruktion.....	207

12.2 Abbildungsverzeichnis

Abbildung 3.1: Beispiele für Blütendiagramm (a) und Baumdiagramm (b). (a aus Vogel & Angermann, 1992, S. 558; b aus Bickel et al., 1995, S. 125).....	9
Abbildung 3.2: Kategorisierung der Diagrammtypen mit Bezug zu realistischen Bildern und Texten.....	12
Abbildung 3.3: Liniendiagramm mit zwei abhängigen Variablen, die auf zwei verschiedenen Achsen eingetragen sind.....	17
Abbildung 3.4: Liniendiagramm zur Populationsentwicklung von Schildläusen und Marienkäfern in einem Räuber-Beute-System.....	34
Abbildung 3.5: Beispiel für ein Item zur Achsenbelegung (adaptiert nach Padilla et al., 1986).....	37
Abbildung 5.1: Itemcharakteristik-Kurve (<i>item characteristic curve, ICC</i>) für ein Item <i>i</i> ; monoton steigend.....	60
Abbildung 5.2: ICCs dreier Items mit unterschiedlicher Schwierigkeit. Je weiter rechts, desto schwieriger das Item.....	61

Abbildung 5.3: ICCs dreier Items unterschiedlicher Trennschärfe. Je steiler die Kurve, desto trennschärfer das Item.	62
Abbildung 5.4: Kategorienfunktionen eines ordinalen Items mit drei Kategorien.	65
Abbildung 5.5: Schematische Darstellung der zwei mehrdimensionalen Rasch-Modelle (verändert nach Adams et al., 1997). Komponenten/Dimensionen 1 bis 3. a) between-item-Modell: jedes Item ist genau einer Fähigkeitskomponente/Modelldimension zugeordnet. b) within-item-Modell: jedes Item ist einer oder mehreren Fähigkeitskomponenten/Modelldimensionen zugeordnet.	67
Abbildung 5.6: Diagramm <i>Maus</i> mit exemplarischem Item zur Identifizierung, welches das Erkennen der dargestellten Relation abfragt.	72
Abbildung 5.7: Diagramm <i>Rentiere</i> (im Testbogen ohne eingezeichnete Datenpunkte sowie Interpolationslinie).	73
Abbildung 5.8: Items zum Diagramm <i>Papageitaucher</i> , die verschiedene Ordnungen des Ablesens repräsentieren.	73
Abbildung 5.9: Diagramm <i>Eier</i> mit Aufgabenstellung zur Konstruktion respektive Integration.	75
Abbildung 5.10: Aufgabenstellung zur Konstruktion des Diagramms <i>Kiefernspinner</i> . Zum Zeichnen wurde Millimeterpapier bereitgestellt.	75
Abbildung 5.11: Diagramm <i>Papageitaucher</i> mit zusätzlichem Text und zwei Items zur Integration.	78
Abbildung 5.12: Beispiel für ein Item zum biologischen Vorwissen.	79
Abbildung 5.13: Beispiel für ein Item zum Umgang mit abstrakten Diagrammen.	80
Abbildung 5.14: Auszug aus der Itemstatistik für ein Single-Choice-Item zur Identifizierung. Die Spalte <i>Label</i> enthält die Rubriken 1 bis 4, Spalte <i>Score</i> deren Punktzahl. Für jede Rubrik sind in der Spalte <i>Pt Bis</i> deren punkt-biseriale Korrelation sowie die zugehörige <i>t</i> -Statistik mit Signifikanzwert <i>p</i> aufgeführt.	86
Abbildung 5.15: Formeln für die Indices zum Modellvergleich. L = Likelihood, n_p = Zahl der Modellparameter (vgl. 5.1.2.4), N = Stichprobengröße.	95
Abbildung 5.16: Beispiel für den graphischen Modelltest auf Personenhomogenität. Die Stichproben 2A und 2B bezeichnen zwei Teilstichproben einer zufälligen Zweiteilung einer fiktiven Gesamtstichprobe. Im dargestellten Fall ist Personenhomogenität gegeben.	96
Abbildung 5.17: Schematische Darstellung des within-item-Modells $2dimDTx3dim$, das drei strukturelle Komponenten Informationsentnahme (IE), Konstruktion (K) und Integration (Int) der Diagrammkompetenz sowie die Diagrammtypen Liniendiagramm (LD) und Säulendiagramm (SD) unterscheidet. Ein Item wird jeweils einem Diagrammtyp und einer Kompetenzkomponente zugeordnet.	98
Abbildung 5.18: Zusammenhänge zwischen unabhängiger (X), mediierender (M) und abhängiger Variable (Y). Zu berechnende Regressionen mit zugehörigen Koeffizienten in drei unterschiedlichen Farben.	100
Abbildung 5.19: Beispiel für ein Boxplot.	103
Abbildung 6.1: Schaubild zu den berechneten Zusammenhängen zwischen den Komponenten des $3dim$ -Modells. IE = Informationsentnahme, K = Konstruktion, Int = Integration. Zum Abgleich mit der methodischen Beschreibung (vgl. 5.3.3.2, Relationen der Komponenten) wurden die Benennungen der Regressionskoeffizienten beibehalten.	119
Abbildung 6.2: Mittlere Itemschwierigkeiten der Komponenten im $3dim$ -Modell.	126
Abbildung 6.3: Mittlere Itemschwierigkeiten der Komponenten im $5dim$ -Modell.	127
Abbildung 6.4: Boxplot zu den Schwierigkeiten der vier Ordnungen des Ablesens.	129

Abbildung 6.5: Boxplot zu den Schwierigkeiten der einzelnen Konstruktionsaspekte. Typ: Auswahl des Diagrammtyps, Achsenbel.: Achsenbelegung, Beschr.: Beschriftung, Skal.: Skalierung, Werte: Werteintrag, Skizz.: Skizzierung.....	131
Abbildung 6.6: Schwierigkeiten der einzelnen Items zur Integration. Die Items <i>I_P2c</i> , <i>IS_T_a</i> , <i>IS_T_b</i> und <i>IS_T_e</i> bezeichnen Items, die dem Integrationsaspekt <i>Resultat Text</i> zugeordnet sind. Die Items <i>I_Ke_W</i> , <i>I_Ke_x</i> und <i>IS_T_f</i> sind Items zum Aspekt <i>Resultat Diagramm</i>	132
Abbildung 6.7: Schwierigkeiten der Items zu verschiedenen Diagrammtypen aufgeschlüsselt nach Komponenten des <i>3dim</i> -Modells. IE: Informationsentnahme, K: Konstruktion, Int: Integration. Zur einfacheren Darstellung sind die Personenparameter um 10 nach oben verschoben.....	134
Abbildung 6.8: Vergleich der Personenfähigkeiten nach Klassenstufen für die Komponenten des <i>3dim</i> -Modells (a) und die Komponenten des <i>5dim</i> -Modells (b). Zur einfacheren Darstellung sind die Personenparameter um 10 nach oben verschoben.....	136
Abbildung 6.9: Beispiel für die fälschliche Auswahl eines Liniendiagramms.	141
Abbildung 6.10: Beispiel für eine fehlerhafte Achsenbelegung.	142
Abbildung 6.11: Beispiel für Achsenbeschriftungen nur durch die Angabe der Einheiten der Variablen.	143
Abbildung 6.12: Beispiel für die Erstellung einer Skala, die nur die im Text gegebenen Werte auflistet. Aufgabe war es, die Höhe einer Pflanze über eine bestimmte Beobachtungszeit darzustellen.	144
Abbildung 6.13: Beispiel für eine Antwort zum Werteintrag im Diagramm <i>Rentiere</i>	145

12.3 Testinstrumente

12.3.1 Übersicht über die Items in den Testheften der Pilotierung

Tabelle 12.1: Verteilung sämtlicher Items auf die Testhefte der Pilotierung. Items zur Diagrammkompetenz: Hellgrau unterlegt die Items zur Informationsentnahme, weiß unterlegt die Items zur Konstruktion, dunkelgrau unterlegt die Items zur Integration.

Testheft 1		Testheft 2		Testheft 3		Testheft 4		Testheft 5		Testheft 6	
Diagramm	Item	Diagramm	Item	Diagramm	Item	Diagramm	Item	Diagramm	Item	Diagramm	Item
Kaninchen	Id_K_a	Kaninchen	Id_K_a	Kaninchen	Id_K_a	Kaninchen	Id_K_a	Kaninchen	Id_K_a	Kaninchen	Id_K_a
	Id_K_b		Id_K_b		Id_K_b		Id_K_b		Id_K_b		Id_K_b
	Id_K_c		Id_K_c		Id_K_c		Id_K_c		Id_K_c		Id_K_c
Rentiere	Id_Re	Maus	Id_M_a	Temperatur	Id_T_a	Rentiere	Id_Re	Rentiere	Id_Re	Rentiere	Id_Re
Gewebe	IdS_Ga		Id_M_b		Id_T_b	Körpertemperatur	IdS_Ca	Hund	A_Hu_a	Läuse	A_La_a
	IdS_Gb		Id_M_c		Id_T_c		IdS_Cb		A_Hu_b		A_La_c
Läuse	A_La_a		Id_M_d		Id_T_d		Asseln		Id_As		Läuse
	A_La_c	Rentiere	Id_Re	Rentiere	Id_Re	A_La_c		A_La_c	Rentiere	K_Re_W	
	A_La_e	Läuse	A_La_b	Läuse	A_La_a	A_La_l	A_La_l	A_La_l	K_Re_X		
	A_La_h		A_La_c		Läusekäfer	A_LK_b	Paramecium2	A_P2a1	Lernermerkmal	Item	
	A_La_i		A_La_d			A_LK_d		A_P2a2	biologisches Vorwissen	rb1	
	A_La_l		A_La_f	A_LK_h	A_P2a3	pop1					
	A_La_m		A_La_g	A_LK_i	A_P2a4	ps1					
Hund	A_Hu_a	A_La_k	Läusekäfer	A_LK_e	Population	A_Po_a	af1				
	A_Hu_b	A_La_l		A_LK_f		A_Po_b	konk1				
	A_Hu_c	A_La_n		A_LK_g	RäuberBeute	A_RB_c	A_Po_c	af2bsp			
Fischherz	K_FH_A	Paramecium2	A_P2b1	A_LK_k	Früchte	Au_Fr	Früchte	Au_Fr	ps2		
	K_FH_B										

Fortsetzung der Tabelle auf der folgenden Seite.

Testheft 1		Testheft 2		Testheft 3		Testheft 4		Testheft 5		Testheft 6				
Diagramm	Item	Diagramm	Item	Diagramm	Item	Diagramm	Item	Diagramm	Item	Diagramm	Item			
Kiefernspinner	K_Ki_T	Paramecium2	A_P2b5	Papageitaucher	AS_P_a	Früchte	K_Pr_BC	Früchte	K_Re_X	biologisches Vorwissen	rb2			
	K_Ki_A		A_P2c1		bis		K_Pr_L	Schnecke	K_Sc_L		konk2			
	K_Ki_BC		A_P2c2		AS_P_h		K_Pr_S	Blüten- durchmesser	KS_B_A	Lesekompetenz	LKa			
	K_Ki_S		A_P2c3	K_As_W	K_Pr_T				KS_B_BC		LKb			
	K_Ki_W	A_P2c4	K_As_X	K_Pr_W	KS_B_S		Re1							
	K_Ki_X	AS_Z_a	K_Re_W	K_Pr_X	KS_B_T		Re2							
Rentiere	K_Re_W	Artenzahlen	bis	Rentiere	K_Re_X	K_PW_A	Pflanzenwuchs		Eier		KS_B_W	Rechen- fähigkeiten	Re3	
	K_Re_X		AS_Z_h	Transpiration	K_Tr_S	K_PW_BC				KS_B_X	Re4			
Ozon	KS_O_B	Kröte	Au_Kr	Winterschlaf	KS_W_A	K_PW_S		Rentiere		Flusskrebse	KS_F_B		Umgang mit abstrakten Diagrammen	Re5
	KS_O_S	Aal	K_Aa_A		KS_W_BC	K_PW_T					K_Re_W			Paramecium1
Paramecium2	I_P2b		Rentiere		K_Aa_BC	KS_W_S		KS_W_T		K_Re_X	Paramecium2			I_P1_S
	I_P2c	K_Re_W			KS_W_W		Kälte		I_Ke_W	Paramecium3	I_P2a	uad13		
Gleichwarm	I_GW_L	Geschlechts- reife	K_Re_X					KS_W_X	I_Ke_X		I_P3_L	I_P3_W		uad_au
	I_GW_W		KS_G_A		Bakterien	I_Bak	I_RB_a	I_P3_X		S_uad1				
	I_GW_X		KS_G_BC	Photosynthese	I_PS_a	RäuberBeute			I_RB_b	Eier	IS_F_S	bis		
Ozon	IS_O_L		KS_G_S		I_PS_b		I_RB_c	S_uad4						
	Papageitaucher	KS_G_T	I_PS_c	I_RB_d										
		IS_T_a	Papageitaucher	Photosynthese	I_RB_e									
		IS_T_b												
		IS_T_c												
		IS_T_d												
IS_T_e														
IS_T_f														

12.3.2 Items zu den Lernermerkmalen

12.3.2.1 Biologisches Vorwissen

Item *pop1*: Eine Population ist...

- ...eine Gruppe von Tieren einer Art, die alle von einem Vorfahr abstammen.
- ...eine Gruppe von Tieren einer Art, die alle gleich alt sind.
- ...eine Gruppe von Tieren einer Art, die in einem definierten Gebiet leben.
- ...eine Gruppe von Tieren einer Art, die alle gleich groß sind.

Item *pop2*: Die Populationsgröße ist...

- ...die Anzahl der Nachkommen der Population.
- ...die durchschnittliche Größe der Tiere in der Population.
- ...die Anzahl der Populationen in einem Lebensraum.
- ...die Anzahl der Tiere, die zu der Population gehören.

Item *rb2*: Das erste Gesetz von Lotka-Volterra besagt, ...

- ...dass sich die Individuenzahlen von Räuber und Beute über einen längeren Zeitraum hinweg annähern.
- ...dass die Individuenzahlen von Räuber und Beute periodisch und phasenverschoben schwanken.
- ...dass die Individuenzahlen von Räuber und Beute sich nach einem bestimmten Zeitraum auf einen konstanten Wert einstellen.
- ...dass die Individuenzahlen von Räuber und Beute sich gegenseitig nicht beeinflussen.

Item *konkl*: Eine Blaumeise sucht ihre Nahrung auf den höheren, äußeren Zweigen eines Baumes, eine Kohlmeise sucht eher auf den unteren Zweigen, die nah am Stamm sind. Welche Aussage trifft zu?

- Die beiden Meisenarten können zusammen auf einem Baum leben, weil sie verschiedene Nahrung bevorzugen.
- Die beiden Meisenarten können nicht zusammen auf einem Baum leben, weil die Zweige sie nicht halten könnten.
- Die beiden Meisenarten können nicht zusammen auf einem Baum leben, weil sie die gleiche Nahrung suchen und um die Nahrung konkurrieren.
- Die beiden Meisenarten können zusammen auf einem Baum leben, weil sie verschiedene Plätze für die Nahrungssuche wählen.

12.3.2.2 Lesekompetenz

Das Item zur Lesekompetenz ist den Aufgabenbeispielen zur PISA 2000-Studie entnommen (OECD, 2006), wobei der Text gekürzt wurde.

Lies bitte den folgenden Informationstext für die Mitarbeiter der Firma ACOL und beantworte danach die Frage!

Programm zur freiwilligen Gripeschutzimpfung bei ACOL

Wie Sie sicherlich wissen, kann die Grippe im Winter schnell und weiträumig um sich greifen. Ihre Opfer liegen dann oft wochenlang krank im Bett.

Der beste Weg, das Virus zu bekämpfen, ist ein fitter und gesunder Körper. Tägliche Bewegung und eine Ernährung mit viel Obst und Gemüse sind sehr zu empfehlen, um das Immunsystem in seinem Kampf gegen diesen Krankheitserreger zu unterstützen.

Die Firma ACOL hat beschlossen, ihren Mitarbeitern die Gelegenheit zur Gripeschutzimpfung als einen zusätzlichen Weg anzubieten, um dieses tückische Virus an der Ausbreitung unter uns zu hindern. ACOL hat eine Krankenschwester engagiert, die in der Woche ab dem 17. November einen halben Tag lang während der Arbeitszeit die Impfungen in den Räumen der Firma ACOL vornimmt. Dieses Programm ist kostenlos und steht allen Mitarbeitern zur Verfügung.

Nach medizinischen Erkenntnissen verursacht die Impfung keine Grippeinfektion. Manchmal kommt es jedoch zu einigen Nebenwirkungen wie Müdigkeit, leichtem Fieber und Schmerzen im Arm.

Item *LKb*: Im Informationsblatt heißt es, wenn man sich gegen das Grippevirus schützen wolle, sei eine Gripeschutzimpfung...

- ...wirksamer als Bewegung und gesunde Ernährung, aber riskanter.
- ...eine gute Sache, aber kein Ersatz für Bewegung und gesunde Ernährung.
- ...genauso wirksam wie Bewegung und gesunde Ernährung und weniger anstrengend.
- ...nicht nötig, wenn man sich viel bewegt und gesund ernährt.

12.3.2.3 Rechenfähigkeiten

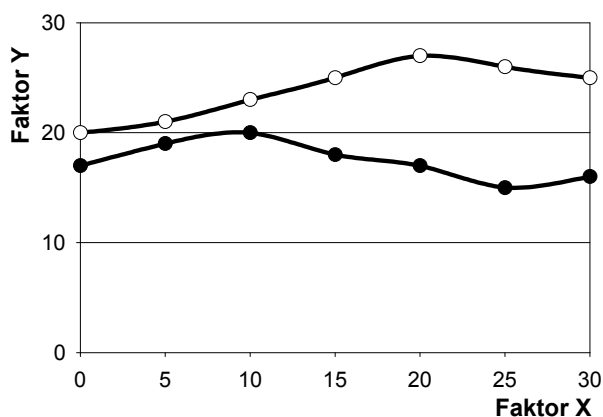
Item *Re4*: $16\ 000 + 87\ 000 - 10\ 300 =$ _____

12.3.2.4 Fähigkeiten zum Umgang mit abstrakten Diagrammen

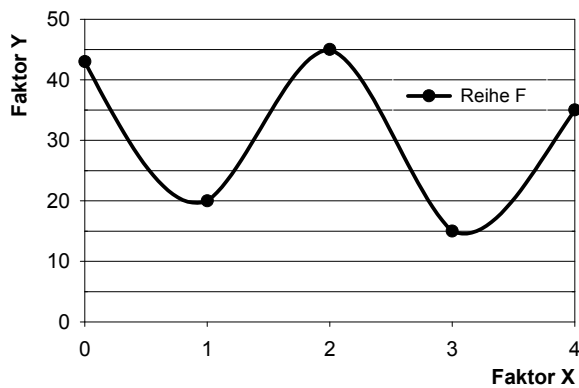
Item *uad10*: Die abhängige Größe in einem Diagramm ist...

- ...der Faktor, der auf der Y-Achse eingetragen wird.
- ...die Beschriftung der X-Achse.
- ...der Faktor, der konstant gehalten wird.
- ...die Beschriftung der Linie.

Item *uad11*: Falls du der Meinung bist, das folgende Diagramm sei nicht vollständig, vervollständige es durch beliebige Angaben!

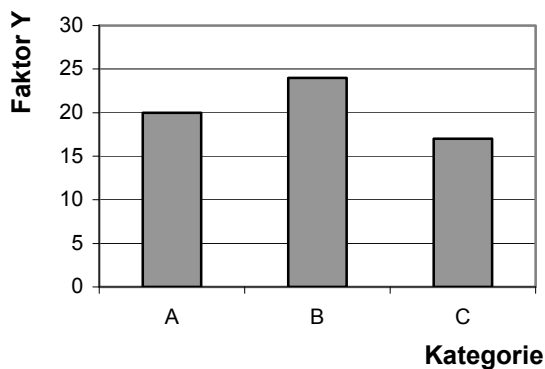


Item *uad6*: Welche der Aussagen trifft zu?



- Reihe F nimmt bei 2 ihr Maximum an und der Y-Wert beträgt an dieser Stelle 43.
- Reihe F nimmt bei 0 ihr Maximum an und der Y-Wert beträgt an dieser Stelle 30.
- Reihe F nimmt bei 4 ihr Maximum an und der Y-Wert beträgt an dieser Stelle 35.
- Reihe F nimmt bei 2 ihr Maximum an und der Y-Wert beträgt an dieser Stelle 45.

Item *S_uad3*: Vergleiche für das folgende Diagramm den Unterschied zwischen A und B mit dem zwischen B und C!



- Der Unterschied zwischen A und B beträgt 4 und ist größer als der zwischen B und C.
- Der Unterschied zwischen A und B ist gleich groß wie der zwischen B und C.
- Der Unterschied zwischen B und C beträgt 4 und ist größer als der zwischen A und B.
- Der Unterschied zwischen B und C beträgt 7 und ist größer als der zwischen A und B.

12.3.2.5 Anstrengung

Das Item zur Selbsteinschätzung der Anstrengung in Anlehnung an das „Anstrengungsthermometer 2“ (PISA-Konsortium Deutschland, 2006a, S. 125f) formuliert.

Noch eine letzte Frage...

Stell dir bitte eine konkrete Situation (innerhalb oder außerhalb der Schule) vor, die für dich persönlich sehr wichtig ist und in der du dein Bestes geben und dich maximal anstrengen würdest. In dieser Situation würdest du dann auf einer Skala von 1 bis 10 den Wert 10 ankreuzen.

Wie sehr hast du dich im Vergleich zu der gerade vorgestellten Situation bei diesem Test angestrengt? Gib bitte einen Wert von 1 (für gar nicht angestrengt) bis 10 (für maximal angestrengt) an!

12.4 Itemkennwerte

12.4.1 Pilotierung: Items zur Diagrammkompetenz

12.4.1.1 Liste der von ConQuest ausgeschlossenen Items

IdS_Ca, A_La_a, A_La_b, A_La_c, A_La_d, A_La_f, A_La_h, A_La_k, A_La_n, A_LK_a, A_LK_c, A_LK_d, A_LK_f, A_Po_a, A_Po_b, A_P2a3, A_P2a4, A_P2b2, A_P2b3, A_P2b4, A_P2b5, A_P2c3, A_P2c4, A_RB_c, A_RB_d, AS_P_a, AS_P_b, AS_P_c, AS_P_d, AS_Z_a, AS_Z_b, AS_Z_c, AS_Z_d, AS_Z_e, AS_Z_f, AS_Z_g, AS_Z_h, Au_Kr, IS_F_S.

12.4.1.2 Statistische Kennwerte der Items

In den nachfolgenden Tabellen sind die statistischen Kennwerte sämtlicher Items zur Diagrammkompetenz aufgeführt. Sie sind nach Komponenten geordnet.

Itemwerte, die außerhalb der aufgeführten Grenzen liegen, sind durch * markiert. Hinsichtlich der Ausschlusskriterien bezüglich WMNSQ, t-Wert, Schwierigkeit, PMK und PBK₀ führen sie direkt zum Ausschluss des Items, welches in der Spalte „selektiert“ nicht auftritt. Hinsichtlich der zweitgenannten Beachtungskriterien führen sie dazu, das Item als kritisch einzustufen. Dies wird in der Spalte „selektiert“ durch * angezeigt. Items dieser Spalte, die schließlich für den Haupttest ausgewählt wurden, werden in schwarzer Schrift dargestellt, Items, die ihren Kennwerten nach hätten ausgewählt werden können, aber nicht ausgewählt wurden, in grauer Schrift. Items, die trotz eines erfüllten Ausschlusskriteriums in den Haupttest eingingen, werden mit ** markiert.

Tabelle 12.2: Itemkennwerte der Informationsentnahme.

Diagramm	Item	WMNSQ	t	Schwierigkeit	PMK	PBK ₀	selektiert
Kaninchen	<i>Id_K_a</i>	1.06	0.5	-1.541	0.26	-0.20	<i>Id_K_a</i>
	<i>Id_K_b</i>	0.96	0.0	-3.359*	0.42	-0.31	
	<i>Id_K_c</i>	0.98	-0.1	-1.541	0.45	-0.19	<i>Id_K_c</i>
Maus	<i>Id_M_a</i>	0.98	0.1	-1.375	0.39	/	<i>Id_M_a</i>
	<i>Id_M_b</i>	1.07	0.3	-2.480	0.14*	/	** <i>Id_M_b</i>
	<i>Id_M_c</i>	0.86	-1.2	0.562	0.53	/	<i>Id_M_c</i>
	<i>Id_M_d</i>	1.16	0.6	-1.375	0.32	-0.56	<i>Id_M_d</i>
Temperatur	<i>Id_T_a</i>	1.06	0.3	-2.154	0.13*	/	
	<i>Id_T_b</i>	1.11	0.4	-1.806	0.19*	/	
	<i>Id_T_c</i>	1.04	0.3	-0.851	0.42	-0.46	<i>Id_T_c</i>
	<i>Id_T_d</i>	0.83	-1.5	-0.306	0.63	-0.58	<i>Id_T_d</i>
Rentiere	<i>Id_Re</i>	0.96	-0.9	-0.174	0.34	-0.17	<i>Id_Re</i>
Asseln	<i>Id_As</i>	1.08	0.7	-0.306	0.21*	/	
Körpertemperatur	<i>IdS_Cb</i>	1.00	0.1	-1.104	0.39	/	<i>IdS_Cb</i>
	<i>IdS_Cc</i>	1.13	0.6	-1.322	0.20*	/	
Gewebe	<i>IdS_Ga</i>	1.05	0.3	-0.912	0.21*	/	** <i>IdS_Ga</i>
	<i>IdS_Gb</i>	0.98	0.0	-1.116	0.64	-0.55	<i>IdS_Gb</i>
Läuse	<i>A_La_e</i>	0.94	-0.3	-0.912	0.35	/	<i>A_La_e</i>
	<i>A_La_g</i>	1.16	0.5	-2.013	-0.10*	/	
	<i>A_La_i</i>	0.84	-0.8	-0.912	0.59	-0.42	<i>A_La_i</i>
	<i>A_La_l</i>	0.98	-0.3	0.653	0.48	-0.16	<i>A_La_l</i>
	<i>A_La_m</i>	0.99	0.0	-0.538	0.39	-0.42	<i>A_La_m</i>
Läusekäfer	<i>A_LK_b</i>	1.06	0.6	-0.196	0.44	-0.45	<i>A_LK_b</i>
	<i>A_LK_e</i>	1.09	0.4	-1.806	0.23*	-0.46	
	<i>A_LK_g</i>	1.00	0.1	-0.135	0.31	-0.44	<i>A_LK_g</i>
	<i>A_LK_h</i>	1.17	0.5	-2.194	0.17*	/	** <i>A_LK_h</i>
	<i>A_LK_i</i>	0.94	-0.5	-0.365	0.58	-0.19	<i>A_LK_i</i>
	<i>A_LK_k</i>	0.97	-0.2	0.374	0.52	-0.60	<i>A_LK_k</i>
	<i>A_LK_l</i>	0.94	-0.5	0.137	0.69	-0.51	<i>A_LK_l</i>
<i>A_LK_m</i>	1.00	0.1	-0.135	0.34	-0.44	<i>A_LK_m</i>	
Population	<i>A_Po_c</i>	1.03	0.2	-2.149	0.35	-0.66	<i>A_Po_c</i>
Hund	<i>A_Hu_a</i>	1.11	1.2	0.111	0.23*	/	** <i>A_Hu_a</i>
	<i>A_Hu_b</i>	1.07	0.4	1.472	0.39	/	<i>A_Hu_b</i>
	<i>A_Hu_c</i>	0.96	-0.4	0.308	0.48	/	<i>A_Hu_c</i>
Paramecium2	<i>A_P2a1</i>	0.86	-0.4	-1.503	0.57	-0.59	<i>A_P2a1</i>
	<i>A_P2a2</i>	0.77	-1.2	-1.023	0.52	/	<i>A_P2a2</i>
	<i>A_P2b1</i>	0.99	0.0	-1.116	0.40	-0.40	<i>A_P2b1</i>
	<i>A_P2c1</i>	0.91	-0.5	-0.720	0.51	-0.50	<i>A_P2c1</i>
	<i>A_P2c2</i>	0.87	-1.0	-0.362	0.61	-0.57	<i>A_P2c2</i>
Papageitaucher	<i>AS_P_e</i>	0.86	-0.2	-2.154	0.28	/	<i>AS_P_e</i>
	<i>AS_P_f</i>	0.84	-0.1	-2.618	0.24*	/	
	<i>AS_P_g</i>	0.85	-1.1	-0.480	0.72	-0.44	<i>AS_P_g</i>
	<i>AS_P_h</i>	0.98	-0.1	-0.851	0.32	/	<i>AS_P_h</i>

WMNSQ = weighted mean square, t = Wert des zugehörigen t-Tests, Schwierigkeit = Itemparameter, PMK = Produkt-Moment-Korrelation (Trennschärfe), PBK₀ = höchste signifikante punkt-biseriale Korrelation einer 0-Rubrik. Erklärungen der Markierungen durch * im Text.

Tabelle 12.3: Itemkennwerte der Konstruktion.

Diagramm	Item	WMNSQ	t	Schwierigkeit	τ_1	SE $_{\tau}$	PMK	PBK $_0$	selektiert
Früchte	Au_Fr	0.98	0.0	-1.866			0.41	-0.40	Au_Fr
Fischherz	K_FH_A	1.05	0.3	-1.350			0.20*	/	
	K_FH_BC	1.22	0.9	-0.309	-1.354	0.404	0.29	/	K_FH_BC
Aal	K_Aa_A	1.07	0.5	0.856			0.20*	/	
	K_Aa_BC	0.78	-0.9	0.923	-0.985	0.410	0.62	/	
Transpiration	K_Tr_S	1.50*	1.4	-1.154	0.175	0.505	0.35	-0.46	*K_Tr_S
Pflanzenwuchs	K_PW_T	1.11	0.4	-3.14*			0.02*	/	
	K_PW_A	1.03	0.3	-4.008*			0.03*	/	
	K_PW_BC	1.03	0.2	-2.314	-0.879	0.469	0.52	/	K_PW_BC
	K_PW_S	0.83	-0.6	-0.916	1.405*	0.740	0.60	-0.45	*K_PW_S
Aktivität	K_Ak_S	1.18	0.7	-1.032	0.466	0.552	0.16*	/	
Schnecke	K_Sc_L	1.10	0.5	-1.157			0.34	/	K_Sc_L
Rentiere	K_Re_W	1.05	0.5	0.562	0.140	0.196	0.53	-0.18	K_Re_W
	K_Re_X	1.10	1.5	-0.302			0.36	-0.22	K_Re_X
Asseln	K_As_W	0.85	-0.5	-1.304			0.58	-0.78	K_As_W
	K_As_X	0.97	0.0	-1.545			0.58	-0.78	K_As_X
Paramecium	K_Pr_T	0.68*	-2.3*	-0.729			0.74	-0.63	
	K_Pr_A	0.65*	-2.5*	-0.729			0.83	-0.84	
	K_Pr_BC	0.93	-0.2	1.422	-1.330	0.431	0.72	-0.85	K_Pr_BC
	K_Pr_S	0.75*	-1.2	-0.210	1.469	0.739	0.80	-0.75	*K_Pr_S
	K_Pr_L	0.65*	-1.7	0.008	0.996	0.618	0.79	-0.80	*K_Pr_L
Kiefernspinner	K_Ki_T	0.73*	-1.9	-0.506			0.84	-0.72	*K_Ki_T
	K_Ki_A	0.76	-1.7	-0.506			0.75	-0.70	K_Ki_A
	K_Ki_BC	0.79	-1.0	0.301	1.515*	0.738	0.78	-0.76	*K_Ki_BC
	K_Ki_S	1.32	1.3	0.432	1.497*	0.739	0.60	-0.76	*K_Ki_S
	K_Ki_W	0.81	-0.9	0.115	0.469	0.501	0.87	-0.76	K_Ki_W
	K_Ki_X	0.76	-1.0	1.201			0.70	-0.76	K_Ki_X
Eier	KS_F_BC	0.96	0.0	1.803			0.51	-0.42	KS_F_BC
Ozon	KS_O_BC	1.04	0.2	-0.030	0.025	0.446	0.54	/	KS_O_BC
	KS_O_S	1.25	1.7	-0.137			0.04*	/	
Geschlechtsreife	KS_G_T	0.99	0.0	0.352			0.47	-0.68	KS_G_T
	KS_G_A	0.91	-0.4	-0.073			0.75	-0.44	KS_G_A
	KS_G_BC	1.22	0.9	1.842	-1.669	0.423	0.46	-0.52	KS_G_BC
	KS_G_S	1.25	1.0	-0.053	-1.150	0.408	0.58	-0.52	KS_G_S
Winterschlaf	KS_W_T	1.00	0.1	-0.684			0.51	-0.71	KS_W_T
	KS_W_A	0.72*	-0.3	-3.023*			0.80	-0.80	
	KS_W_BC	1.22	1.0	-0.208	-0.325	0.418	0.49	-0.45	KS_W_BC
	KS_W_S	0.96	0.0	-1.733	-0.188	0.511	0.66	-0.80	KS_W_S
	KS_W_W	0.81	-1.4	0.591			0.45	-0.71	KS_W_W
	KS_W_X	0.94	-0.5	-0.137			0.45	-0.71	KS_W_X
Blütendurchmesser	KS_B_T	0.96	0.1	-2.415			0.49	-0.57	KS_B_T
	KS_B_A	1.08	0.4	-3.764*			0.57	-0.57	**KS_B_A
	KS_B_BC	1.13	0.6	-0.079	0.088	0.478	0.66	-0.57	KS_B_BC
	KS_B_S	0.79	-0.8	-0.878	-0.511	0.445	0.58	-0.57	KS_B_S
	KS_B_W	1.09	0.7	-0.287			0.52	-0.57	KS_B_W
	KS_B_X	0.94	0.0	-2.019			0.41	-0.57	KS_B_X

WMNSQ = weighted mean square, t = Wert des zugehörigen t-Tests, Schwierigkeit = Itemparameter, PMK = Produkt-Moment-Korrelation (Trennschärfe), PBK $_0$ = höchste signifikante punkt-biseriale Korrelation einer 0-Rubrik. Erklärungen der Markierungen durch * im Text.

Tabelle 12.4: Itemkennwerte der Integration.

Diagramm	Item	WMNSQ	t	Schwierigkeit	τ_1	SE $_{\tau}$	PMK	PBK $_0$	selektiert
Räuber-Beute	I_RB_a	1.09	0.5	-1.031			0.59	-0.55	I_RB_a
	I_RB_b	1.12	0.6	-0.028			0.69	-0.57	I_RB_b
	I_RB_c	0.71*	-0.9	0.729			0.81	-0.61	*I_RB_c
	I_RB_d	0.86	-0.4	0.471			0.80	-0.61	I_RB_d
	I_RB_e	0.85	-0.4	0.997			0.72	-0.59	I_RB_e
Photosynthese	I_PS_a	1.35*	1.3	1.312			0.62	-0.68	*I_PS_a
	I_PS_b	0.54*	-1.9	-0.704			0.84	-0.76	*I_PS_b
	I_PS_c	0.58*	-2.0*	0.297			0.80	-0.77	
Paramecium1	I_P1_S	1.38*	1.4	-0.377	-1.221	0.431	0.54	/	
Paramecium2	I_P2a	1.00	0.1	-1.217			0.59	/	I_P2a
	I_P2b	0.62*	-1.9	0.514			0.83	-0.56	*I_P2b
	I_P2c	0.59*	-2.0*	0.275			0.83	-0.77	**I_P2c
Paramecium3	I_P3_L	1.24	0.8	2.187			0.23*	/	
	I_P3_X	0.69*	-1.0	1.555	-0.205	0.530	0.77	/	*I_P3_X
	I_P3_W	1.20	0.7	1.644	-0.987	0.472	0.66	0.00	I_P3_W
Gleichwarm	I_GW_L	1.17	0.6	2.480			0.21*	/	
	I_GW_W	1.04	0.2	2.125			0.45	0.39*	
	I_GW_X	0.99	0.0	1.102	-0.506	0.469	0.80	-0.75	I_GW_X
Kälte	I_Ke_W	1.23	0.7	2.253			0.52	-0.56	I_Ke_W
	I_Ke_X	1.01	0.1	0.220			0.63	-0.56	I_Ke_X
Bakterien	I_Bak	1.20	0.5	4.459*			0.27	0.57*	
Flusskrebis	I_FK_X	0.94	-0.2	-0.192			0.49	-0.40	I_FK_X
Papageitaucher	IS_T_a	1.38*	0.9	-3.170			0.58	-0.73	*IS_T_a
	IS_T_b	0.75*	-0.5	-2.136			0.86	-0.85	*IS_T_b
	IS_T_c	0.44*	-1.6	-2.136			0.86	-0.91	*IS_T_c
	IS_T_d	0.96	0.0	-2.607			0.79	-0.85	IS_T_d
	IS_T_e	0.40*	-1.7	-2.136			0.86	-0.91	*IS_T_e
	IS_T_f	0.97	-0.1	0.754			0.64	-0.88	IS_T_f
Ozon	IS_O_L	1.11	0.5	0.160	-0.413	0.462	0.72	-0.66	IS_O_L

WMNSQ = weighted mean square, t = Wert des zugehörigen t-Tests, Schwierigkeit = Itemparameter, PMK = Produkt-Moment-Korrelation (Trennschärfe), PBK $_0$ = höchste signifikante punkt-biseriale Korrelation einer 0-Rubrik. Erklärungen der Markierungen durch * im Text.

12.4.2 Pilotierung: Items zu den Lernermerkmalen

Die Items zu den vier Lernermerkmalen biologisches Vorwissen, Lesekompetenz, Rechenfähigkeiten und Umgang mit abstrakten Diagrammen wurden wie die Items zur Diagrammkompetenz in der Pilotierung erprobt und einzelnen Rasch-Skalierungen unterzogen. Die Kennwerte sind in der folgenden Tabelle dargestellt.

Tabelle 12.5: Itemkennwerte der Items zu den Lernermerkmalen.

Lernermerkmal	Item	WMNSQ	t	Schwierigkeit	τ_1	SE $_{\tau}$	PMK	PBK $_0$	selektiert
Vorwissen Biologie	<i>rb1</i>	1.26	1.0	-1.388			0.40	-0.40	<i>rb1</i>
	<i>rb2</i>	1.22	1.1	0.477			0.42	-0.40	<i>rb2</i>
	<i>pop1</i>	0.81	-0.7	-1.388			0.58	-0.58	<i>pop1</i>
	<i>pop2</i>	1.14	0.8	0.276			0.31	/	<i>pop2</i>
	<i>ps1</i>	0.95	-0.1	-1.656			0.45	-0.41	<i>ps1</i>
	<i>ps2</i>	1.09	0.4	-1.656			0.35	/	<i>ps2</i>
	<i>af1</i>	0.88	-0.6	0.080			0.79	-0.59	<i>af1</i>
	<i>af2bsp</i>	1.14	0.5	-1.963			0.29	/	<i>af2bsp</i>
	<i>konk1</i>	1.06	0.4	-0.114			0.69	-0.63	<i>konk1</i>
	<i>konk2</i>	1.08	0.3	-1.963			0.29	/	<i>konk2</i>
Lesekompetenz	<i>LKa</i>	3.21*	1.6	-7.058*			0.65	-0.45	
	<i>LKb</i>	0.97	0.1	0.223			0.91	-0.45	<i>LKb</i>
Rechenfähigkeiten	<i>Re1</i>	0.94	0.1	-2.899			0.41	/	<i>Re1</i>
	<i>Re2</i>	1.07	0.3	-1.691			0.55	-0.55	<i>Re2</i>
	<i>Re3</i>	0.82	-0.3	-2.394			0.70	-0.70	<i>Re3</i>
	<i>Re4</i>	1.06	0.4	0.300			0.62	-0.55	<i>Re4</i>
	<i>Re5</i>	1.09	0.3	-2.394			0.46	-0.41	<i>Re5</i>
Umgang mit abstrakten Diagrammen	<i>dka_au</i>	0.99	0.3	-3.436*			-0.13*	/	
	<i>dka1</i>	1.00	0.1	0.223	-2.363	0.504	0.40	/	<i>dka1</i>
	<i>dka2</i>	0.96	0.1	-2.675			0.28	/	<i>dka2</i>
	<i>dka3</i>	0.92	-0.1	-2.201			0.44	/	<i>dka3</i>
	<i>dka4</i>	0.95	-0.1	-1.293			0.31	/	<i>dka4</i>
	<i>dka5</i>	1.07	0.3	-2.201			0.06*	/	
	<i>dka6</i>	0.97	-0.2	-0.650			0.41	/	<i>dka6</i>
	<i>dka7</i>	0.80	-0.4	-2.201			0.49	-0.51	<i>dka7</i>
	<i>dka8</i>	1.15	0.9	-0.850			0.28	/	<i>dka8</i>
	<i>dka9</i>	0.97	-0.1	-0.650			0.35	/	<i>dka9</i>
	<i>dka10</i>	0.82	-1.5	-0.091			0.59	/	<i>dka10</i>
	<i>dka11</i>	0.87	-1.0	-0.459			0.55	-0.41	<i>dka11</i>
	<i>dka12</i>	0.83	-0.4	-1.844			0.50	-0.45	<i>dka12</i>
	<i>dka13</i>	0.89	-0.6	-0.850			0.61	-0.58	<i>dka13</i>
	<i>S_dka1</i>	0.96	0.2	-3.436*			-0.05*	/	
	<i>S_dka2</i>	0.97	0.3	-3.436*			-0.13*	/	
<i>S_dka3</i>	1.12	0.5	1.293			0.39	-0.49	<i>S_dka3</i>	
<i>S_dka4</i>	1.00	0.0	0.459			0.42	-0.51	<i>S_dka4</i>	

WMNSQ = weighted mean square, t = Wert des zugehörigen t-Tests, Schwierigkeit = Itemparameter, PMK = Produkt-Moment-Korrelation (Trennschärfe), PBK $_0$ = höchste signifikante punkt-biseriale Korrelation einer 0-Rubrik. Erklärungen der Markierungen durch * im Text zu Tabelle 12.2.

Bezogen auf das biologische Vorwissen lagen die Kennwerte aller Items in gültigen Bereichen. Die Itemauswahl fand deshalb hauptsächlich nach inhaltlichen Kriterien statt. Beispielsweise wurden die Items zur Photosynthese (*ps1* und *ps2*) entfernt, da die entsprechenden Items zur Diagrammkompetenz, die sich inhaltlich auf die Photosynthese bezogen, nicht für den Haupttest ausgewählt wurden. Insgesamt wurden ein Item zum Lotka-Volterra-Gesetz (*rb2*), zwei Items zum Populationsbegriff (*pop1* und *pop2*) sowie ein Item zum Konkurrenzbegriff (*konk1*) für den Haupttest beibehalten.

Von den zwei Items zur Lesekompetenz erwies sich nur eines als geeignet (*LKb*). Es wurde für den Haupttest übernommen. Hinsichtlich der Rechenfähigkeiten wurde das Items *Re4* ausgewählt, weil es mit Abstand die höchste Schwierigkeit bei hoher Trennschärfe aufwies.

Hinsichtlich des Umgangs mit abstrakten Diagrammen wurden vier Items eliminiert, die eine zu geringe Trennschärfe bzw. drei davon eine zu geringe Schwierigkeit zeigten. Die weitere Auswahl erfolgte nach inhaltlichen Abwägungen.

12.4.3 Haupttestung: Items zur Diagrammkompetenz

Tabelle 12.6: Itemkennwerte der Items zur Diagrammkompetenz im Haupttest.

Diagramm	Item	WMNSQ	t	Schwierigkeit	τ_1	SE $_{\tau}$	PMK
Maus	Id_M_a	1,13	1,2	-2,008			0,11*
	Id_M_b	1,01	0,1	-2,311			0,24
	Id_M_c	1,20	3,8*	0,639			0,16*
Rentiere	Id_Re	1,00	0,1	-0,971			0,37
Gewebe	IdS_Ga	1,08	1,4	0,761			0,28
	IdS_Gb	1,02	0,5	0,588			0,34
Läuse-(Käfer)	A_LK_b	1,10	1,7	-0,803			0,26
	A_La_l	1,13	2,8*	0,370			0,28
	A_LK_g	1,14	2,7*	-0,469			0,28
	A_LK_h	0,89	-0,8	-2,348			0,36
	A_LK_i	1,11	2,3*	-0,367			0,30
	A_LK_k	1,10	2,1*	0,222			0,30
	A_LK_m	1,04	0,9	-0,334			0,37
Rentiere	A_Re	1,05	1,0	0,537			0,35
Paramecium2	A_P2c1	0,79	-5,1*	0,173			0,63
	A_P2c2	0,80	-4,9*	0,255			0,61
	A_P2c3	0,84	-3,9*	0,092			0,56
	A_P2c4	0,93	-1,3	0,691			0,43
Hund	A_Hu_a	1,16	2,6*	0,959			0,19*
	A_Hu_b	1,04	0,4	2,062			0,28
Papagei-taucher	AS_P_e	0,90	-1,0	-1,947			0,45
	AS_P_g	0,97	-0,5	-1,049			0,42
Rentiere	K_Re_W	1,17	1,9	1,255	0,082	0,151	0,36
	K_Re_X	1,01	0,3	-0,351			0,39
Kiefernspinner	K_Ki_T	0,77	*-3,8*	-1,089			0,64
	K_Ki_A	0,84	-2,8*	-0,877			0,57
	K_Ki_BC	0,94	-0,9	0,007	0,091	0,136	0,61
	K_Ki_S	0,98	-0,3	-0,346	0,506	0,154*	0,60
	K_Ki_W	1,11	1,0	1,529	0,05	0,159	0,33
	K_Ki_X	0,97	-0,7	-0,006			0,46

WMNSQ = weighted mean square, t = Wert des zugehörigen t-Tests, Schwierigkeit = Itemparameter, PMK = Produkt-Moment-Korrelation (Trennschärfe), PBK₀ = höchste signifikante punkt-biseriale Korrelation einer 0-Rubrik. Erklärungen der Markierungen durch * im Text zu Tabelle 12.2.

Fortsetzung der Tabelle auf der folgenden Seite.

Diagramm	Item	WMNSQ	t	Schwierigkeit	τ_1	SE $_{\tau}$	PMK
<i>Blütendurchmesser</i>	<i>KS_B_T</i>	1,07	1,6	-0,022			0,32
	<i>KS_B_A</i>	0,83	-1,8	-1,860			0,52
	<i>KS_B_BC</i>	1,10	1,5	0,166	0,165	0,139*	0,50
	<i>KS_B_S</i>	1,11	1,5	-0,646	-0,591	0,124	0,45
	<i>KS_B_W</i>	1,08	1,2	1,131			0,26
	<i>KS_B_X</i>	1,10	2,2*	0,386			0,24
<i>Geschlechtsreife</i>	<i>KS_G_T</i>	0,94	-1,0	0,922			0,46
	<i>KS_G_A</i>	0,77	-4,9*	-0,452			0,65
	<i>KS_G_BC</i>	0,89	-1,4	1,469	-0,726	0,132	0,56
	<i>KS_G_S</i>	0,84	-2,4*	0,481	-0,596	0,123	0,65
<i>Paramecium2</i>	<i>I_P2c</i>	1,09	0,5	2,085	1,118	0,277*	0,27
<i>Kälte</i>	<i>I_Ke_W</i>	0,99	-0,2	1,313			0,36
	<i>I_Ke_X</i>	1,01	0,2	0,622			0,39
<i>Papageitaucher</i>	<i>IS_T_a</i>	0,89	-1,4	-1,384			0,48
	<i>IS_T_b</i>	0,94	-1,3	-0,087			0,46
	<i>IS_T_e</i>	0,94	-0,9	-0,933			0,46
	<i>IS_T_f</i>	1,10	1,5	1,111			0,27

WMNSQ = weighted mean square, t = Wert des zugehörigen t-Tests, Schwierigkeit = Itemparameter, PMK = Produkt-Moment-Korrelation (Trennschärfe), PBK_0 = höchste signifikante punkt-biseriale Korrelation einer 0-Rubrik. Erklärungen der Markierungen durch * im Text zu Tabelle 12.2.

Einige der Items weisen kritische Werte auf. Da aber über alle Items hinweg die Reliabilität des Tests ausreichend hoch war (WLE-Reliabilität: Rel. = .890, EAP/PV-Reliabilität: Rel. = .966), wurden auch die kritischen Items in die Analysen einbezogen.

12.5 Ergänzungen zur Ergebnisdarstellung

Tabelle 12.7: Regressionsanalyse zur Bestimmung der Effektstärke des Zusammenhang von Informationsentnahme und Konstruktion.

Prädiktor	B	SE _B	β
Konstruktion	.316	.039	.435***

$R^2_{\text{korr.}} = .186, F(1,289) = 66.874, p < .001$

Tabelle 12.8: Regressionsanalysen zur Mediationswirkung der Informationsentnahme auf den Zusammenhang zwischen Konstruktion und Integration.

1. einfache Regression mit Konstruktion als Prädiktor für Integration.

Prädiktor	B	SE _B	β
Konstruktion	.399	.066	.334***

$R^2_{\text{korr.}} = .108, F(1,289) = 35.986, p < .001$

2. Multiple Regression mit Konstruktion und Informationsentnahme als Prädiktoren für Integration.

Prädiktor	B	SE _B	β
Konstruktion	.103	.062	.086
Informationsentnahme	.936	.085	.570***

$R^2_{\text{korr.}} = .371, F(2,289) = 85.841, p < .001$

3. Einfache Regression mit Konstruktion als Prädiktor für Informationsentnahme.

Prädiktor	B	SE _B	β
Konstruktion	.316	.039	.435***

$R^2_{\text{korr.}} = .186, F(1,289) = 66.874, p < .001$

Tabelle 12.9: Regressionsanalysen zur Mediationswirkung des Ablesens auf den Zusammenhang zwischen Identifizierung und Integration.

1. Einfache Regression mit Identifizierung als Prädiktor für Integration.

Prädiktor	B	SE _B	β
Identifizierung	.482	.082	.329***
$R^2_{\text{korr.}} = .105, F(1,289) = 34.829, p < .001$			

2. Multiple Regression mit Identifizierung und Ablesen als Prädiktoren für Integration.

Prädiktor	B	SE _B	β
Identifizierung	.148	.075	.101*
Ablesen	.782	.072	.560***
$R^2_{\text{korr.}} = .366, F(2,289) = 84.104, p < .001$			

3. Einfache Regression mit Identifizierung als Prädiktor für Ablesen.

Prädiktor	B	SE _B	β
Identifizierung	.427	.057	.406***
$R^2_{\text{korr.}} = .162, F(1,289) = 56.741, p < .001$			

Tabelle 12.10: Regressionsanalysen zur Mediationswirkung des Ablesens auf den Zusammenhang zwischen Aufbau und Integration.

1. Einfache Regression mit Aufbau als Prädiktor für Integration.

Prädiktor	B	SE _B	β
Aufbau	.377	.060	.349***
$R^2_{\text{korr.}} = .119, F(1,289) = 39.851, p < .001$			

2. Multiple Regression mit Aufbau und Ablesen als Prädiktoren für Integration.

Prädiktor	B	SE _B	β
Aufbau	.128	.056	.118*
Ablesen	.770	.072	.552***
$R^2_{\text{korr.}} = .369, F(2,289) = 85.191, p < .001$			

3. Einfache Regression mit Aufbau als Prädiktor für Ablesen.

Prädiktor	B	SE _B	β
Aufbau	.323	.041	.418***
$R^2_{\text{korr.}} = .172, F(1,289) = 60.785, p < .001$			

Tabelle 12.11: Regressionsanalysen zur Mediationswirkung des AbleSENS auf den Zusammenhang zwischen Dateneintrag und Integration.

1. Einfache Regression mit Dateneintrag als Prädiktor für Integration.

Prädiktor	B	SE _B	β
Dateneintrag	.325	.087	.214***
$R^2_{\text{kor.}} = .043, F(1,289) = 13.827, p < .001$			

2. Multiple Regression mit Dateneintrag und AbleSEN als Prädiktoren für Integration.

Prädiktor	B	SE _B	β
Dateneintrag	.069	.075	.045
AbleSEN	.821	.069	.588***
$R^2_{\text{kor.}} = .359, F(2,289) = 84.719, p < .001$			

3. Einfache Regression mit Dateneintrag als Prädiktor für AbleSEN.

Prädiktor	B	SE _B	β
Dateneintrag	.312	.061	.287***
$R^2_{\text{kor.}} = .079, F(1,289) = 25.802, p < .001$			

Tabelle 12.12: Modellvergleiche der Gesamtmodelle, die jeweils die Aspekte einer Komponente differenzieren.

Modell	Zahl der geschätzten Parameter	AIC _c	Differenz AIC _c	BIC	Differenz BIC	-2 log L
3dim Gesamt	62	16219	-	16412	-	16060.28086
2+6dimK	92	16220	-1	16470	-58	15948.41943
2+2dimInt	66	16225	-6	16427	-16	16053.34608
2+5dimIE	84	16246	-27	16483	-72	16007.51851
13dim Gesamt	147	16458	-239	16688	-277	15855.30020

Tabelle 12.13: Deskriptive Statistik zu den Itemschwierigkeiten in den einzelnen Aspekten der Konstruktion.

Konstruktionsaspekt	N	M	SD	Median	Min	Max	Spannweite
Auswahl des Typs	3	-0.04	1.13	0.02	1.05	-1.20	2.25
Achsenbelegung	3	-1.18	0.85	-0.95	-0.47	-2.12	1.65
Achsenbeschriftung	3	0.67	0.89	0.26	1.70	0.07	1.63
Skalierung	3	-0.17	0.69	-0.35	0.60	-0.75	1.35
Werteintrag	3	1.52	0.26	1.48	1.79	1.28	0.51
Skizzierung	3	0.05	0.41	0.04	0.47	-0.35	0.82

Tabelle 12.14: Deskriptive Statistik der Itemschwierigkeiten hinsichtlich der zwei Aspekte der Integration.

Integrations- aspekt	N	M	SD	Median	Min	Max	Spannweite
<i>Resultat Text</i>	4	-0,07	1,98	-0,64	-1,73	2,72	4,45
<i>Resultat Diagramm</i>	3	1,21	0,42	1,32	0,74	1,56	0,82
Insgesamt	7	0,48	1,58	0,74	-1,73	2,72	4,45

Tabelle 12.15: Deskriptive Statistik der Itemschwierigkeiten in den einzelnen Komponenten des *3dim*-Modells nach Diagrammtypen differenziert. LD: Liniendiagramm, SD: Säulendiagramm.

Komponente (3im)	Diagramm- typ	N	M	SD	Median	Min	Max	Spann- weite
Informations- entnahme	LD	18	-0,18	1,16	0,16	-2,32	2,08	4,41
	SD	4	-0,39	1,31	-0,21	-1,92	0,78	2,71
	Gesamt	22	-0,21	1,16	0,16	-2,32	2,08	4,41
Konstruktion	LD	8	0,04	0,93	-0,15	-1,06	1,55	2,61
	SD	10	0,18	0,97	0,30	-1,83	1,49	3,32
	Gesamt	18	0,12	0,93	0,03	-1,83	1,55	3,38
Integration	LD	3	1,36	0,73	1,34	0,65	2,10	1,46
	SD	4	-0,30	1,10	-0,49	-1,36	1,13	2,49
	Gesamt	7	0,41	1,25	0,65	-1,36	2,10	3,46

Tabelle 12.16: Deskriptive Statistik der Itemschwierigkeiten in den einzelnen Komponenten des *5dim*-Modells nach Diagrammtypen differenziert. LD: Liniendiagramm, SD: Säulendiagramm.

Komponente (5dim)	Diagramm- typ	N	M	SD	Median	Min	Max	Spann- weite
Identifizierung	LD	4	-1,14	1,33	-1,46	-2,28	0,66	2,95
	SD	2	0,70	0,12	0,70	0,61	0,78	0,17
	Gesamt	6	-0,53	1,40	-0,17	-2,28	0,78	3,07
Ablesen	LD	14	0,10	0,99	0,22	-2,32	2,08	4,41
	SD	2	-1,47	0,63	-1,47	-1,92	-1,02	0,90
	Gesamt	16	-0,10	1,08	0,16	-2,32	2,08	4,41
Aufbau	LD	4	-0,55	0,50	-0,59	-1,06	0,03	1,10
	SD	8	0,03	1,02	0,10	-1,83	1,49	3,32
	Gesamt	12	-0,16	0,90	-0,16	-1,83	1,49	3,32
Dateneintrag	LD	4	0,63	0,92	0,65	-0,33	1,55	1,87
	SD	2	0,78	0,52	0,78	0,41	1,15	0,74
	Gesamt	6	0,68	0,76	0,78	-0,33	1,55	1,87
Integration	LD	3	1,36	0,73	1,34	0,65	2,10	1,46
	SD	4	-0,30	1,10	-0,49	-1,36	1,13	2,49
	Gesamt	7	0,41	1,25	0,65	-1,36	2,10	3,46

Tabelle 12.17: Deskriptive Statistik der Personenfähigkeiten in den Komponenten des Modells *3dim* getrennt nach Geschlecht. m = Schüler, w = Schülerinnen.

Komponente (<i>3dim</i>)	Geschlecht	N	M	SD	Median	Min	Max	Spannweite
Informations-entnahme	m	139	0,38	1,13	0,27	-2,65	3,18	5,83
	w	149	0,13	1,01	0,05	-3,29	3,18	6,47
Konstruktion	m	139	-0,29	1,47	-0,02	-4,26	2,65	6,91
	w	149	-0,14	1,51	0,32	-4,26	2,26	6,53
Integration	m	139	-0,36	1,77	-0,33	-3,71	3,35	7,06
	w	149	-0,63	1,78	-0,33	-3,71	2,54	6,25

Tabelle 12.18: Deskriptive Statistik der Personenfähigkeiten in den Komponenten des Modells *5dim* getrennt nach Klassenstufen. m = Schüler, w = Schülerinnen.

Komponente (<i>5dim</i>)	Geschlecht	N	M	SD	Median	Min	Max	Spannweite
Identifizierung	m	139	0,62	1,25	0,91	-3,32	3,20	6,51
	w	149	0,42	1,15	0,05	-3,32	3,20	6,51
Ablesen	m	139	0,28	1,31	0,33	-2,91	4,10	7,02
	w	149	0,00	1,22	0,03	-4,21	2,83	7,04
Aufbau	m	139	0,03	1,62	0,36	-3,91	4,35	8,26
	w	149	0,17	1,67	0,59	-3,91	3,02	6,93
Dateneintrag	m	139	-0,71	1,15	-0,69	-2,80	1,15	3,95
	w	149	-0,54	1,17	-0,69	-2,80	2,68	5,48
Integration	m	139	-0,36	1,78	-0,34	-3,72	3,38	7,10
	w	149	-0,63	1,78	-0,34	-3,72	2,57	6,28

Tabelle 12.19: Referenzzahlen für die Häufigkeiten der auftretenden Fehler bei der Konstruktion.

Konstruktionsaspekt	Diagramm	Zahl der Bearbeitungen
Auswahl des Diagrammtyps	<i>Kiefernspinner</i>	216
	<i>Geschlechtsreife</i>	242
	<i>Blütendurchmesser</i>	144
Achsenbelegung	<i>Kiefernspinner</i>	219
	<i>Kiefernspinner Pilotierung</i>	22
	<i>Geschlechtsreife</i>	241
	<i>Blütendurchmesser</i>	175
Achsenbeschriftung	<i>Kiefernspinner</i>	205
	<i>Geschlechtsreife</i>	203
	<i>Blütendurchmesser</i>	170
Skalierung	<i>Kiefernspinner</i>	218
	<i>Geschlechtsreife</i>	247
	<i>Blütendurchmesser</i>	177
Werteintrag	<i>Kiefernspinner</i>	213 / 190*
	<i>Blütendurchmesser</i>	238 / 158*
	<i>Rentiere</i>	249 / 216*

Hinweis: Die durch * markierten Zahlen stehen für die Anzahl der Bearbeitungen, die nicht die volle Punktzahl erhalten haben.

Danksagung

Zunächst sei den Schülerinnen und Schülern gedankt, die sorgfältig und angestrengt die Tests bearbeitet haben! Ohne sie wäre diese Arbeit nicht möglich gewesen.

Mein aufrichtiger Dank gilt meinen Betreuern, Herrn Prof. Dr. Helmut Prechtl und Frau Prof. Dr. Claudia Nerdel, zum einen für ihre Gelassenheit, die mir die Freiheit gab, meinen eigenen Gedanken und Ideen zu folgen, zum anderen für ihren Druck in den richtigen Momenten sowie Ansporn zum Publizieren. Sie haben mich nicht nur durch ihre konstruktive Kritik, sondern auch durch ihre Bereitschaft, sich mit unzähligen Schildlaus-Aufgaben auseinanderzusetzen, bei dieser Arbeit unterstützt!

Prof. Dr. Horst Bayrhuber und Prof. Dr. Ute Harms danke ich dafür, dass sie mir durch sehr flexible Arbeitszeiten die Möglichkeit gegeben haben, meine produktivsten Tagesphasen auszunutzen.

Meinen herzlichen Dank möchte ich Dr. Markus Lücken für die umfangreiche methodische Beratung aussprechen, die sich so oft aus einer *ganz kleinen Frage* ergeben hat. Zudem danke ich Steffen Brandt für die Anleitung beim Umgang mit ConQuest sowie Dr. Oliver Walter für die Beantwortung meiner Fragen im Bereich der probabilistischen Testauswertung.

Für die Hilfe bei der Dateneingabe möchte ich mich bei Filia Tzanidakis, Jan-Hendrik Bake, Truels Dentler und Sandra Nitz bedanken, die es zu Tages- und Nachtzeiten auf sich genommen haben, stundenlang nur Zahlencodes zu sehen und zu hören!

Für Fehlersuche und -funde bei Durchsicht der Texte sei Jan-Hendrik Bake, Maike Ehmer, Dr. Ingrid Glowinski, Gesa Kramer, Dr. Markus Lücken, Dr. Iris Mackensen-Friedrichs, Sandra Nitz und Dr. Cornelia Sommer herzlich gedankt!

Meinen Schreibtischnachbarinnen Gesa und Maike sowie meinen Flurnachbarn Burkhard, Markus und Sandra danke ich für die diversen kleinen Prokrastinationshilfen, die so viele nette Momente während der Arbeit ermöglicht haben! Mein Dank gilt auch den Anhängern der Mittagsrunde für die Momente, in denen sie mir durch unterhaltsame Spaziergänge zur Mensa sowie durch wissenschaftliche und nichtwissenschaftliche *kognitive Desserts* zu einer wohltuenden gedanklichen Pause vom Schreiben verholfen haben.

Ein ganz besonderes Dankeschön geht an Maike für so Vieles und insbesondere für ihr offenes Ohr trotz eigener Anspannung. Sie hat mich durch ihre Empathie besonders in den letzten Monaten emotional aufgefangen!

Mein tiefer Dank gilt Jan-Hendrik nicht nur für die anregenden Diskussionen über und um diese Arbeit herum. Vor allem möchte ich mich für seinen Trost und seine Aufmunterung in zweifelnden und unsicheren Momenten und für die Geborgenheit bedanken, aus der ich so viel Kraft und Anreiz zur Fertigstellung dieser Arbeit geschöpft habe!

Herzlich bedanken möchte ich mich nicht zuletzt bei meinen Eltern, Brigitte und Manfred Lachmayer, und meiner Großmutter Waltraut Giese, die mich durch ihre Anerkennung und ihre Liebe gestärkt und mir Halt gegeben haben!

Lebenslauf: Simone Lachmayer

Persönliche Daten

- Geburtsdatum 17.04.1980
- Geburtsort Augsburg
- Staatsangehörigkeit deutsch
- Wohnort Schwefelstr. 9, 24118 Kiel

Schulbildung und Studium

- Aug. 1986 – Juni 1990 Besuch der Grundschule Heikendorf
- Aug. 1990 – Juni 1999 Besuch des Heinrich-Heine-Gymnasiums Heikendorf
- Juni 1999 Abitur
- Okt. 1999 – Jan. 2005 Studium der Biologie und Mathematik für das Lehramt an Gymnasien an der Christian-Albrechts-Universität zu Kiel
- Juni 2004 – Okt. 2004 Staatsexamensarbeit in der Abteilung Biologiedidaktik des Leibniz-Instituts für die Pädagogik der Naturwissenschaften (IPN) mit dem Titel „Auswirkungen des Einsatzes instruktionaler Begleitmaterialien auf den Erwerb von Faktenwissen und Verständnis beim computergestützten Lernen des Themas Olfaktion in der Oberstufe“
- 24. Jan. 2005 Erstes Staatsexamen für das Lehramt an Gymnasien in Biologie und Mathematik
- seit April 2005 Promotionsstudium der Biologiedidaktik an der CAU Kiel

Wissenschaftliche Tätigkeit

- April 2004 – April 2005 Arbeit als wissenschaftliche Hilfskraft am IPN
- seit Mai 2005 Arbeit als wissenschaftliche Angestellte in der Abteilung Biologiedidaktik des IPN

Kiel, 27. Mai 2008

Erklärung

Hiermit erkläre ich, dass die vorliegende Dissertation abgesehen von der Beratung durch meinen Betreuer, Herrn Prof. Dr. Helmut Prechtel, nach Inhalt und Form mein eigenes Werk ist. Bei ihrer Anfertigung habe ich keine anderen als die angegebenen Quellen und Hilfsmittel verwendet.

Ferner erkläre ich, dass ich zuvor noch keinen Promotionsversuch unternommen habe und dass die vorliegende Arbeit weder vollständig noch in Teilen an anderer Stelle im Rahmen eines Prüfungsverfahrens vorgelegt, veröffentlicht oder zur Veröffentlichung eingereicht worden ist.

Simone Lachmayer
27. Mai 2008